ERROR DISTRIBUTIONS AND PARAMETER ESTIMATION IN RAINFALL-RUNOFF MODELING

BY

APURBA KUMAR BORAH

Bachelor of Engineering
University of Udaipur
Rajasthan, India
1982

Master of Engineering
Asian Institute of Technology
Bangkok, Thailand
1985

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
May, 1990

ERROR DISTRIBUTIONS AND PARAMETER
ESTIMATION IN RAINFALL-RUNOFF
MODELING

Thesis Approved:

Thesis Advisor

Dean of the Graduate College

ACKNOWLEDGEMENTS

This study is concerned with the estimation of parameters in rainfall-runoff modeling. It was financed by the Oklahoma Water Resources Research Institute and the Oklahoma Agricultural Experiment Station. I am grateful to the Agricultural Engineering Department for providing me with a research assistantship.

I wish to extend my sincere thanks and deep appreciation to my major advisor, Regents Professor Dr. C. T. Haan for his guidance and assistance throughout the study. Without his help and encouragement, this study would have been never completed. Appreciation is also extended to the other committee members, Dr. Avdesh K. Tyagi, Dr. Bruce N. Wilson and Dr. Kevin Lansey for their valuable suggestions and cooperation.

I am grateful to my family for their understanding and encouragement during the course of this research. Finally, heartfelt appreciation is expressed to my wife Santwana for her love, encouragement and moral support.

TABLE OF CONTENTS

Chapter Page

LIST OF TABLES

LIST OF FIGURES

Figure | Page

CHAPTER I

INTRODUCTION

Statement of the Problem

Many hydrologic models have been developed in the last three decades. This rapid development has taken place with the advent of digital computers. Hydrologic models are being applied to a wide variety of decisions for both structural and non-structural designs in water resources systems. However, modeling is not without problems. A great deal of uncertainty exists in hydrologic modeling due to inherent variability, model error and parameter error.

Inherent variability is obvious because processes being modeled are stochastic in nature. Formulations of probability models for some of the hydrologic processes explain the presence of uncertainty in hydrologic models. Simplification of the complex hydrologic processes leads to model error. Inadequate structure of the model reflects our lack of complete knowledge about the various natural processes occurring in a watershed. Another source of error in modeling is parameter error. Uncertain parameters yield uncertain model predictions.

Parameter estimation is an important but difficult aspect of hydrologic modeling. The parameters of a model are not always directly related to measurable watershed characteristics. Therefore, they are estimated as a function of model prediction and observed data generally by a curve fitting procedure. Since model parameters are inferred as a function of observed data which are stochastic in nature, the parameter estimates should be treated as random variables (Haan,1989).

Uncertainty is introduced into parameters during the calibration process due to model structure, data used for estimation, objective function chosen, fitting criterion, parameter interaction and misspecification of the error model. This randomness or uncertainty in parameters can be characterized by probability density functions. Knowledge of parameter uncertainty can be used to estimate the uncertainty of the model predictions. Model predictions should be analyzed in a probabilistic manner instead of as point estimates.

Objectives

The objectives of this study were to :

1. Evaluate the impact of various error distributions in rainfall and streamflow on parameter estimates.
2. Evaluate parameter estimation techniques in the presence of errors in input data.

Scope of Study

The USGS Precipitation Runoff Modeling Systems (PRMS) was selected for this study. The model was run in the daily mode to yield daily runoff predictions. From preliminary studies, three of the more sensitive parameters SMAX, REMX and SC1 were selected. Four years of data (1974-1977) from Chickasha R-6 watershed (Oklahoma) were used for this study.

Initially the three parameters were optimized to find representative values for the watershed. This set of parameters was assumed to be the true values. Then four years of precipitation data were routed through PRMS to generate an error-free runoff sequence. Errors from a particular distribution were introduced separately into each value of rainfall and generated runoff records. Error distributions considered for this study were normal, lognormal, double exponential, and uniform distributions without correlation and normal distribution with correlated errors. Parameters were estimated using the three series of data sets : error-free rainfall and contaminated runoff, contaminated rainfall and error-free runoff, and both contaminated rainfall and runoff records. In each case, 100 different simulations were made. Estimation techniques employed for this study were the method of absolute errors, ordinary least squares and maximum likelihood techniques. The source of parameter uncertainty was evaluated from the spread of the parameters about their true value. Uncertainty in parameters was characterized by determining their probability density function. The Kolmogorov-Smirnov test was used to determine the significance of these distributions. Estimation techniques were evaluated according to their ability to estimate the parameters with low variations. These processes were repeated for each error distribution and parameter estimation technique.

CHAPTER II

REVIEW OF LITERATURE

Introduction

This research is concerned with parameter estimation in hydrologic modeling. Accordingly, this chapter deals with previous research on parameter estimation and its difficulties. Initially, a model classification is provided to formulate a model in perspective to its parameter estimation. Sources of parameter uncertainties are then discussed. Finally, a review of the state of the art in parameter estimation is provided. Theory on parameter estimation techniques is described in detail in the succeeding chapter.

Mathematical statements are used to formulate a problem in engineering systems. Generally, a solution can be obtained by mathematical techniques if the mathematical representation of the problem is well founded. However, all physical problems can not be completely expressed in mathematical statements due to a lack of understanding of these processes. Therefore, an abstract representation of the real system known as a 'model' is developed and an approximate solution of the real problem is sought. In the field of hydrology, three types of models have been used to represent hydrologic problems, namely : physical, analog and mathematical (Clarke,1973; Chow, Maidment and Mays,1988). The current trend is towards mathematical models based on theory and experiments. The above model classification with terms physical, analog and mathematical can not be taken literally. Mathematical expressions are also used extensively in both analog and physical models. Mathematical models are basically digital simulation models programmed on digital computers.

Physical Model

Physical models are scale models which represent the real system in a reduced scale such as hydraulic model of a spillway. There have been several attempts to build small laboratory scale models to study the rainfall-runoff process. Mamisao (1952), Amorocho and Orlob (1961), Chery (1966) and Lynch and Sopper (1974) were among those in surface water hydrology who utilized physical models to conduct studies of the response of a watershed. However, great difficulty has been encountered in transfering the results of those studies to natural watersheds.

Analog Model

Analog models use another system whose physical structure is quite different from that of the real system but whose response can be related to that of the real system. This task became easier with the development of analog computers. Riley, Chadwick and Israelson (1967) developed an electric analog computer to simulate the response of a small watershed in Southern Arizona. Analog models were specifically built for a given watershed and they were not transferable to other watersheds.

Mathematical Model

Mathematical models represent the behavior of the system in mathematical form by linking the input and output through a set of equations. They are widely used due to the availability of computers and diverse capabilities of modeling techniques. Mathematical models are classified in various ways. Some of the terms found in hydrologic literature are deterministic, stochastic, analytic, synthetic, empirical, parametric, conceptual, lumped, distributed, statistical, numerical, regression etc. The following model classification is used based on their concept of development.

Theoretical vs. Empirical

Theoretical models are entirely derived from basic physical laws such as conservation of mass, conservation of energy, laws of thermodynamics etc. There exists no completely theoretical operational model in hydrology. Empirical models are based on observation. Relationships between measured input and output are established through transformation functions. In general, transformation functions are not required to have any physical meaning. All hydrological models contain empirical relations. There is very little clear distinction between these two classes of models. Clarke (1973) defined them as conceptual and empirical. Sorooshian and Dracup (1978) classified them as synthetic vs. analytic in comparison to theoretical vs. empirical.

Deterministic vs. Stochastic

In the second level, models are classified based on randomness of the process. Stochastic models represent at least one of the processes by a probability distribution function. On the other hand, in deterministic models the random behavior of the variables is ignored and the process is considered to follow a definite law of certainty. Thus, with a given input, a deterministic model always produces the same output.

Lumped vs. Distributed

The third model subdivision is based on the distribution of model input variables and/or parameters. Distributed models consider the spatial variability of input variables or model parameters while lumped models use average values representing the entire system being modeled. Parameters used in lumped models are thus not a true measure of the underlying physical system. Such approximations limit the accuracy of model predictions. None of the existing rainfall-runoff models are completely distributed models. They are all lumped models to some extent with varying degrees of lumping.

Time-Variant vs. Time-Invariant

The time of input application affects the input-output relationship of time-variant system. The input-output relationship of time-invariant systems have no dependency on the time of input application. In hydrology, most of the systems are time-variant, but in most of the models, these variations are not considered for the sake of simplicity.

Model Formulation

Classification of hydrologic models has already been discussed. Regardless of how models are classified, they can generally be represented as (Haan,1989) :

$$\underline{O} = \underline{f}\,(\underline{I}, \underline{P}, t) + \underline{e} \qquad (1)$$

where $\underline{O}$ is a matrix of hydrologic response to be modeled, $\underline{f}$ is a collection of functional relationships, $\underline{I}$ is a matrix of inputs, $\underline{P}$ is a vector of parameters, t is time, and $\underline{e}$ is a matrix of errors. Response in $\underline{O}$ may range from a single number such as a peak flow or a runoff volume to a continuous record of flow, soil water content, evapotranspiration, and other quantities. Model classification depends on the nature of $\underline{f}$.

The distinction between $\underline{I}$ and $\underline{P}$ is not always clear. Generally $\underline{I}$ represents inputs, some of which are time varying such as rainfall, temperature, land use, etc. while $\underline{P}$ represents coefficients particular to a watershed that must be estimated by some means.

The error term, $\underline{e}$, represents the difference between what is actually observed, $\underline{O}$, and what the model predicts, $\underline{\hat{O}}$.

$$\underline{\hat{O}} = \underline{f}\,(\underline{I},\underline{P},\ t) \tag{2}$$

$$\underline{e} = \underline{O} - \underline{\hat{O}} \tag{3}$$

Thus $\underline{e}$ is a function of both $\underline{I}$ and $\underline{P}$. The vector $\underline{e}$ can be eliminated from the Eqn. 1, provided the following conditions are satisfied.

1. The $\underline{f}$ must be an exact representation of the watershed response.
2. The true values of the parameters $\underline{P}$ are known.
3. The independent variables are error-free.
4. The dependent variables can be expressed directly as a function of inputs $\underline{I}$ and the parameters $\underline{P}$.

None of the above conditions can be fully satisfied to achieve a perfect fit of the model. The stochastic part will always be present in a model. As already indicated, models contain empirical equations and thus can not be exact. It is impossible to know the exact value of the parameters because errors will be introduced into the parameters during measurement of observed data. The strict physical meaning of the parameters is lost due to the empirical nature of the model equations. Finally, it is not always possible to express dependent variables in terms of inputs and parameters in a deterministic form.

The major contributing sources of uncertainties in rainfall-runoff models are discussed in the following sections.

Precipitation Errors

The U.S. National Weather Service (NWS) collects and publishes precipitation data in the United States. There are about 12,000 locations over which rainfall data have been compiled (Riggs, 1985).

The most common error in precipitation is due to its spatial variation. Precipitation is usually measured from a few points in a large watershed and an average value is assumed to be representative of the entire basin. Often, measurements are available from a single gage and the gage may not be centrally located or may even be located some distance from the watershed. Erroneous precipitation records used for parameter estimation result in non-optimal parameters. Studying the effect of spatial variability of rainfall, Dawdy and Bergmann (1969) concluded that it has a substantial effect on the parameters of rainfall-runoff models. Aitken (1972) also reported the presence of random and systematic error in rainfall and runoff records. However, Kuczera (1982) found that bias due to these random errors in rainfall with a monthly coefficient of variation of 12 percent was of secondary importance in comparison to parameter uncertainty.

Another major source of error in precipitation records is due to malfunctioning of rain gages. Jackson and Aron (1971) attributed these rain gage errors to clock and/or weighing device malfunctions. Error caused by clock malfunction can be of significant importance, specifically in studies involving the temporal distribution of rainfall and runoff.

Evapotranspiration Errors

Evapotranspiration is calculated using pan evaporation or air temperature data. The U.S. National Weather Service collects these data in the United States. While calculating potential evapotranspiration, errors are introduced by empirical or semiempirical equations. However, the effect of these errors has been found not to be severe, especially during the periods of major storms (Ibbitt, 1972).

Streamflow Errors

The U.S. Geological Survey (USGS) is the responsible agency for collecting streamflow measurements in the United States. The U.S.D.A. also publishes streamflow records on their experimental watersheds. Jackson and Aron (1971) attributed the following causes for streamflow errors:

1. Instrument errors in the velocity measuring equipment.
2. Instrument errors in the stage measuring equipment.
3. Errors due to procedures used in measuring the velocity profile and average velocity.
4. Errors due to preparation of rating tables from a finite set of measurements.
5. Errors due to assuming steady flow past the gage site.

In traditional parameter estimation problems, uncertainties in the input vectors are neglected. It is assumed that mainly streamflow contributes to the error term. This may be a reasonable assumption concerning evapotranspiration, but assuming precipitation is measured without error is certainly questionable. This assumption has been made in a majority of nonlinear estimation problems, irrespective of its justification. Troutman (1982) showed how erroneous precipitation data results in biased parameter estimates in rainfall-runoff modeling. However, Ibbitt (1972) reported that the errors in precipitation data were taken care of by storage action of the model over several time intervals. His observation may be model specific. Since parameters of rainfall-runoff models are estimated through curve fitting techniques, errors in input precipitation are bound to affect the estimates. Studying the influence of precipitation errors on parameter estimates is one of the objectives of this research.

Parameter Interaction

Parameter interaction is another source of uncertainty in hydrologic modeling. In the presence of parameter interaction, it is difficult to obtain unique parameter estimates. Interdependence of parameters results in nonoptimal convergence of automatic search techniques (Johnston and Pilgrim, 1976; Sorooshian et. al., 1983). Parameter correlation can be studied by examining the behavior of the sum of squares error function (Mandeville et. al., 1970). With correlation between two parameters, this function gives a series of concentric circles on a plane. This indicates equal sensitivity in each parameter. When model parameters are interdependent, the resulting parameter response surface is elliptical rather than circular. The axis of the ellipse inclines to the parameter direction. A simple scale change is not enough to transform the elliptic pattern to circular contours because hydrologic models are highly nonlinear and the presence of threshold parameters create discontinuities in the error function surface. To overcome this problem, Gupta and Sorooshian (1983) recommended appropriate reparameterization of the model. This approach is doubtful for a conceptual rainfall-runoff model because these models are approximations of natural processes occurring in a watershed. Due to shortcomings of a model component, other components of the model are forced to compensate for the model inadequacy. Thus to some degree parameter interaction is always present. The method of principal components can be used to treat the collinearity problem in the regression model. Haan and Allen (1972) compared principal component and multiple regression in the context of discarding insensitive variables.

Parameter Estimation : The State of the Art

Parameter estimation in rainfall-runoff models is not a simple task. Traditional criterion of minimizing the sum of squares between observed and predicted values seem to be a straight forward numerical procedure. But difficulties encountered in parameter estimation in rainfall-runoff modeling has been cited extensively in hydrologic literature (Dawdy and Thompson,1966; Pickup,1973; Johnston and Pilgrim,1976 and others).

Dawdy and O'Donnell (1965) were among the earliest researchers to apply an automatic technique for evaluating parameters of a conceptual model. The fitting technique was an iterative trial and error search method developed by Rosenbrock (1960). The least squares criterion was used as the objective function. They evaluated the fitting technique based on its ability to obtain true parameter values starting from erroneous initial parameter values.

In the field of operations research, a series of powerful automatic fitting techniques were developed during the late 1960s. Beard (1967) investigated these computer based procedures for finding optimal values of parameters for a hydrologic model. Ibbitt and O'Donnel (1971) conducted a comprehensive study that compared nine different optimization techniques using simple least squares criterion. The optimization techniques used for their study were : Beard's univariate search technique, Rosenbrock's search technique, a modified version of Rosenbrock's technique, Powell's direct search method, Fletcher and Powell's method, Barnes' least squares method and Karnopp's random search method. They found that the modified version of the Rosenbrock's method showed relative superiority over the others. Nash and Sutcliffe (1970) also reported the use of the modified Rosenbrock's technique with simple least squares as the fitting criterion to investigate the performance of a simple conceptual model. They observed that models containing few independent parameters provided better estimates of runoff volume.

Chapman (1970) was a pioneer among Australian hydrologists to conduct research in the area of automatic parameter estimation. He utilized fitting criterion other than the simple least squares. The other criterion used was the logarithmic transformation of observed and estimated flows before optimization. He concluded that the logarithmic transformation gave three times more weight on the threshold storms than to the larger storms. He further concluded that the fitting criterion would have little effect on the optimal values of the parameters if the model is a realistic simulation of the catchment response. If there exists no errors in the model formulation as well as in the hydrological data, his observation would be true.

Jackson and Aron (1971) reported an extensive and interesting review of parameter estimation techniques in hydrology. They investigated the source of input error and concluded that erroneous input had significant effect on the estimated parameters.

Subjectivity in the selection of an objective function can not be denied. A few researchers have recently tried to deal with this problem. Johnston and Pilgrim (1976) conducted a comparative study of various objective functions while using the Boughton model (an Australian model). The objective functions were functions of various power settings of the deviation of observed and generated flows. They found that squaring the deviation (simple least squares) provided a better objective function than other power settings.

Diskin and Simon (1977) studied the effects of twelve objective functions on parameter estimation for a hydrologic simulation model. Different objective functions estimated twelve different optimal parameter sets with optimal parameters being optimal only in the context of the selected objective function.

They argued for considering more than one objective function in the optimization procedure. Two reasons were cited for this. First, the best function for a given application is not known in advance. Second, optimization procedures using a number of objective functions are less likely to concentrate on local minimum associated with one objective function. Although it seems plausible that one objective function will estimate the parameters and the others can be used to measure model performance; the computational cost involved acts as a deterrent for practical application. Of course, it is interesting to note that they acknowledge the random nature of parameters by considering more than one objective function.

Most of the research described so far analyzed parameter estimation problems from a mathematical rather than statistical point of view. However, more recently some researchers have adopted a statistical framework for parameter estimation, considering a parameter as a random variable (rv) with a probability density function (pdf). Sorooshian and Dracup (1980) considered the stochastic nature of model residuals for parameter estimation and showed that a stochastically proper objective function yielded improved parameter estimates. Troutman (1985) discussed extensively errors in rainfall-runoff modeling and observed that errors should be considered random variables characterized by pdfs. Therefore parameter estimation should be based on probabilistic structure of the errors. Haan (1989) made an interesting and extensive review of parameter uncertainty in hydrologic modeling. He suggested that both model and parameter evaluation should rely on statistical considerations.

Concluding Remarks

The form of the distribution of the error term (Eqn. 1) is important since the estimated parameters are thought to be affected by it. Normality assumption of errors has been used extensively in hydrologic literature. The normality assumption is rarely appropriate. With the normality assumption for the error term in nonlinear models, estimated parameters can also be approximated by the normal distribution which in turn facilitates estimating confidence intervals on the estimates. There is very little reference to the effect of other error distributions on parameter estimates and their distributions. There is a need to evaluate the performance of various estimation techniques with respect to the distribution of errors. Very few researchers have tried to evaluate the performance of various estimation techniques in rainfall-runoff modeling. When they do they adhered to normality assumptions. For example, Sorooshian (1980) used a very simple two parameter model for this kind of study. Application of this approach to conceptual operational rainfall-runoff modeling is rare. Instead of rationalizing for a particular method, one should use the parameter estimation technique that works better in the context of the chosen error distribution. This is the main theme of this dissertation. From the literature review, it is obvious that uncertain parameters result from input data errors, model structure, parameter interaction and estimation techniques. Variability in the parameters can be characterized by the pdf. Once pdfs of parameters are known, uncertain model output should be analyzed in a statistical framework.

CHAPTER III

PARAMETER ESTIMATION TECHNIQUES

Introduction

Parameter estimation is the process of deriving model parameters for a particular application. It is the most important requirement for a model before it can be applied to a watershed for a planning, design or operational purpose. It has already been stated that hydrologic models represent the complex natural processes in a simplified way. Therefore, all the model parameters are not always related to the physical watershed characteristics. Only very few parameters can be obtained directly from field measurement. The remaining parameters are sought through indirect ways from the model results. Kuczera (1982) classified the approaches currently employed for parameter estimation into three groups:

a. priori one

b. curve fitting and

c. mixed approach

The simplest procedure is priori one in which model parameters are evaluated from measurable watershed characteristics, published tables and charts. Watershed drainage area, channel specific capacity, and Manning's n are some examples.

Curve fitting is the process of fitting the model functions (in terms of its parameters) to data with some arbitrary criteria. This includes trial and error calibration, least squares, absolute value difference, method of moments and maximum likelihood. In most cases a mixed approach is used combining both priori one and curve fitting techniques.

Generally, an objective function of the error term is optimized to obtain the optimal parameter set. Specifying an objective function, the user tries to produce residuals having certain statistical properties so the parameters can be regarded optimal. A general form of an objective function can be written as

$$O.F. = n^{-1} \sum_{i=1}^{n} \left[\underline{Y}_i - f(\underline{X}_i;\underline{P}) \right]^{\xi} \tag{4}$$

where ξ is any positive number, $\underline{Y}_i$ is a nx1 vector of observations, $\underline{X}_i$ is a nxk matrix of inputs and $\underline{P}$ is an mx1 vector of parameters.

For most rainfall-runoff models this equation has no closed form solution. One has to apply numerical techniques to estimate $\underline{P}$.

The method of moments and Bayesian analysis are rarely used for parameter estimation in conceptual rainfall-runoff models. The method of moments is used to estimate distribution parameters of hydrologic random variables. However, its efficiency is very poor because most of the hydrologic variables are from skewed distributions. Bayesian analysis has been used to quantify parameter uncertainty in the parameters of flood frequency models (Davis et al., 1972, Vicens et al., 1975, Wood, 1976, Wood and Rodriguez Iturbe, 1975, Bodo and Unny, 1976, Edwards, 1988 and others). To apply Bayesian analysis, one must have prior information about the parameters that can be expressed in probabilistic terms. This prior distribution is then used to determine the posterior density that characterizes the parameter uncertainty. Misspecification of the prior probability density of the parameters may result in erroneous analysis. The method of least squares, method of absolute errors and maximum likelihood techniques are discussed in the following sections because they are widely used in rainfall-runoff models.

Method of Least Squares

Least Squares (LS) is the most widely used method for estimating parameters in hydrologic models. Parameters are estimated by minimizing the error sum of squares between the observed and model predicted values. When $\xi = 2$ in Eqn. 4, the least squares objective function results.

$$\text{O.F.} = n^{-1} \sum_{i=1}^{n} \left[\underline{Y}_i - \underline{f(X_i;P)} \right]^2 \tag{5}$$

Solution of the above equation depends on whether the model is linear or nonlinear in the parameters.

Linear Models

A linear model may be written in vector notation as

$$\hat{Y}_i = \underline{X_i P} + \epsilon_i \tag{6}$$

where ϵ_i is the residual of the ith prediction. Estimation of $\underline{P}$ from the least squares criterion results in the following objective function:

$$\underset{\hat{\underline{P}}}{\text{MIN}} \left[(\underline{Y} - \underline{X\hat{P}}) \, (\underline{Y} - \underline{\hat{P}})^T \right] \tag{7}$$

Minimization of Eqn. 7 results in the well known linear normal equations, given by

$$(\underline{X}^T \underline{X})\underline{\hat{P}} = \underline{X}^T \underline{Y} \tag{8}$$

Now the estimation of $\underline{P}$ is easily accomplished by

$$\underline{\hat{P}} = (\underline{X}^T \underline{X})^{-1} \underline{X}^T \underline{Y} \tag{9}$$

Least Squares Assumptions

Non-singularity of $\underline{X}^T\underline{X}$ is necessary to obtain $\underline{P}$ from Eqn. 9. Up to this point, no other assumptions have been postulated about the nature of the residuals. If some assumptions about the stochastic nature of the residuals are made, the $\underline{P}$ will possess some statistically appealing characteristics. The assumptions made about the residuals are:

1. The residuals have mean zero.
2. The variance of the residual is constant.
3. The residuals are statistically independent of each other.

The above assumptions are known as the least squares assumptions. If these assumptions are obeyed, then the least squares estimates are unbiased and minimum variance estimators. Draper and Smith (1966) show that confidence interval and statistical hypothesis tests may be conducted easily if it may further be assumed that

4. The residuals are normally distributed, i.e. $N(0, \sigma^2)$.

Nonlinear Models

So far the least squares criterion in the context of the linear models has been discussed. However, rainfall-runoff models are highly nonlinear in their parameters and can not be specified in the form of Eqn. 6. The least squares criterion can still be used for parameter estimation however. The objective function to be minimized in this case assumes the form of

$$\min_{\hat{\underline{P}}} \left[\underline{Y} - f(\underline{X};\underline{P})\right] \left[\underline{Y} - f(\underline{X};\underline{P})\right]^T \qquad (10)$$

The resulting normal equation given by Draper and Smith (1966) is in the nonlinear form

$$\sum_{i=1}^{n} \left\{ \left[\underline{Y}_i - f(\underline{X}_i;\underline{P}) \right] \left[\frac{\partial f(\underline{X}_i;\underline{P})}{\partial \underline{P}} \right] \right\}_{\underline{P}=\hat{\underline{P}}} = 0 \tag{11}$$

There is no general closed form solution of the nonlinear normal equations. One is required to use a numerical approach to estimate $\underline{P}$. There is an abundance of literature describing methods for estimating $\underline{P}$ by solution of systems of nonlinear equations (Bard 1974; Beck and Arnold 1977; and others).

Again, the stochastic nature of the residuals dictate the quality of the estimate $\hat{\underline{P}}$. Draper and Smith (1966) reported that if the residuals obey assumptions 1-4 listed earlier, the nonlinear least squares estimate $\hat{\underline{P}}$ may be regarded as identical to the maximum likelihood estimate of $\underline{P}$. Thus the least squares estimates have the same optimal properties as the maximum likelihood estimates; specifically, unbiasedness, minimum variance and asymptotic efficiency.

Violation of the Least Squares Assumptions

The least squares assumptions are very strong assumptions and often are not satisfied by the residuals of hydrologic models (Clarke 1973; and Sorooshian and Dracup 1980). Clarke (1973) aptly noted that parameters of hydrologic models are optimized using the least squares criterion, regardless of its justification. As a result, the parameter estimates are not statistically optimal in various aspects; namely,

1. If the residuals have nonzero mean, parameter estimates are biased.
2. If the variance of the residuals depend on the response, the resulting parameter estimates don't have minimum variance.
3. If the residuals are correlated, both bias and non-minimum variance are introduced into the estimates.

Violation of least squares assumptions can be checked by construction of residual plots and analysis of runs of the residuals (Draper and Smith, 1966). If any of these assumptions are violated, the calibration procedures should be modified (Beck and Arnold, 1977). Power transformations (Box and Hill, 1974) are often employed for this purpose. Sorooshian and Dracup (1980) discuss application of the above transformations for correlated and heteroscedastic errors in hydrologic models.

Method of Absolute Errors

An absolute value objective function can be obtained by putting $\xi=1$ in Eqn. 4. Unlike least squares, parameters are estimated by minimizing the sum of absolute errors.

$$\text{O.F.} = n^{-1} \sum_{i=1}^{n} \left| \underline{Y}_i - f(\underline{X}_i; \underline{P}) \right| \qquad (12)$$

Absolute error optimization can easily be transformed into a linear programming problem with the desired constraint, such as linearity, inequality and nonnegativity. It is specially helpful in hydrologic modeling where some parameters can not assume a negative value (nonnegativity constraint). Linear programming techniques are advantageous because well established algorithms and computer codes are available. Instead of linear programming, the above objective function can also be solved by numerical techniques.

The method of absolute error estimation was recognized by Fourier in the 1820s. For brevity and to contrast with least squares, Fischer (1961) coined the term "Method of Least Lines" for this procedure. Method of least lines has been used in deriving unit hydrograph ordinates by various hydrologist (Eagleson et al. 1966; Deininger 1969; Singh 1976; Mays and Coles 1980 and others). Fischer (1961) reported that the optimality for the method of least squares namely random sampling and the normality assumption often don't exist. In this circumstance least squares gives undue weight to extreme observations. It is just as logical to minimize the sum of absolute errors in this instance.

Method of Maximum Likelihood

The maximum likelihood technique is based on the structure of the residuals of the model. The residuals are assumed to come from a known density function. However, the parameters of the density function are not known. The parameter estimates are obtained by maximizing the distribution taken as a function of the model parameters. This function is called the likelihood function (Mood et al.,1974). By definition, a maximum likelihood estimate is an estimate that has a maximum possibility of being near to a true parameter value (Freund, 1962).

With the normality assumption regarding the errors with mean zero and autocovariance matrix COV, the likelihood function is

$$L(\{\hat{\underline{P}}\},COV) = (2\pi)^{-n/2}\,|COV|^{-1/2}$$

$$EXP\left[-0.5(Q_{mes} - Q_{comp})^T\, COV^{-1}\,(Q_{mes} - Q_{comp})\right] \qquad (13)$$

where COV is nxn nonsingular autocovariance matrix, |COV| is determinant of COV, COV^{-1} is inverse matrix of COV, Q_{mes} is a vector of measured daily runoff volume and Q_{comp} is a vector of computed daily runoff volume.

Taking the natural logarithm of the above equation results in the Log Likelihood function

$$Ln\ L(\underline{P},COV) = -\frac{n}{2}Ln(2\pi)-\frac{1}{2}Ln\left|COV\right|$$

$$-\frac{1}{2}\left[(Q_{mes} - Q_{comp})^T COV^{-1}(Q_{mes} - Q_{comp})\right] \qquad (14)$$

This logarithmic function is simpler than the likelihood function and is maximized to get the parameters. If the following assumptions regarding the autocovariance matrix are made, i.e., the residuals are serially independent with constant variance of σ^2, it can be shown that maximization of Eqn.14 is same as minimization of the simple least squares criterion.

$$COV=E(\underline{e.e}^T)=\begin{cases} E(e_t.e_{t+s})=\sigma^2 & \text{for } s=0 \\ E(e_t.e_{t+s})=0 & \text{for } s\neq 0 \end{cases} \tag{15}$$

Simple least squares may not be appropriate when the above assumption is violated. Violations of least squares assumptions may occur in various ways which have been discussed earlier.

Elements of the COV matrix are required in order to maximize Eqn. 14 with respect to the unknown parameters. The elements of the COV matrix can not be obtained directly because streamflows are a single realization over time. Indirectly, the elements can be evaluated along with the other parameters if we have knowledge about the autocorrelation process of the residuals. A first order autoregressive process is frequently used for hydrologic time series.

$$e_t = \rho\, e_{t-1}+V_t \tag{16}$$

where $|\rho|<1$ is the first lag serial autocorrelation coefficient for the residuals and V_t is a random component. The random vector (V_t) is assumed to come from a multivariate normal distribution with mean zero and variance σ_V^2 Sorooshian and Dracup (1980) utilized the following autocovariance matrix for stationary correlated residuals.

$$\mathrm{COV} = \begin{bmatrix} 1 & \rho & \rho^2 & . & \rho^{n-1} \\ & 1 & \rho & . & \rho^{n-2} \\ & & 1 & . & \rho^{n-3} \\ & & & & 1 \end{bmatrix} \frac{\sigma_v^2}{1-\rho^2} \tag{17}$$

This covariance matrix is known as a "Steady state" autocovariance and has been utilized by Goldberger (1964), Box and Jenkins (1970) and Johnson (1972). With further simplification of the Eqn. 14, Sorooshian and Dracup (1980) formulated the following objective function for Maximum Likelihood Estimation (MLE)

$$\underset{\underline{P}\ \sigma_v\ \rho}{\mathrm{MAX}} \mathrm{MLE} = -\frac{n}{2}\mathrm{Ln}(2\pi) - \frac{1}{2}\mathrm{Ln}\frac{\sigma_v^{2n}}{1-\rho^2} - \frac{1}{2\sigma_v^2}(1-\rho^2)\epsilon_1^2 - \frac{1}{2\sigma_v^2}\sum_{t=2}^{n}\left[(Q_t - \rho Q_t)_{mes} - (Q_t - \rho Q_t)_{comp}\right]^2 \tag{18}$$

Unknown parameters are obtained by maximizing Eqn. 18. The advantage of MLE is that σ_v^2 and ρ can also be evaluated along with other unknown model parameters.

The first term of Eqn. 18 can be regarded as a constant for a particular application. For a large data set (large n), the 2nd term is insignificant. The third term can also be neglected as it contains the square of a very small number (ϵ_1). Thus, Eqn. 18 reduces to the form

$$\underset{\underline{P}\ \sigma_v\ \rho}{\mathrm{MAX}} \mathrm{MLE} = -\frac{1}{2\sigma_v^2}\sum_{t=2}^{n}\left[(Q_t - \rho Q_t)_{mes} - (Q_t - \rho Q_t)_{comp}\right]^2 \tag{19}$$

Maximization of a negative function is the same as minimization of its positive function value. Therefore

$$\underset{\underline{P}\ \sigma_v\ \rho}{\text{MIN}}\ \text{MLE} = \frac{1}{2\sigma_v^2}\sum_{t=2}^{n}\left[(Q_t - \rho Q_t)_{mes} - (Q_t - \rho Q_t)_{comp}\right]^2 \qquad (20)$$

It is evident from Eqn. 20 that when the residuals are uncorrelated, i.e., ρ=0, maximum likelihood estimation is equivalent to the least squares objective function.

Properties of Estimates

To make any kind of comparison as to which parameter estimation technique is best, we must know some desirable properties of the estimates. These properties can then be used to describe the performance of various estimation techniques. Thus, the concept of unbiasedness, minimum variance, consistency, efficiency and sufficiency of the estimates will be discussed briefly. Detailed mathematical treatment of these properties can be found in the literature (Mood et al.,1974; Haan,1977; and Beck and Arnold,1977).

Unbiasedness

In reality, the true values of the parameters are unknown and have to be estimated from the sample realizations. Intuitively, the best technique is one which gives estimates that are in some sense close to the true values. Unbiasedness is a measure of the closeness of the estimate . An estimate $\hat{P}$ of a parameter P is said to be unbiased if $E(\hat{P}) = P$. If bias exists, it is given by $E(\hat{P}) - P$. Unbiasedness does not mean that the specific estimate $\hat{P}$ is equal to P. It only means that the average of many independent estimate for $\hat{P}$ will equal P.

Consistency

An estimate $\hat{P}$ is obtained by a calibration procedure. Therefore, the estimate will have sampling variability and will depend on the particular input data used in the calibration period. A different calibration period would result in a different estimate of $\hat{P}$. $\hat{P}$ is said to be a strongly consistent estimator of P if the difference between $\hat{P}$ and P grows smaller as the calibration period grows larger. Consistency is an asymptotic property since a sufficiently large sample is almost certain to produce estimates close to the true values if the estimates are consistent.

Efficiency

Another measure of the closeness of an estimate to its true value is its variance. Different estimation techniques will lead to estimate $\hat{P}$ with different variances. The best estimate is the one with minimum variance. Because if $\hat{P}$ is unbiased and has a minimum variance, $\hat{P}$ is concentrated near P and have a better chance of being close to true value than those with a larger variance. An estimator $\hat{P}$ is called the most efficient estimator for P if it is unbiased and has a variance which is smaller than any other unbiased estimator for P.

Sufficiency

All the information from a sample which is relevant should be used to estimate $\hat{P}$. In this case, $\hat{P}$ is called the sufficient estimator for P.

The concepts of unbiasedness and minimum variance are small sample properties; on the other hand, consistency is an asymptotic (large sample) property. Troutman (1985) stated that it is not usually possible to obtain estimators with unbiasedness and minimum variance in a nonlinear modeling context. Consistency of estimates is however suitable to assess optimality of a particular estimation technique. Sufficient estimators in rainfall-runoff modeling are perhaps nonexistent. Only 10 to 35 percent of annual rainfall is transformed to runoff volumes. Thus, it is doubtful that estimation of all model parameters solely from a runoff record is sufficient. Kuczera (1983) showed that improved parameters can be obtained by augmenting the estimate other time series such as water table data in parameter estimation of rainfall-runoff model.

CHAPTER IV

RAINFALL-RUNOFF MODEL AND WATERSHED DESCRIPTION

Description of Rainfall-Runoff Model

The model used for this study was the USGS Precipitation Runoff Modeling Systems (PRMS) which is a physically based, highly nonlinear computer model. PRMS can be used to simulate daily flow, storm peaks and volumes, and sediment yields. All components of the model are based on either physical laws or empirical relationships. This model can be regarded as a distributed model in the sense that the watershed can be divided into various homogeneous units, based on hydrologic conditions. The sum of the responses of all units, weighted on a unit area basis, produces the total system response from the watershed.

PRMS can operate on either a daily or storm time scales. The daily mode simulates hydrologic components as daily averages. Streamflow is computed as a mean daily flow. On the other hand, the storm mode simulates hydrologic processes at intervals shorter than a day. In this study, PRMS was run in the daily mode to yield daily runoff predictions. As such, the following sections will deal the model structure and input in relation to the daily mode only. Detailed description of the model can be found in Leavesley et al.(1983).

Model Input

Model input can be classified into two groups - input hydrologic data and input parameters.

Input hydrologic data includes daily rainfall and minimum and maximum daily air temperature. Daily streamflow data are required for optimization purposes only. If snowmelt is considered, daily solar radiation data are required as input. Daily minimum and maximum air temperature are used to estimate evaporation. Instead of air temperature data, pan evaporation data can also be utilized directly.

Input parameters include quantitative data on vegetation, soils and hydrologic characteristics of the watershed. Some parameters can be directly obtained from measurable watershed characteristics, while others are estimated during calibration.

Model Structure

PRMS conceptualizes the watershed system as a series of reservoirs called the impervious zone, soil zone, subsurface zone and groundwater reservoir. The output from these reservoirs combine to generate the total system response. The impervious zone reservoir has no infiltration capacity. Once the maximum retention storage capacity of this reservoir is satisfied, surface runoff can occur. The soil zone reservoir extends to the predominant rooting depth and has a maximum retention capacity of SMAX which occurs at field capacity. This zone is divided into two layers. The upper layer is the recharge zone having a maximum retention capacity of REMX. Losses from the recharge zone occur through evaporation and transpiration. The lower zone of soil reservoir loses water through transpiration only. Once SMAX is satisfied, the excess water recharges the subsurface and groundwater reservoirs. Streamflow is the sum of the surface runoff, subsurface flow and groundwater flow.

The model is shown schematically in Figure 1.

A brief description of the hydrologic processes used in the model is presented in the following paragraphs.

Surface Runoff

Daily surface runoff is computed using a contributing area concept. The percentage of an HRU (hydrologic response unit) contributing to surface runoff can be computed as a nonlinear function of antecedent soil moisture and rainfall. The contributing area (CAP) is computed from

$$CAP = SCN * 10^{(SC1 * SMIDX)} \tag{21}$$

where SCN and SC1 are coefficients , SMIDX is the sum of the current available water in the soil zone (SMAV) plus one-half of the daily net precipitation (PTN) and CAP is the contributing area as a decimal fraction of total HRU area.

$$SRO = CAP * PTN \tag{22}$$

where PTN is the daily net precipitation (inch).

Total precipitation (PPT) is reduced through interception storage and throughfall available from the predominant vegetation. Therefore net daily precipitation (PTN) is a function of the cover density and storage available for the vegetation on an HRU. It is computed by

$$PTN = [PPT * (1 - COVDN) + (PTF * COVDN)] \tag{23}$$

where PPT is the total daily precipitation received on an HRU (inch), COVDN is the seasonal cover density and PTF is the precipitation falling through the canopy (inch).

Daily flow from impervious areas is computed using total daily precipitation (PPT). Impervious area is assigned a maximum retention storage capacity (RETIP). Once RETIP is satisfied, the remaining PPT becomes runoff.

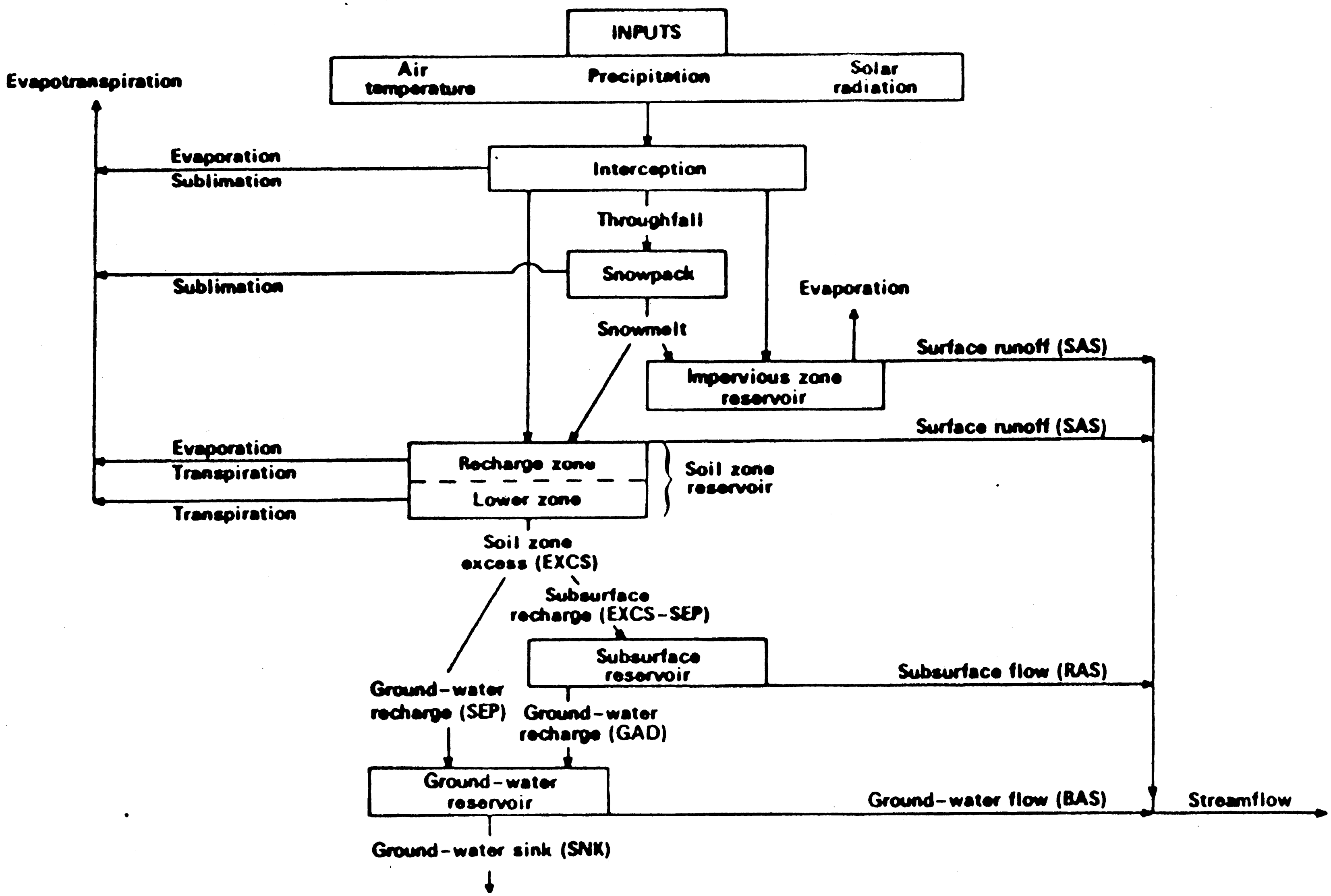

Figure 1. Flow diagram for PRMS

Evapotranspiration

Evapotranspiration (ET) is one of the most important processes of the hydrologic cycle. It influences the spatial and temporal distribution of soil water as well as the antecedent hydrologic conditions. There are various empirical methods to estimate ET using pan evaporation, air temperature and solar radiation data. In this study , air temperature data are used to calculate ET as they are readily available for the watershed. The Hamon formulation (Hamon, 1961) is used to estimate daily potential evapotranspiration (PET) as a function of daily mean air temperature and possible hours of sunshine. PET is computed by

$$PET = CTS(MO) * DYL^2 * VDSAT \qquad (24)$$

where CTS is a coefficient for month MO , DYL is possible hours of sunshine, in units of 12 hours and VDSAT is the saturated water-vapor density at the daily mean air temperature in g/m^3.

Federer and Lash (1978) proposed the following equation to compute VDSAT.

$$VDSAT = 216.7 * \frac{VPSAT}{TAVC+273.3} \qquad (25)$$

where VPSAT is the saturated vapor pressure in mb at TAVC and TAVC is the daily mean air temperature, ^0C.

VPSAT is calculated as an exponential function of daily mean air temperature (Murray,1967)

$$VPSAT = 6.108 * EXP\left(17.26939 * \frac{TAVC}{TAVC+237.3}\right) \qquad (26)$$

Actual evapotranspiration (AET) is computed as a function of the available water to satisfy PET. With nonlimiting availability of water, AET equals PET. Interception storage and retention storage on impervious areas are used first to satisfy PET. Remaining PET demand then is fulfilled from the soil zone storage. AET is computed for the soil zone as a function of the ratio of currently available water to its maximum available water holding capacity. AET computed for the recharge zone is used first to fulfill PET, any remaining demand is met by AET from the lower zone of soil reservoir. This concept of AET-PET computations for sand, loam and clay soils as a function of the soil-water ratio was introduced by Zahner (1967).

Infiltration

Computations of infiltration depend on the form of the precipitation input. When daily rainfall occurs on a snowfree HRU, infiltration is computed as the difference between the net precipitation and surface runoff. For snowmelt, infiltration is nonlimiting until the soil reaches field capacity. Once field capacity is reached, a user-defined daily infiltration capacity (SRX) controls the daily infiltration volume. Snowmelt in excess of SRX contributes to surface runoff. All infiltration in excess of SMAX is routed to the subsurface and groundwater reservoirs.

Subsurface Flow

The source of subsurface flow is the soil water in excess of field capacity. Inflow to the subsurface reservoir occurs when the excess soil water is greater than the daily recharge rate (SEP) to the groundwater reservoir. Subsurface flow is calculated using a reservoir routing system. For the subsurface flow system, the continuity of mass equation is expressed as

$$RAS = INFLOW - \frac{d(RES)}{dt} \qquad (27)$$

where RAS is the rate of outflow from the subsurface reservoir (inch/ Δt), INFLOW is the rate of inflow to the subsurface reservoir (inch/ Δt) , and RES is the storage volume in the subsurface reservoir (in).

RAS is again expressed as a nonlinear function of RES using the relationship

$$RAS = RCF * RES + RCP * RES^2 \tag{28}$$

where RCF and RCP are routing coefficients.

These two equations are combined to solve for RES. Once RES is known, subsurface flow RAS is computed from the continuity equation. For daily flow simulation, Δt equals 24 hours.

Groundwater Flow

Inflow to the groundwater reservoir occurs from both the soil zone and the subsurface reservoir. Inflow from the soil zone occurs when field capacity is exceeded and is limited by a maximum daily recharge rate (SEP). Recharge from the subsurface reservoir to the groundwater reservoir (GAD) is computed by the relationship

$$GAD = RSEP * REXP * \frac{RES}{RESMX} \tag{29}$$

where RSEP is a daily recharge coefficient , RES is the current storage in the subsurface reservoir (in), RESMX and REXP are coefficients.

Baseflow (BAS) is computed by

$$BAS = RCB * GW \tag{30}$$

where RCB is the reservoir routing coefficient and GW is the groundwater reservoir storage (acre-in).

Loss of water from the groundwater reservoir to points beyond the area of interest is treated using a groundwater sink. Daily accretion to the sink (SNK) is computed by

$$SNK = GSNK * GW \tag{31}$$

where GSNK is a seepage constant.

Watershed Description

The model was applied to Chickasha R-6 watershed located in the Washita River Basin of Southwest Central Oklahoma. The Chickasha R-6 watershed is about 9 miles northeast of Chickasha in Grady County, Oklahoma. This is an USDA-ARS watershed that is 27 acres (11 hectares) in area. Average slope of the watershed is about 3 percent. Figure 2 is a topographic map of the watershed. The watershed was selected because of its uniformity in land use and soils.

The Chickasha R-6 is a native grassland watershed. Approximately 42% of the area supports a cover of short grasses consisting primarily of buffalo grass and blue grama grass. The remainder of the area is covered with mild-tall grasses consisting mainly of little bluestem.

USDA-ARS (1966) describes the soils of this watershed. They are 53 percent Grant silt loam, 42 percent Renfrow silt loam and 5 percent Kingfisher silt loam. Hydrologic data have been monitored in this watershed since 1966. Four years of data (1974-1977) were used in this study.

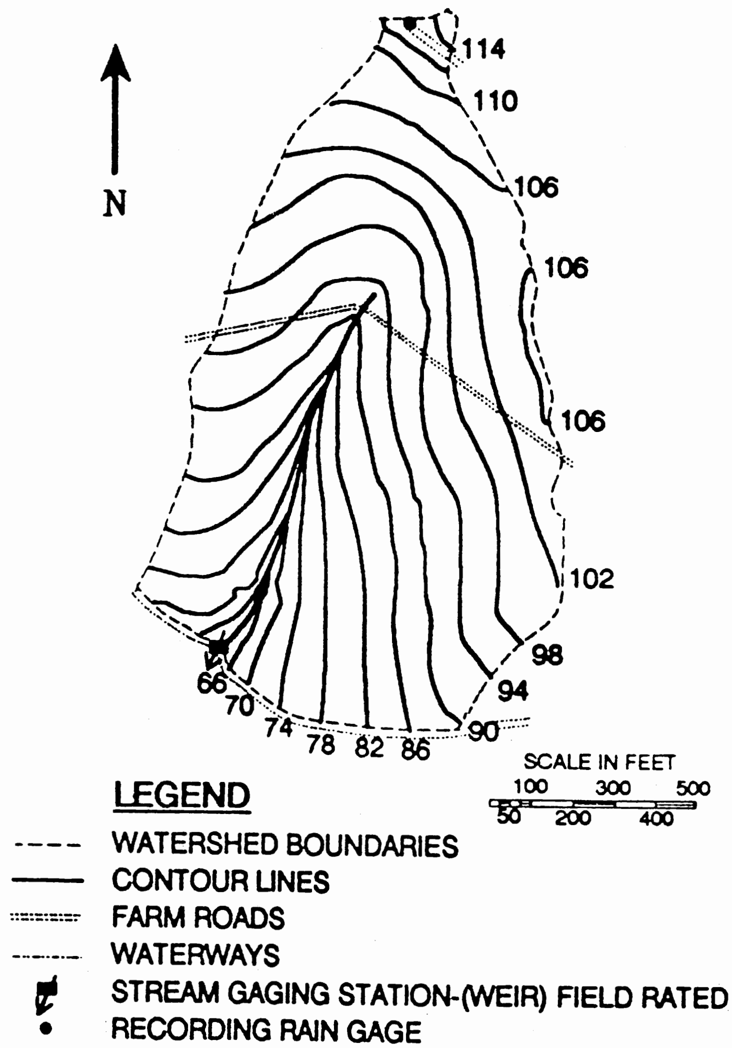

Figure 2. Topographic Map for Chickasha R-6 Watershed

CHAPTER V

METHODOLOGY

Monte Carlo experiments used to evaluate the parameter estimation techniques are described in this chapter. This research heavily draws on random data generation. Therefore, data generation techniques are presented first followed by a discussion of the general procedure.

Random Error Generation

The integral of a probability density function (pdf) between $-\infty$ and ∞ is uniformly distributed over the interval (0, 1). The procedure used to generate a random observation, e, from a pdf is to:

1. Select a random number R_u from a uniform distribution in the interval (0, 1).
2. Set R_u equal to the integral, and
3. Solve for e.

This process is known as the inverse transform of the pdf (Haan, 1977). Many computer routines are available to generate random numbers in the interval (0, 1). An analytic integration of the probability distribution functions is not always possible and numerical methods are often applied instead.

Normal Distribution

Pdf of a normal random variable, e is given by

$$P_e(e) = (2\pi\sigma^2)^{-1/2} \text{ EXP } \frac{-(e-\mu)^2}{2\sigma^2} \text{ for } -\infty < e < \infty \tag{32}$$

An analytic inverse transform can not be found for normal distribution. So the following procedures are used to generate a normal random variable.

1. Generate a random number, R_n, from the standard normal distribution based on a procedure given by Wolfe and Koelling (1983).

2. Generate a normal observation from the relationship

$$e = \sigma R_n + \mu \tag{33}$$

where μ and σ are the mean and standard deviation of normal distribution of e.

In this study, μ was taken as the error-free rainfall or runoff value and σ was varied from 10% to 30% of the error-free value. Random normal errors thusly generated were added to rainfall and runoff data respectively. These contaminated data were then used for parameter estimation.

<u>Lognormal Distribution</u>

Generation of lognormal random variables is based on a numerical method developed for normal deviate generation. If Y= ln(X) is normally distributed with mean μ_1 and variance σ_1^2, then X is lognormally distributed with mean μ and variance σ^2. This property is used for lognormal data generation.

First a normal deviate is generated based on Eqn. 33 and μ and σ, the mean and standard deviation of logarithmically transformed data. They are estimated from the mean and standard deviation of untransformed data using the relationship from Chow (1954).

$$\mu_1 = \frac{1}{2} \text{Ln} \frac{\mu^2}{C_v^2+1} \quad (34)$$

$$\sigma_1 = \text{Ln}\ (C_v^2 + 1) \quad (35)$$

where μ is the mean of the untransformed data and C_v is the coefficient of variation of the untransformed data. A random Y is generated from a normal distribution with mean μ_1 and variance σ_1^2 and transformed to e from the relationship

$$e = \text{EXP}(Y) \quad (36)$$

where e is the lognormally distributed error.

For a particular data point, e is added to the error-free value, μ_1 to obtain the contaminated observation. C_v was varied from 0.1 to 0.3.

Uniform Distribution

The cumulative distribution function of a uniform random variable e is

$$P_e(e) = \frac{(e-\alpha)}{(\beta-\alpha)} \text{ for } \alpha < e < \beta \quad (37)$$

Deriving e from Eqn. 37, the following relationship is obtained.

$$e = \alpha + (\beta - \alpha) * P_e(E) \quad (38)$$

A random uniform variate is determined by selecting a number $P_e(e)$ in the interval (0, 1) and calculating e from Eqn. 38. Estimates for α and β are

$$\hat{\alpha} = \bar{e} - \sqrt{3}\, s \quad (39)$$

$$\hat{\beta} = \bar{e} + \sqrt{3}\, s \quad (40)$$

The standard deviation s was varied from 0.1 to 0.3 of error-free value.

Double Exponential Distribution

Cumulative distribution of a double exponential variable, e is given by

$$P_e(e) = \frac{1}{2} \text{EXP} \frac{(-|e|)}{\beta} \tag{41}$$

Now e can be generated as

$$e = - \beta * \text{Ln} (2P_e(e)) \tag{42}$$

The mean of e is zero and the variance is $2\beta^2$.

A random number $P_e(e)$ in the interval (0, 1) is selected and e is determined from Eqn. 42.

Correlated Errors

An autoregressive AR(1) model was used for correlated errors.

$$e_t = \rho e_{t-1} + v_t \tag{43}$$

where ρ is the correlation between two errors one time interval apart and v_t is a normally distributed random component $N(0, \sigma^2)$. In this study σ ranged from 0.1 to 0.3. Correlation coefficients, ρ used were 0.3, 0.5 and 0.8. Random component v_t was generated as discussed in the normal distribution section. Here two successive errors were correlated. For day 2, error from day 1 and the random component v_t constituted the total error. The error for a particular data point was only the random component if the previous value of rainfall or runoff was zero. Errors were not added to zero values of rainfall or runoff. Correlation was introduced into the second and subsequent value of rainfall or runoff for nonzero sequences only.

Listings for the computer programs used for generation of random numbers from various distributions are given in Appendix A.

General Procedure

The model used in this research was the USGS PRMS model. Hydrologic data (1974 - 1977) records were taken from Chickasha R-6 watershed in Oklahoma. A small watershed (27 acres) was selected so it could be regarded as a homogeneous unit on its hydrologic characteristics. Although the model has watershed partitioning capability, only one hydrologic unit (HRU) was considered to facilitate interpretation of the results. Description of the model and watershed were given in chapter IV.

In this study, PRMS was run in the daily mode to yield daily runoff predictions. PRMS has more than 50 parameters but all of them are not sensitive in the daily mode. From preliminary runs, three of the more sensitive parameters SMAX, REMX and SC1, were selected. A brief explanation of these parameters is given in Table 1. Additional model inputs were daily maximum and minimum air temperature, daily rainfall, and daily streamflow for optimization purposes.

Initially the model was run in the optimization mode using the observed flow data to find representative values for the three parameters. This set of parameters (Table 1) was assumed to be true values for the watershed. To study the effect of data errors on parameter estimation, an error-free standard for making comparisons must be available. Therefore, four years of rainfall and air temperature data were routed through PRMS to generate a synthetic runoff sequence that can be viewed as error-free for the optimized parameters in Table 1. Obviously this set of parameter values should perfectly fit the model to this simulated data.

Parameter estimation techniques employed for this study were the method of least squares (OLS), method of absolute errors (MAE) and maximum likelihood techniques (MLE). Detailed description of these methods were presented in Chapter

TABLE 1

PARAMETERS OF THE PRMS MODEL

No	Variable	Typical value	Description
1	SMAX	6.93 inch	Maximum available water holding capacity of the active soil zone.
2	REMX	2.00 inch	Maximum storage capacity of the recharge zone.
3	SC1	0.45	Coefficient in nonlinear relationship between contributing area and soil moisture.

III. All three methods exactly recovered the true parameter values when both rainfall and runoff data were error-free.

Errors were introduced separately into each value of the rainfall and synthetic runoff records. Error distribution considered were normal, lognormal, double exponential, uniform distribution and autoregressive errors. For any particular data point, the error was selected from a distribution whose mean was the error-free value and whose coefficient of variation was some percentage of the error-free value. Coefficient of variation was varied from 10% to 30%. For correlated error, ρ was set to 0.3, 0.5 and 0.8. For example, a 20% normal error in runoff equals a random number from the normal distribution whose mean is the error-free runoff value and whose standard deviation is 20% of the error-free value. In this study such an error is termed as a 20 percent error. Obviously the actual error varied randomly according to the distribution. Errors were not added to zero values. Errors in temperature input were assumed to be negligible and hence those data sets were not contaminated.

Three error situations were analyzed. The first series of tests used error free rainfall data and a contaminated runoff record. In the second scenario, only precipitation records were contaminated. The last type of test used error contaminated rainfall and runoff records to estimate the parameters. In each case, 100 different, independent simulations were made. For each simulation the parameters were estimated using the Rosenbrock (1960) optimization scheme. The parameters converged within 100 iterations although it generally took many fewer iterations than this. These procedures were repeated for each error distribution and parameter estimation technique. There were altogether 105 combinations of these experiments, each having 100 independent optimizations of the parameter set.

CHAPTER VI

RESULTS AND DISCUSSIONS

This chapter describes the results of the Monte Carlo simulations and analysis used to evaluate the three parameter estimation techniques under various error scenarios. The performance of the estimation techniques is discussed in the first section. Parameter correlations are then presented. Following the discussion of parameter interaction, parameter distributions are discussed. Finally an analysis of uncertainty of model output resulting from uncertain parameters is presented.

Estimation Technique Evaluation

The summary of the 105 Monte Carlo experiments are presented in Table 2, 4, 6, 8, and 10. For each case, the mean and standard deviation (SD) of the parameters are shown. Each row in these Tables is based on 100 independent observations obtained from Monte Carlo studies. The error model used for rainfall (R) and streamflow (Q) are given in column 1 and 2. Information regarding the error distribution is given in the upper part of the tables. Column 3 contains the estimation techniques used in this study. They were ordinary least squares (OLS), method of absolute error (MAE) and maximum likelihood estimation (MLE). The remaining columns contain the mean and SD of the parameters SMAX, REMX and SC1 respectively.

The performance of the three estimation techniques in the context of a given error model is discussed in the following sections. The estimator SD describes the spread of the estimates around their average value. The mean of the estimator provides its bias from the true mean. The mean squared error (MSE) statistic is the most common measure of overall accuracy of the estimators. The MSE describes the variance of the estimates around their true value. MSE can also be shown as a measure of both estimator bias and variance (Mood et al.,1974). Consequently the estimation technique which provides estimates with minimum MSE is judged to be better than the other techniques. The equation used to calculate MSE is

$$MSE = BIAS^2 + VARIANCE \tag{44}$$

Estimated BIAS and MSE statistics of the parameters resulting from different estimation techniques for each error model are shown separately. Table 3, 5, 7, 9 and 11 contain these results. Error models are shown in column 1 and column 2 of these tables. Column 3 presents the objective function used for parameter optimization. The remaining six columns show the BIAS and MSE statistics for SMAX, REMX, and SC1 respectively. Distributions of the error models are given at the top of each table. Parameter estimation techniques were compared for each error model presented on a particular table. These tables could not be compared among themselves because different amounts of errors were introduced by various error models. The magnitudes of the statistics vary from table to table. A specific discussion of each table is given in the following paragraphs.

Normal Distribution

The most widely used error model in hydrologic investigations is the normal distribution. In this study, the results for normal errors are shown in Table 2 and Table 3. Notice that the mean of the estimates are biased. The magnitude of the imposed error affects the population of the best fit parameters. The estimator bias and standard deviation (SD) increase with increasing error levels. Interestingly, the error contaminated precipitation data increased the variance of the parameters more than the same degree of contamination in the runoff record. The parameter REMX

TABLE 2

SUMMARY STATISTICS OF SMAX, REMX AND SC1, 1

% Err. R	Q	Obj. fn.	SMAX MEAN	SD	REMX MEAN	SD	SC1 MEAN	SD
				Normal Distribution of Errors				
0	0		6.930		2.000		0.450	
0	10	OLS	6.942	0.227	1.992	0.172	0.450	5.254E-3
0	10	MAE	6.890	0.182	2.028	0.135	0.451	4.444E-3
0	10	MLE	6.942	0.225	1.997	0.176	0.450	5.311E-3
0	20	OLS	6.974	0.513	2.070	0.426	0.449	1.494E-2
0	20	MAE	6.874	0.294	2.073	0.270	0.451	1.115E-2
0	20	MLE	6.972	0.522	2.074	0.372	0.449	1.489E-2
0	30	OLS	7.261	0.711	2.113	0.511	0.443	2.012E-2
0	30	MAE	6.993	0.478	2.089	0.335	0.448	1.746E-2
0	30	MLE	7.231	0.660	2.097	0.502	0.443	1.777E-2
10	0	OLS	7.075	0.624	2.062	0.526	0.448	1.652E-2
10	0	MAE	6.903	0.356	2.079	0.353	0.449	1.446E-2
10	0	MLE	7.082	0.621	2.039	0.532	0.448	1.642E-2
20	0	OLS	7.228	1.165	2.243	0.713	0.431	2.939E-2
20	0	MAE	6.973	0.755	2.214	0.548	0.432	2.670E-2
20	0	MLE	7.175	1.179	2.241	0.711	0.434	2.847E-2
30	0	OLS	7.078	1.248	2.310	0.824	0.414	4.298E-2
30	0	MAE	7.028	1.110	2.245	0.668	0.406	4.357E-2
30	0	MLE	7.023	1.220	2.278	0.824	0.417	3.979E-2
10	10	OLS	7.063	0.664	2.031	0.542	0.448	1.785E-2
10	10	MAE	6.880	0.372	2.074	0.397	0.448	1.583E-2
10	10	MLE	7.046	0.656	2.039	0.553	0.449	1.896E-2

TABLE 3

BIAS AND MEAN SQUARED ERROR OF THE PARAMETERS, 1

% Err. R	% Err. Q	Obj. fn.	SMAX BIAS	SMAX MSE	REMX BIAS	REMX MSE	SC1 BIAS	SC1 MSE
				Normal Distribution of Errors				
0	10	OLS	0.012	0.0517	-0.008	0.0297	0.000	2.7619E-05
0	10	MAE	-0.040	0.0349	0.028	0.0190	0.001	2.0388E-05
0	10	MLE	0.012	0.0505	-0.003	0.0310	0.000	2.8216E-05
0	20	OLS	0.044	0.2655	0.070	0.1861	-0.001	2.2447E-04
0	20	MAE	-0.056	0.0894	0.073	0.0784	0.001	1.2503E-04
0	20	MLE	0.042	0.2743	0.074	0.1439	-0.001	2.2280E-04
0	30	OLS	0.331	0.6146	0.113	0.2735	-0.007	4.5961E-04
0	30	MAE	0.063	0.2324	0.089	0.1200	-0.002	3.1075E-04
0	30	MLE	0.301	0.5254	0.097	0.2612	-0.007	3.6324E-04
10	0	OLS	0.145	0.4098	0.062	0.2800	-0.002	2.7602E-04
10	0	MAE	-0.027	0.1277	0.079	0.1307	-0.001	2.1096E-04
10	0	MAE	0.152	0.4084	0.039	0.2845	-0.002	2.7484E-04
20	0	OLS	0.298	1.4456	0.243	0.5678	-0.019	1.2096E-03
20	0	MAE	0.043	0.5719	0.214	0.3461	-0.018	1.0478E-03
20	0	MLE	0.245	1.4511	0.241	0.5632	-0.016	1.0827E-03
30	0	OLS	0.148	1.5785	0.310	0.7752	-0.036	3.1287E-03
30	0	MAE	0.098	1.2417	0.245	0.5057	-0.044	3.8164E-03
30	0	MLE	0.093	1.4963	0.278	0.7569	-0.033	2.6920E-03
10	10	OLS	0.133	0.4588	0.031	0.2951	-0.002	3.2172E-04
10	10	MAE	-0.050	0.1410	0.074	0.1628	-0.002	2.5500E-04
10	10	MLE	0.116	0.4442	0.039	0.3068	-0.001	3.6152E-04

was overestimated most of the time. On the other hand, bias in SMAX and SC1 are consistently in different directions i.e. overpredictions of SMAX is associated with underestimates of SC1 and visa-versa. This suggests there may be correlation between them. Another interesting point for normal error is that MAE produced the estimates with minimum variance and MSE for each set of data along with a relatively small bias. MAE was superior to other techniques because MSE statistics for each parameter estimated by MAE were smaller than the others (Table 3). Performance of OLS and MLE is not significantly different from each other in all cases. This follows because, as was shown in the Chapter III, for normal error, minimization of OLS is same as the maximization of MLE.

Lognormal Distribution

Table 4 contains the results for lognormal error. It is evident that the estimated mean value of the parameters is little affected by the error imposed. The standard deviation of the estimated parameters however, increase dramatically as the size of the error increases. Again the contaminated precipitation increased the variance of the parameters more than the contaminated runoff record. Bias of the estimates from mixed error scenarios are similar to either error contaminated rainfall or runoff data, but variations of the parameter are generally higher when both rainfall and runoff are contaminated. Like normal errors, MAE provides the better estimates in the presence of lognormal errors. MSE statistic for each parameter estimated by MAE was smaller (Table 5). Performance of MLE and OLS was similar.

Double Exponential Distribution

The results for the double exponential error are presented in Table 6 and Table 7. Biases of the estimates are higher in all sets of the data. This is because the

TABLE 4

SUMMARY STATISTICS OF SMAX, REMX AND SC1, 2

% Err. R	Q	Obj. fn.	SMAX MEAN	SD	REMX MEAN	SD	SC1 MEAN	SD
			Lognormal Distribution of Errors					
0	0		6.930		2.000		0.450	
0	10	OLS	6.962	0.248	2.031	0.209	0.451	6.569E-3
0	10	MAE	6.917	0.133	2.034	0.126	0.454	3.321E-3
0	10	MLE	6.958	0.243	2.040	0.227	0.451	6.647E-3
0	20	OLS	6.996	0.483	2.148	0.415	0.451	1.109E-2
0	20	MAE	6.874	0.257	2.066	0.229	0.451	9.955E-3
0	20	MLE	6.975	0.485	2.149	0.392	0.452	1.100E-2
0	30	OLS	7.028	0.684	2.187	0.532	0.448	1.893E-2
0	30	MAE	6.850	0.415	2.097	0.333	0.448	1.596E-2
0	30	MLE	7.039	0.677	2.190	0.543	0.448	1.848E-2
10	0	OLS	6.984	0.479	2.079	0.501	0.450	1.508E-2
10	0	MAE	6.893	0.281	2.136	0.417	0.448	1.300E-2
10	0	MLE	6.977	0.493	2.067	0.502	0.450	1.675E-2
20	0	OLS	7.205	1.146	2.152	0.657	0.434	3.677E-2
20	0	MAE	7.104	0.710	2.186	0.534	0.431	2.902E-2
20	0	MLE	7.213	1.131	2.161	0.648	0.434	3.425E-2
30	0	OLS	7.107	1.420	2.203	0.800	0.417	5.018E-2
30	0	MAE	6.925	0.986	2.224	0.648	0.414	4.814E-2
30	0	MLE	7.096	1.359	2.171	0.792	0.417	4.877E-2
10	10	OLS	6.970	0.514	2.056	0.485	0.450	1.541E-2
10	10	MAE	6.903	0.307	2.160	0.408	0.449	1.302E-2
10	10	MLE	6.983	0.516	2.060	0.482	0.450	1.576E-2

TABLE 5

BIAS AND MEAN SQUARED ERROR OF THE PARAMETERS, 2

% Err. R	Q	Obj. fn.	SMAX BIAS	SMAX MSE	REMX BIAS	REMX MSE	SC1 BIAS	SC1 MSE
				Lognormal Distribution of Errors				
0	10	OLS	0.032	0.0625	0.031	0.0448	0.001	4.3796E-05
0	10	MAE	-0.013	0.0179	0.034	0.0169	0.004	2.7029E-05
0	10	MLE	0.028	0.0597	0.040	0.0529	0.001	4.4817E-05
0	20	OLS	0.066	0.2377	0.148	0.1945	0.001	1.2326E-04
0	20	MAE	-0.056	0.0692	0.066	0.0570	0.001	1.0010E-04
0	20	MLE	0.045	0.2376	0.149	0.1760	0.002	1.2323E-04
0	30	OLS	0.098	0.4773	0.187	0.3186	-0.002	3.6315E-04
0	30	MAE	-0.080	0.1789	0.097	0.1205	-0.002	2.5836E-04
0	30	MLE	0.109	0.4703	0.190	0.3312	-0.002	3.4783E-04
10	0	OLS	0.054	0.2320	0.079	0.2573	0.000	2.2739E-04
10	0	MAE	-0.037	0.0803	0.136	0.1921	-0.002	1.7219E-04
10	0	MLE	0.047	0.2447	0.067	0.2560	0.000	2.8051E-04
20	0	OLS	0.275	1.3889	0.152	0.4541	-0.010	1.5987E-03
20	0	MAE	0.174	0.5350	0.186	0.3197	-0.019	1.2187E-03
20	0	MLE	0.283	1.3593	0.169	0.4464	-0.016	1.4423E-03
30	0	OLS	0.177	2.0478	0.203	0.6812	-0.033	3.6264E-03
30	0	MAE	-0.005	0.9720	0.224	0.4699	-0.036	3.6065E-03
30	0	MLE	0.166	1.8733	0.171	0.6562	-0.033	3.4613E-03
10	10	OLS	0.040	0.2657	0.056	0.2380	0.000	2.3745E-04
10	10	MAE	-0.027	0.0947	0.160	0.1923	-0.001	1.7068E-04
10	10	MLE	0.053	0.2689	0.060	0.2358	0.000	2.4866E-04

TABLE 6

SUMMARY STATISTICS OF SMAX, REMX AND SC1, 3

% Err. R	Q	Obj. fn.	SMAX MEAN	SD	REMX MEAN	SD	SC1 MEAN	SD
			Double Exponential Distribution of Errors					
0	0		6.930		2.000		0.450	
0	10	OLS	6.717	0.163	1.958	0.045	0.448	5.219E-3
0	10	MAE	6.930	0.082	2.002	0.052	0.450	3.704E-3
0	10	MLE	6.654	0.627	1.958	0.043	0.448	5.282E-3
0	20	OLS	6.766	0.277	1.923	0.164	0.449	7.313E-3
0	20	MAE	6.899	0.174	2.020	0.132	0.451	7.620E-3
0	20	MLE	6.769	0.278	1.922	0.164	0.449	7.400E-3
0	30	OLS	6.836	0.453	1.885	0.225	0.449	1.134E-2
0	30	MAE	6.838	0.265	2.026	0.151	0.453	1.337E-2
0	30	MLE	6.813	0.422	1.889	0.218	0.450	1.166E-2
10	0	OLS	7.333	0.539	2.111	0.414	0.440	1.883E-2
10	0	MAE	7.014	0.287	2.074	0.391	0.444	1.228E-2
10	0	MLE	7.333	0.537	2.152	0.717	0.439	1.891E-2
20	0	OLS	7.186	0.742	2.247	0.567	0.428	3.193E-2
20	0	MAE	6.973	0.623	2.150	0.440	0.432	2.800E-2
20	0	MLE	7.163	0.704	2.234	0.550	0.430	2.821E-2
30	0	OLS	6.843	1.010	2.279	0.608	0.418	3.520E-2
30	0	MAE	6.846	0.890	2.333	0.577	0.421	3.997E-2
30	0	MLE	6.857	0.992	2.316	0.611	0.417	3.560E-2
10	10	OLS	7.370	0.598	2.088	0.426	0.441	1.965E-2
10	10	MAE	7.035	0.283	2.071	0.370	0.446	1.164E-2
30	10	MLE	7.325	0.569	2.085	0.432	0.441	1.918E-2

TABLE 7

BIAS AND MEAN SQUARED ERROR OF THE PARAMETERS, 3

% R	Err. Q	Obj. fn.	SMAX BIAS	SMAX MSE	REMX BIAS	REMX MSE	SC1 BIAS	SC1 MSE
				Double Exponential Distribution of Errors				
0	10	OLS	-0.213	0.0718	-0.042	0.0038	-0.002	3.2082E-05
0	10	MAE	0.000	0.0067	0.002	0.0027	0.000	1.3807E-05
0	10	MLE	-0.276	0.4689	-0.042	0.0036	-0.002	3.3187E-05
0	20	OLS	-0.164	0.1037	-0.077	0.0330	-0.001	5.5173E-05
0	20	MAE	-0.031	0.0313	0.020	0.0177	0.001	5.9757E-05
0	20	MLE	-0.161	0.1033	-0.078	0.0329	-0.001	5.7004E-05
0	30	OLS	-0.094	0.2137	-0.115	0.0640	-0.001	1.2945E-04
0	30	MAE	-0.092	0.0784	0.026	0.0233	0.003	1.8591E-04
0	30	MLE	-0.117	0.1915	-0.111	0.0601	0.000	1.3618E-04
10	0	OLS	0.403	0.4530	0.111	0.1834	-0.010	4.5865E-04
10	0	MAE	0.084	0.0895	0.074	0.1586	-0.006	1.8566E-04
10	0	MLE	0.403	0.4501	0.152	0.5370	-0.011	4.7010E-04
20	0	OLS	0.256	0.6155	0.247	0.3823	-0.022	1.5076E-03
20	0	MAE	0.043	0.3903	0.150	0.2162	-0.018	1.0940E-03
20	0	MLE	0.233	0.5498	0.234	0.3570	-0.020	1.2161E-03
30	0	OLS	-0.087	1.0278	0.279	0.4474	-0.032	2.2950E-03
30	0	MAE	-0.084	0.7987	0.333	0.4437	-0.029	2.4331E-03
30	0	MLE	-0.073	0.9902	0.316	0.4729	-0.033	2.3369E-03
10	10	OLS	0.440	0.5511	0.088	0.1893	-0.009	4.7088E-04
10	10	MAE	0.105	0.0910	0.071	0.1419	-0.004	1.5070E-04
10	10	MLE	0.395	0.4795	0.085	0.1933	-0.009	4.4708E-04

double exponential distribution introduced larger error to the data. Regarding the effect of precipitation on parameter variance, trends similar to lognormal and normal error is observed. Errors in the rainfall affects the parameter more than the errors in the runoff. Again MAE performed better than the OLS and MLE, producing lower MSE statistic most of the error situations (Table 7).

Uniform Distribution

Table 8 and Table 9 show the results for the uniform distribution of errors. Notice that all the estimates are biased. Biases are specially higher for the parameter SMAX. For runoff error, SMAX was underestimated while overestimation of SMAX occurred in precipitation error. Effect of precipitation error on the parameter variations is more severe than the runoff error. Again overall performance of MAE is better than those of either OLS or MLE based on MSE statistic in Table 9.

Correlated Error

In this case an autoregressive {AR(1)} model is used for the errors. The results are shown in Table 10. Three different autocorrelation coefficients were used in the error model: ρ= 0.3, 0.5 and 0.8. Estimates for precipitation errors were based on 25 simulations instead of 100. Due to larger variations in parameters and slow convergence of the solution, only 25 simulations were run in these cases. Correlated errors introduced a large amount of error in the data. Therefore, performance of all the estimation techniques are not as good as the other error models. Estimates are severely affected by errors in data, specially in the case of precipitation errors. This is because formulation of objective function neglects the errors in precipitation data. Overall performance of MLE was better than OLS and MAE because MLE resulted in smaller MSE statistics (Table 11) for most of the error scenarios. The

TABLE 8

SUMMARY STATISTICS OF SMAX, REMX AND SC1, 4

% Err. R	Q	Obj. fn.	SMAX MEAN	SD	REMX MEAN	SD	SC1 MEAN	SD
			Uniform Distribution of Errors					
0	0		6.930		2.000		0.450	
0	10	OLS	6.698	0.235	1.958	0.090	0.446	6.599E-3
0	10	MAE	6.846	0.199	2.034	0.134	0.451	5.877E-3
0	10	MLE	6.692	0.241	1.963	0.091	0.447	8.077E-3
0	20	OLS	6.774	0.364	1.944	0.601	0.443	1.271E-2
0	20	MAE	6.822	0.379	1.979	0.221	0.449	1.391E-2
0	20	MLE	6.761	0.359	1.894	0.255	0.444	1.248E-2
0	30	OLS	6.798	0.561	1.868	0.390	0.443	1.986E-2
0	30	MAE	6.773	0.498	2.042	0.671	0.448	2.029E-2
0	30	MLE	6.781	0.542	1.871	0.378	0.443	1.993E-2
10	0	OLS	7.393	0.679	2.009	0.559	0.448	1.696E-2
10	0	MAE	7.139	0.492	1.935	0.516	0.448	1.958E-2
10	0	MLE	7.398	0.671	2.006	0.562	0.448	1.715E-2
20	0	OLS	7.385	1.067	2.150	0.780	0.438	3.453E-2
20	0	MAE	7.286	0.901	1.988	0.603	0.432	3.756E-2
20	0	MLE	7.379	1.058	2.142	0.778	0.438	3.528E-2
30	0	OLS	7.138	1.379	2.188	0.867	0.423	4.709E-2
30	0	MAE	7.226	1.129	2.068	0.748	0.414	4.523E-2
30	0	MLE	7.193	1.338	2.121	0.835	0.426	5.032E-2
10	10	OLS	7.322	0.772	1.948	0.571	0.448	2.221E-2
10	10	MAE	7.121	0.529	1.924	0.522	0.449	2.124E-2
10	10	MLE	7.287	0.711	1.967	0.570	0.449	2.102E-2

TABLE 9

BIAS AND MEAN SQUARED ERROR OF THE PARAMETERS, 4

% Err. R	Q	Obj. fn.	SMAX BIAS	SMAX MSE	REMX BIAS	REMX MSE	SC1 BIAS	SC1 MSE
				Uniform Distribution of Errors				
0	10	OLS	-0.232	0.1091	-0.042	0.0099	-0.004	5.9555E-05
0	10	MAE	-0.084	0.0464	0.034	0.0192	0.001	3.5034E-05
0	10	MLE	-0.238	0.1145	-0.037	0.0096	-0.003	7.6795E-05
0	20	OLS	-0.156	0.1571	-0.056	0.3646	-0.007	2.0768E-04
0	20	MAE	-0.108	0.1556	-0.021	0.0492	-0.001	1.9498E-04
0	20	MLE	-0.169	0.1572	-0.106	0.0764	-0.006	1.9810E-04
0	30	OLS	-0.132	0.3320	-0.132	0.1697	-0.007	4.4926E-04
0	30	MAE	-0.157	0.2728	0.042	0.4515	-0.002	4.1589E-04
0	30	MLE	-0.149	0.3161	-0.129	0.1595	-0.007	4.4462E-04
10	0	OLS	0.463	0.6745	0.009	0.3121	-0.002	2.9132E-04
10	0	MAE	0.209	0.2857	-0.065	0.2706	-0.002	3.8590E-04
10	0	MLE	0.468	0.6692	0.006	0.3162	-0.002	2.9826E-04
20	0	OLS	0.455	1.3446	0.150	0.6310	-0.012	1.3412E-03
20	0	MAE	0.356	0.9379	-0.012	0.3642	-0.018	1.7459E-03
20	0	MLE	0.449	1.3202	0.142	0.6258	-0.012	1.3841E-03
30	0	OLS	0.208	1.9451	0.188	0.7863	-0.027	2.9679E-03
30	0	MAE	0.296	1.3618	0.068	0.5637	-0.036	3.3413E-03
30	0	MLE	0.263	1.8596	0.121	0.7121	-0.024	3.0891E-03
10	10	OLS	0.392	0.7498	-0.052	0.3283	-0.002	4.9774E-04
10	10	MAE	0.191	0.3162	-0.076	0.2786	-0.001	4.5222E-04
10	10	MLE	0.357	0.6333	-0.033	0.3262	-0.001	4.4401E-04

TABLE 10

SUMMARY STATISTICS OF SMAX, REMX AND SC1, 5

% Err. R	Q	Obj. fn.	SMAX MEAN	SD	REMX MEAN	SD	SC1 MEAN	SD
				Correlated Errors				
0	0		6.930		2.000		0.450	
Correlation = 0.3								
0	10	OLS	8.089	0.309	1.616	0.059	0.503	5.900E-3
0	10	MAE	8.550	0.287	1.588	0.067	0.498	7.117E-3
0	10	MLE	7.871	0.189	1.615	0.069	0.504	5.223E-3
10	0	OLS	7.091	0.760	3.447	0.072	0.174	9.889E-3
10	0	MAE	7.011	0.233	3.462	0.077	0.172	6.997E-3
10	0	MLE	7.068	0.773	3.450	0.076	0.175	1.036E-3
10	10	OLS	3.647	0.742	1.887	0.939	0.335	4.999E-2
10	10	MAE	3.661	0.842	3.186	0.528	0.327	4.992E-2
10	10	MLE	4.048	1.375	1.834	0.911	0.320	5.180E-2
Correlation = 0.5								
0	10	OLS	7.966	0.220	1.601	0.037	0.505	5.498E-3
0	10	MAE	8.502	0.257	1.573	0.051	0.499	6.979E-3
0	10	MLE	7.846	0.162	1.601	0.036	0.504	4.976E-3
10	0	OLS	5.913	1.889	3.319	0.545	0.204	5.617E-2
10	0	MAE	7.353	0.318	3.477	0.052	0.165	7.859E-3
10	0	MLE	7.122	0.397	3.422	0.157	0.168	9.574E-3
Correlation = 0.8								
0	30	OLS	8.224	0.624	1.599	0.210	0.505	1.586E-2
0	30	MAE	8.740	0.635	1.577	0.169	0.497	1.377E-2
0	30	MLE	7.703	0.379	1.694	0.222	0.507	1.530E-2
10	0	OLS	6.971	0.338	3.437	0.125	0.161	5.632E-3
10	0	MAE	7.160	0.330	3.452	0.088	0.158	6.675E-3
10	0	MLE	6.881	0.399	3.445	0.122	0.163	6.733E-3

TABLE 11

BIAS AND MEAN SQUARED ERROR OF THE PARAMETERS, 5

% Err. R	Q	Obj. fn.	SMAX BIAS	MSE	REMX BIAS	MSE	SC1 BIAS	MSE
					Correlated Errors			
Correlation = 0.3								
0	10	OLS	1.159	1.4383	-0.384	0.1512	0.053	2.8544E-03
0	10	MAE	1.620	2.7073	-0.412	0.1744	0.048	2.3259E-03
0	10	MLE	0.941	0.9208	-0.385	0.1528	0.054	2.9110E-03
10	0	OLS	0.161	0.6041	1.447	2.0984	-0.276	7.6274E-02
10	0	MAE	0.081	0.0609	1.462	2.1440	-0.278	7.7166E-02
10	0	MLE	0.138	0.6171	1.450	2.1080	-0.275	7.5901E-02
10	10	OLS	-3.283	11.3319	-0.113	0.8953	-0.115	1.5701E-02
10	10	MAE	-3.269	11.3955	1.186	1.6856	-0.123	1.7597E-02
10	10	MLE	-2.882	10.1971	-0.166	0.8571	-0.130	1.9635E-02
Correlation = 0.5								
0	10	OLS	1.036	1.1216	-0.399	0.1603	0.055	3.0883E-03
0	10	MAE	1.572	2.5364	-0.427	0.1849	0.049	2.4497E-03
0	10	MLE	0.916	0.8650	-0.399	0.1606	0.054	2.9300E-03
10	0	OLS	-1.017	4.6036	1.319	2.0370	-0.246	6.3918E-02
10	0	MAE	0.423	0.2799	1.477	2.1830	-0.285	8.1173E-02
10	0	MLE	0.192	0.1942	1.422	2.0474	-0.282	7.9841E-02
Correlation = 0.8								
0	30	OLS	1.294	2.0632	-0.401	0.2050	0.055	3.2436E-03
0	30	MAE	1.810	3.6786	-0.423	0.2076	0.047	2.4080E-03
0	30	MLE	0.773	0.7416	-0.306	0.1429	0.057	3.5288E-03
10	0	OLS	0.041	0.1161	1.437	2.0794	-0.289	8.3437E-02
10	0	MAE	0.230	0.1614	1.452	2.1169	-0.292	8.5309E-02
10	0	MLE	-0.049	0.1616	1.445	2.1019	-0.287	8.2644E-02

desired degree of overall correlation in the data set could not be achieved due to large number of zero values in the data set.

Concluding Remarks

It is evident from the above discussion that the magnitude of the imposed error affects the population of the best fit parameters. One can expect this variation to be transferred to variations in the model output. In a practical sense, the result means that the probability of a parameter value deviating from its true value by some amount, δ, increases as the magnitude of random errors in the input data increases. In practice the independent variables are assumed to be error-free, a questionable assumption. The magnitude of the errors investigated in this study is relatively small. A 20 percent error in precipitation, for example, would be a large error for a point measurement; however, for an estimate on a watershed scale an error of this magnitude could easily be exceeded. Often precipitation records from a single gage in a watershed, or even from a location some miles from the watershed, must be used to estimate precipitation. Under these conditions, considerable error in rainfall is not only possible, but is likely. Troutman (1982) showed that runoff predication errors were inflated using erroneous precipitation data. However these results are in contrast to the observation by Ibbitt (1972). He reported that there were no significant variation among the parameters using erroneous (10% normal) precipitation, runoff and pan evaporation data. This may be because he utilized only 60 data points to estimate the 9 parameters of the model. Further his observation may be model specific.

The MAE was found to be superior under all the error distribution assumption except in the case of correlated errors. The watershed considered for this study was small (27.2 acres) and its annual water yield was around 2". Higher values dominates the OLS objective function. MAE performed better in the normal errors perhaps because of the low flow values. Performance of MLE is similar to OLS in the normal error case because correlation among the data set was low. Poor performance of MLE with non-normal errors was because of the misspecification of the error model. The MLE objective function was based on the assumption of correlated normal errors. Performance of MAE is obvious in non-normal error cases. It was in agreement with the observations found in literature. MLE worked better in the correlated error case as its objective function took care of the correlations among the residuals. However due to many no flow situations, the desired degree of correlation could not be achieved. As a result, OLS and MAE perhaps worked well in individual correlated error situations.

Independence of the errors rather than the distribution seems more important for the quality of the estimates. The magnitude of MSE was smaller for each of the independent error situations regardless of the error distributions. However, in the presence of the correlated errors, the quality of the estimates was very poor. Violation of the assumption of uncorrelated errors contributed significant errors in the parameter estimates.

Parameter Interaction

Tables 12 through 16 contain the correlations among the parameters. In each table, the rows represent the results obtained from 100 independent data sets for a particular estimation technique and error model used in column 1 and 2. Column 3 represents the objective functions used in this study, namely ordinary least squares (OLS), method of absolute errors (MAE) and maximum likelihood estimation (MLE). The remaining columns (4 through 6) show the correlations among the parameters. For example, ρ (SM,SC) in the column 5 represents the correlation between SMAX and SC1. Information regarding the error distribution is given in the upper part of the tables.

Normal Distribution

Table 12 shows the correlations obtained when errors were taken randomly from the normal distribution. Notice that the correlations in the column 5 are consistently high for every set of data. Thus the parameters SMAX and SC1 are correlated. ρ for SMAX and SC1 varied from -0.207 to -0.697. The correlations between SMAX and REMX (column 4) and between REMX and SC1 (column 6) are neither consistently high nor consistently low, suggesting there is no strong correlation structure among them. Henceforth, the parameter REMX is considered independent. The correlation between SMAX and SC1 is not surprising as they are functionally related in the model.

Lognormal Distribution

Table 13 represents the parameter correlations when the error models are from lognormal distribution. Irrespective of the error variances and estimation techniques, correlations between SMAX and SC1 are consistently higher, ranging from -0.008 to -0.697. There is very little interaction either between SMAX and REMX or between REMX and SC1.

Double Exponential Distribution

Correlation results for double exponential error models are presented in Table 14. The correlation between SMAX and REMX ranged from -0.021 to -0.334. This range for REMX and SC1 lies between -0.014 and -0.307. On the other hand, correlation between SMAX and SC1 varies from -0.088 to -0.703, indicating higher interaction between these two parameters. They are inversely related because the correlation coefficients are negative.

TABLE 12

CORRELATIONS OF SMAX, REMX AND SC1, 1

% Error R	Q	Objective function	ρ (SM,RE)	ρ (SM,SC)	ρ (RE,SC)
		Normal Distribution of Errors			
0	10	OLS	-0.220	-0.483	0.141
0	10	MAE	-0.040	-0.399	-0.027
0	10	MLE	-0.286	-0.488	0.211
0	20	OLS	0.071	-0.684	-0.178
0	20	MAE	0.165	-0.419	-0.186
0	20	MLE	0.080	-0.697	-0.224
0	30	OLS	0.201	-0.635	-0.317
0	30	MAE	0.218	-0.526	-0.174
0	30	MLE	0.141	-0.548	-0.284
10	0	OLS	-0.110	-0.377	-0.152
10	0	MAE	-0.106	-0.399	-0.046
10	0	MLE	-0.096	-0.395	-0.128
20	0	OLS	0.236	-0.478	-0.193
20	0	MAE	0.262	-0.576	-0.177
20	0	MLE	0.267	-0.496	-0.251
30	0	OLS	-0.018	-0.320	-0.084
30	0	MAE	0.134	-0.543	0.006
30	0	MLE	0.036	-0.207	0.090
10	10	OLS	-0.078	-0.408	-0.091
10	10	MAE	-0.014	-0.405	-0.150
10	10	MLE	-0.075	-0.437	-0.162

TABLE 13

CORRELATIONS OF SMAX, REMX AND SC1, 2

% Error R	Q	Objective function	ρ (SM,RE)	ρ (SM,SC)	ρ (RE,SC)
		Lognormal Distribution of Errors			
0	10	OLS	-0.173	-0.530	0.101
0	10	MAE	0.175	-0.008	-0.083
0	10	MLE	-0.143	-0.536	0.071
0	20	OLS	0.006	-0.610	-0.039
0	20	MAE	0.283	-0.536	-0.221
0	20	MLE	0.058	-0.595	-0.064
0	30	OLS	0.005	-0.648	-0.249
0	30	MAE	0.318	-0.573	-0.292
0	30	MLE	-0.022	-0.640	-0.237
10	0	OLS	0.257	-0.335	-0.215
10	0	MAE	0.275	-0.305	-0.116
10	0	MLE	0.233	-0.413	-0.202
20	0	OLS	0.207	-0.519	-0.101
20	0	MAE	0.113	-0.480	-0.111
20	0	MLE	0.176	-0.484	-0.112
30	0	OLS	0.080	-0.345	-0.079
30	0	MAE	0.134	-0.327	-0.023
30	0	MLE	0.083	-0.311	-0.024
10	10	OLS	0.103	-0.322	-0.038
10	10	MAE	0.185	-0.268	-0.070
10	10	MLE	0.054	-0.312	-0.031

TABLE 14

CORRELATIONS OF SMAX, REMX AND SC1, 3

% Error R	Q	Objective function	ρ (SM,RE)	ρ (SM,SC)	ρ (RE,SC)
		Double Exponential Distribution of Errors			
0	10	OLS	-0.334	-0.488	-0.038
0	10	MAE	0.127	-0.558	-0.307
0	10	MLE	-0.117	-0.164	-0.033
0	20	OLS	-0.189	-0.415	-0.190
0	20	MAE	-0.037	-0.539	-0.175
0	20	MLE	-0.216	-0.425	-0.134
0	30	OLS	-0.180	-0.351	-0.114
0	30	MAE	0.027	-0.703	-0.202
0	30	MLE	-0.194	-0.391	-0.180
10	0	OLS	-0.207	-0.546	-0.081
10	0	MAE	-0.301	-0.510	-0.107
10	0	MLE	-0.097	-0.532	-0.014
20	0	OLS	-0.090	-0.490	-0.299
20	0	MAE	-0.219	-0.483	-0.206
20	0	MLE	-0.166	-0.418	-0.233
30	0	OLS	-0.093	-0.018	-0.206
30	0	MAE	-0.074	-0.374	-0.150
30	0	MLE	-0.116	-0.088	-0.116
10	10	OLS	-0.046	-0.583	0.043
10	10	MAE	-0.162	-0.398	-0.130
10	10	MLE	-0.021	-0.558	0.036

Uniform Distribution

Table 15 shows the results for uniform error models. Again SMAX and SC1 are correlated because coefficients range from -0.238 to -0.620. These values are smaller for SMAX and REMX and for REMX and SC1. Thus the parameter REMX is statistically independent of both SMAX and SC1. Correlation between SMAX and SC1 are not affected by the error variances, estimation techniques and error scenarios, i.e. whether rainfall, runoff or both are erroneous.

Correlated Errors

The results for the correlated error scenarios are given in table 16. Correlation coefficients used for error models are 0.3, 0.5 and 0.8. Notice that all three coefficients are higher in comparison to normal, lognormal, double exponential and uniform error models. This is perhaps due to large amounts of error introduced by the correlated error models. Correlation between SMAX and SC1 varies from -0.568 to -0.973. On the other hand, correlations between SMAX and REMX are neither high nor low for every set of data. The same can be said about the correlation between REMX and SC1.

Concluding Remarks

From the correlation study, it can be concluded that SMAX and SC1 are correlated. There is low interaction between SMAX and REMX and REMX and SC1 as indicated by their low correlation coefficients. Error models do not have any influence on this result except for the correlated error cases. Larger correlations in Table 16 can perhaps be attributed to the large amounts of error produced by the AR(1) models. The estimation techniques also did not affect the correlation results because interaction between SMAX and SC1 are inherent in the model structure.

TABLE 15

CORRELATIONS OF SMAX,REMX AND SC1, 4

% Error R	Q	Objective function	ρ (SM,RE)	ρ (SM,SC)	ρ (RE,SC)
		Uniform Distribution of Errors			
0	10	OLS	-0.054	-0.412	-0.244
0	10	MAE	-0.073	-0.401	-0.125
0	10	MLE	-0.044	-0.305	-0.232
0	20	OLS	-0.081	-0.336	0.045
0	20	MAE	0.123	-0.503	-0.154
0	20	MLE	0.169	-0.306	-0.059
0	30	OLS	0.155	-0.420	-0.092
0	30	MAE	0.259	-0.363	-0.137
0	30	MLE	0.237	-0.393	-0.070
10	0	OLS	-0.022	-0.404	-0.251
10	0	MAE	-0.025	-0.561	-0.181
10	0	MLE	-0.020	-0.411	-0.262
20	0	OLS	-0.136	-0.357	-0.213
20	0	MAE	-0.187	-0.490	0.028
20	0	MLE	-0.121	-0.355	-0.225
30	0	OLS	-0.260	-0.252	-0.082
30	0	MAE	0.098	-0.442	-0.148
30	0	MLE	-0.202	-0.238	0.052
10	10	OLS	-0.092	-0.620	-0.054
10	10	MAE	-0.125	-0.570	-0.052
10	10	MLE	-0.048	-0.598	-0.144

TABLE 16

CORRELATIONS OF SMAX, REMX AND SC1, 5

% Error R	Q	Objective function	ρ (SM,RE)	ρ (SM,SC)	ρ (RE,SC)
			Correlated Errors		
Correlation = 0.3					
0	10	OLS	0.394	-0.751	-0.289
0	10	MAE	0.432	-0.649	-0.374
0	10	MLE	-0.283	-0.717	0.271
10	0	OLS	0.220	-0.952	-0.278
10	0	MAE	-0.273	-0.614	0.356
10	0	MLE	0.218	-0.954	-0.282
10	10	OLS	-0.592	-0.973	0.547
10	10	MAE	-0.352	-0.968	0.327
10	10	MLE	0.114	-0.911	0.106
Correlation = 0.5					
0	10	OLS	-0.404	-0.686	-0.071
0	10	MAE	0.261	-0.598	-0.400
0	10	MLE	-0.666	-0.568	-0.058
10	0	OLS	0.244	-0.998	-0.226
10	0	MAE	-0.190	-0.864	0.115
10	0	MLE	0.188	-0.973	-0.264
Correlation = 0.8					
0	30	OLS	-0.017	-0.754	-0.104
0	30	MAE	-0.151	-0.623	-0.036
0	30	MLE	-0.167	-0.825	-0.113
10	0	OLS	0.216	-0.834	-0.116
10	0	MAE	0.073	-0.665	-0.129
10	0	MLE	0.102	-0.881	-0.049

SMAX and SC1 are inversely related to each other because of the negative correlation coefficients. For uncertainty analysis, the joint distribution of SMAX and SC1 should be considered to preserve correlation and distributional properties between them.

Parameter Distributions

Normal and lognormal distributions were tested for the parameters SMAX, REMX and SC1. Tables 17 through 21 summarize the results and give the values for the maximum deviation, D1 and D2, between the fitted and empirical normal and lognormal distributions respectively. The Kolmogorov-Smirnov test (Haan, 1977) was used as a criterion of acceptance or rejection of the proposed distribution. Critical value of Kolmogorov-Smirnov test statistic was 0.14 at 5% significance level. Hypothesis that the parameter is from a proposed distribution was accepted when the maximum deviation was less than the critical value of the Kolmogorov-Smirnov test, and visa-versa. In these tables, column 1 and 2 represent the error models. Column 3 contains the objective functions used in this study. D1 and D2 values for each parameter are given in the remaining columns.

Parameter SMAX

Figure 3 through 6 illustrate the distribution of the parameter SMAX. Although the normal distribution was not rejected, the lognormal distribution was found to more accurately describe SMAX. Deviations D2 were smaller than D1 most of the times. With correlated errors and contaminated precipitation, neither the normal nor lognormal distribution can be accepted for SMAX. Nevertheless deviations D2 were smaller than D1 in this case also. The distribution of the error did not have any impact on the resulting distribution of the parameters. The various estimation techniques also did not affect the distribution of SMAX. SMAX can be

TABLE 17

KOLMOGOROV-SMIRNOV TEST STATISTICS, 1

% Error R	Q	Obj. fn.	SMAX D1	D2	REMX D1	D2	SC1 D1	D2
				Normal Distribution of Errors				
0	10	OLS	0.062	0.069	0.114	0.118	0.052	0.054
0	10	MAE	0.150	0.155	0.114	0.099	0.081	0.080
0	10	MLE	0.068	0.075	0.131	0.122	0.067	0.069
0	20	OLS	0.089	0.076	0.088	0.170	0.095	0.105
0	20	MAE	0.078	0.075	0.118	0.093	0.147	0.152
0	20	MLE	0.087	0.074	0.090	0.087	0.100	0.109
0	30	OLS	0.083	0.067	0.110	0.096	0.072	0.082
0	30	MAE	0.086	0.073	0.121	0.096	0.123	0.133
0	30	MLE	0.079	0.061	0.098	0.088	0.056	0.061
10	0	OLS	0.079	0.065	0.107	0.060	0.117	0.125
10	0	MAE	0.075	0.071	0.122	0.130	0.155	0.162
10	0	MLE	0.076	0.063	0.122	0.072	0.081	0.089
20	0	OLS	0.109	0.096	0.094	0.089	0.073	0.087
20	0	MAE	0.122	0.099	0.060	0.082	0.105	0.122
20	0	MLE	0.132	0.109	0.094	0.082	0.056	0.068
30	0	OLS	0.081	0.053	0.083	0.112	0.071	0.084
30	0	MAE	0.151	0.135	0.088	0.074	0.094	0.115
30	0	MLE	0.093	0.068	0.092	0.106	0.069	0.068
10	10	OLS	0.072	0.071	0.079	0.059	0.111	0.119
10	10	MAE	0.092	0.081	0.093	0.084	0.157	0.165
10	10	MLE	0.070	0.063	0.078	0.051	0.118	0.126

TABLE 18

KOLMOGOROV-SMIRNOV TEST STATISTICS, 2

% Error R	Q	Obj. fn.	SMAX D1	SMAX D2	REMX D1	REMX D2	SC1 D1	SC1 D2
			Lognormal Distribution of Errors					
0	10	OLS	0.066	0.071	0.118	0.097	0.060	0.060
0	10	MAE	0.056	0.057	0.127	0.121	0.366	0.329
0	10	MLE	0.078	0.085	0.126	0.105	0.079	0.077
0	20	OLS	0.083	0.072	0.094	0.087	0.066	0.067
0	20	MAE	0.109	0.116	0.108	0.089	0.091	0.087
0	20	MLE	0.081	0.070	0.094	0.082	0.066	0.066
0	30	OLS	0.096	0.092	0.094	0.108	0.091	0.082
0	30	MAE	0.070	0.063	0.134	0.155	0.103	0.110
0	30	MLE	0.090	0.085	0.095	0.095	0.093	0.084
10	0	OLS	0.063	0.076	0.081	0.072	0.082	0.076
10	0	MAE	0.089	0.081	0.042	0.056	0.097	0.102
10	0	MLE	0.047	0.059	0.076	0.065	0.081	0.074
20	0	OLS	0.093	0.064	0.062	0.082	0.096	0.083
20	0	MAE	0.092	0.090	0.074	0.057	0.074	0.083
20	0	MLE	0.101	0.074	0.070	0.065	0.066	0.060
30	0	OLS	0.093	0.062	0.078	0.092	0.082	0.078
30	0	MAE	0.105	0.083	0.118	0.097	0.105	0.130
30	0	MLE	0.092	0.070	0.064	0.086	0.099	0.097
10	10	OLS	0.045	0.039	0.046	0.076	0.077	0.083
10	10	MAE	0.060	0.056	0.091	0.068	0.111	0.116
10	10	MLE	0.066	0.052	0.038	0.077	0.076	0.081

TABLE 19

KOLMOGOROV-SMIRNOV TEST STATISTICS, 3

% Error R	Q	Obj. fn.	SMAX D1	SMAX D2	REMX D1	REMX D2	SC1 D1	SC1 D2
			Double Exponential Distribution of Errors					
0	10	OLS	0.079	0.077	0.300	0.316	0.138	0.135
0	10	MAE	0.106	0.108	0.113	0.119	0.193	0.191
0	10	MLE	0.348	0.439	0.290	0.307	0.127	0.123
0	20	OLS	0.107	0.102	0.367	0.384	0.089	0.086
0	20	MAE	0.066	0.071	0.123	0.137	0.248	0.242
0	20	MLE	0.107	0.102	0.358	0.374	0.087	0.084
0	30	OLS	0.104	0.095	0.307	0.320	0.128	0.121
0	30	MAE	0.077	0.088	0.134	0.121	0.198	0.190
0	30	MLE	0.109	0.100	0.314	0.327	0.133	0.126
10	0	OLS	0.118	0.112	0.078	0.068	0.175	0.184
10	10	MAE	0.128	0.132	0.091	0.088	0.169	0.176
10	10	MLE	0.120	0.121	0.165	0.076	0.182	0.191
20	0	OLS	0.058	0.059	0.070	0.045	0.159	0.167
20	0	MAE	0.112	0.096	0.057	0.053	0.151	0.156
20	0	MLE	0.078	0.069	0.073	0.056	0.145	0.155
30	0	OLS	0.093	0.124	0.088	0.074	0.110	0.124
30	0	MAE	0.112	0.134	0.074	0.055	0.121	0.140
30	0	MLE	0.078	0.106	0.101	0.086	0.121	0.135
10	10	OLS	0.130	0.116	0.109	0.084	0.202	0.214
10	10	MAE	0.080	0.081	0.078	0.088	0.141	0.148
10	10	MLE	0.109	0.098	0.116	0.084	0.211	0.225

TABLE 20

KOLMOGOROV-SMIRNOV TEST STATISTICS, 4

% Error R	Q	Obj. fn.	SMAX D1	SMAX D2	REMX D1	REMX D2	SC1 D1	SC1 D2
				Uniform Distribution of Errors				
0	10	OLS	0.140	0.136	0.263	0.276	0.085	0.087
0	10	MAE	0.106	0.110	0.167	0.153	0.094	0.097
0	10	MLE	0.143	0.139	0.277	0.291	0.099	0.096
0	20	OLS	0.099	0.090	0.339	0.237	0.055	0.061
0	20	MAE	0.069	0.064	0.121	0.138	0.085	0.092
0	20	MLE	0.109	0.099	0.158	0.189	0.045	0.050
0	30	OLS	0.102	0.102	0.085	0.126	0.111	0.118
0	30	MAE	0.077	0.065	0.225	0.155	0.099	0.108
0	30	MLE	0.107	0.108	0.090	0.130	0.087	0.096
10	0	OLS	0.074	0.056	0.058	0.072	0.106	0.101
10	0	MAE	0.045	0.053	0.138	0.113	0.119	0.121
10	0	MLE	0.068	0.051	0.071	0.065	0.110	0.104
20	0	OLS	0.079	0.073	0.061	0.082	0.101	0.086
20	0	MAE	0.104	0.095	0.116	0.065	0.067	0.077
20	0	MLE	0.080	0.081	0.074	0.079	0.082	0.070
30	0	OLS	0.099	0.133	0.105	0.105	0.033	0.047
30	0	MAE	0.106	0.077	0.100	0.093	0.064	0.084
30	0	MLE	0.109	0.081	0.091	0.091	0.060	0.063
10	10	OLS	0.061	0.054	0.077	0.071	0.110	0.101
10	10	MAE	0.068	0.070	0.120	0.088	0.123	0.116
10	10	MLE	0.068	0.051	0.071	0.065	0.110	0.104

TABLE 21

KOLMOGOROV-SMIRNOV TEST STATISTICS, 5

% Error		Obj.	SMAX		REMX		SC1	
R	Q	fn.	D1	D2	D1	D2	D1	D2
				Correlated Errors				
Correlation = 0.3								
0	10	OLS	0.129	0.120	0.277	0.267	0.097	0.099
0	10	MAE	0.116	0.111	0.291	0.280	0.215	0.212
0	10	MLE	0.050	0.053	0.237	0.241	0.074	0.075
10	0	OLS	0.367	0.357	0.235	0.238	0.456	0.459
10	0	MAE	0.151	0.144	0.364	0.366	0.258	0.262
10	0	MLE	0.368	0.358	0.258	0.259	0.417	0.423
10	10	OLS	0.323	0.385	0.295	0.228	0.319	0.292
10	10	MAE	0.358	0.417	0.350	0.369	0.374	0.366
10	10	MLE	0.405	0.344	0.283	0.202	0.335	0.372
Correlation = 0.5								
0	10	OLS	0.113	0.107	0.218	0.221	0.110	0.112
0	10	MAE	0.100	0.101	0.204	0.198	0.188	0.186
0	10	MLE	0.131	0.132	0.276	0.274	0.050	0.048
10	0	OLS	0.307	0.342	0.370	0.407	0.301	0.276
10	0	MAE	0.217	0.220	0.338	0.339	0.256	0.249
10	0	MLE	0.202	0.201	0.311	0.317	0.176	0.172
Correlation = 0.8								
0	30	OLS	0.090	0.075	0.199	0.172	0.173	0.164
0	30	MAE	0.113	0.100	0.213	0.190	0.137	0.133
0	30	MLE	0.088	0.097	0.306	0.285	0.184	0.177
10	0	OLS	0.152	0.145	0.307	0.312	0.174	0.170
10	0	MAE	0.286	0.277	0.318	0.320	0.195	0.207
10	0	MLE	0.181	0.195	0.364	0.364	0.199	0.189

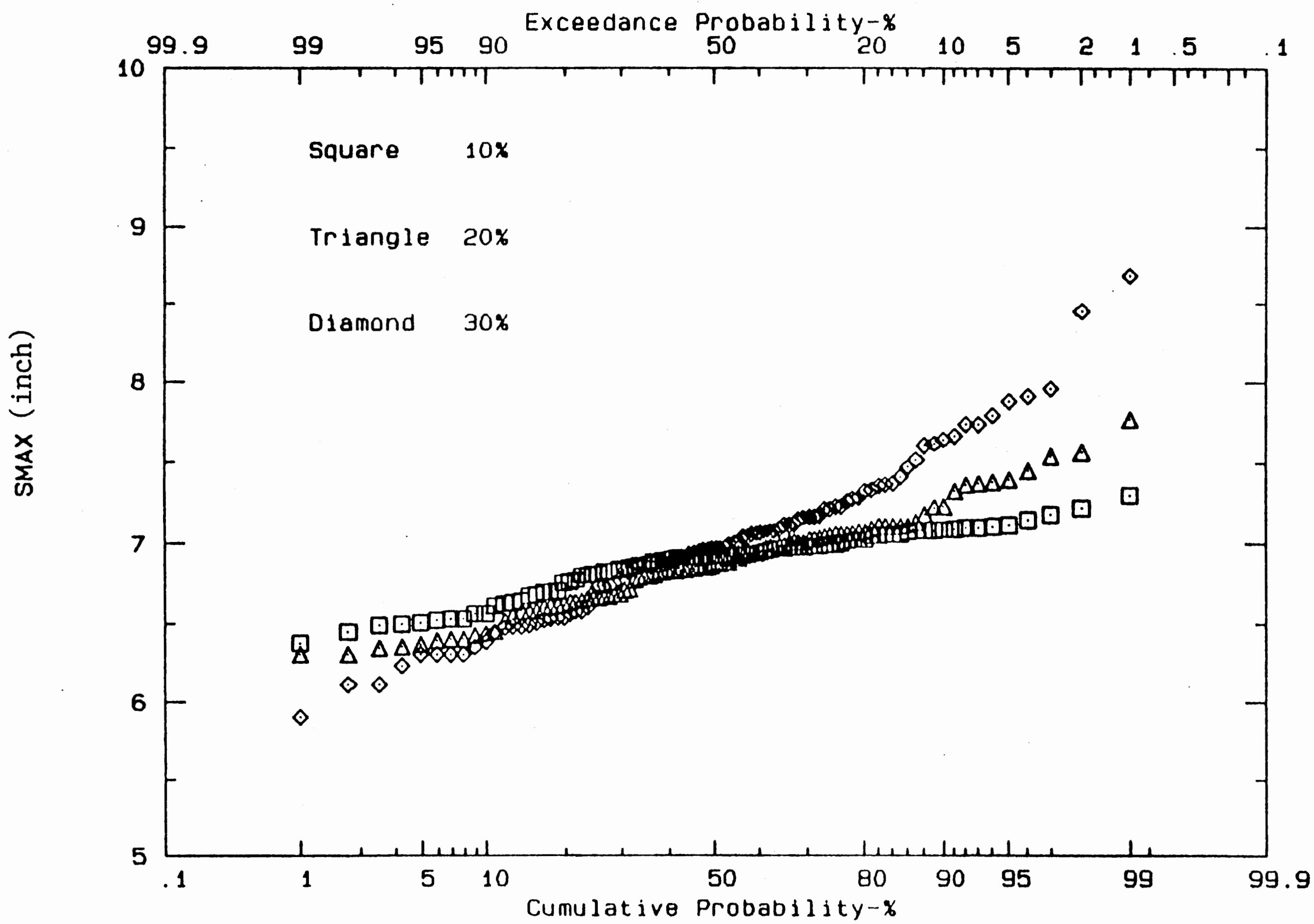

Figure 3. Distribution of SMAX Due to Normal Error in Flow (MAE)

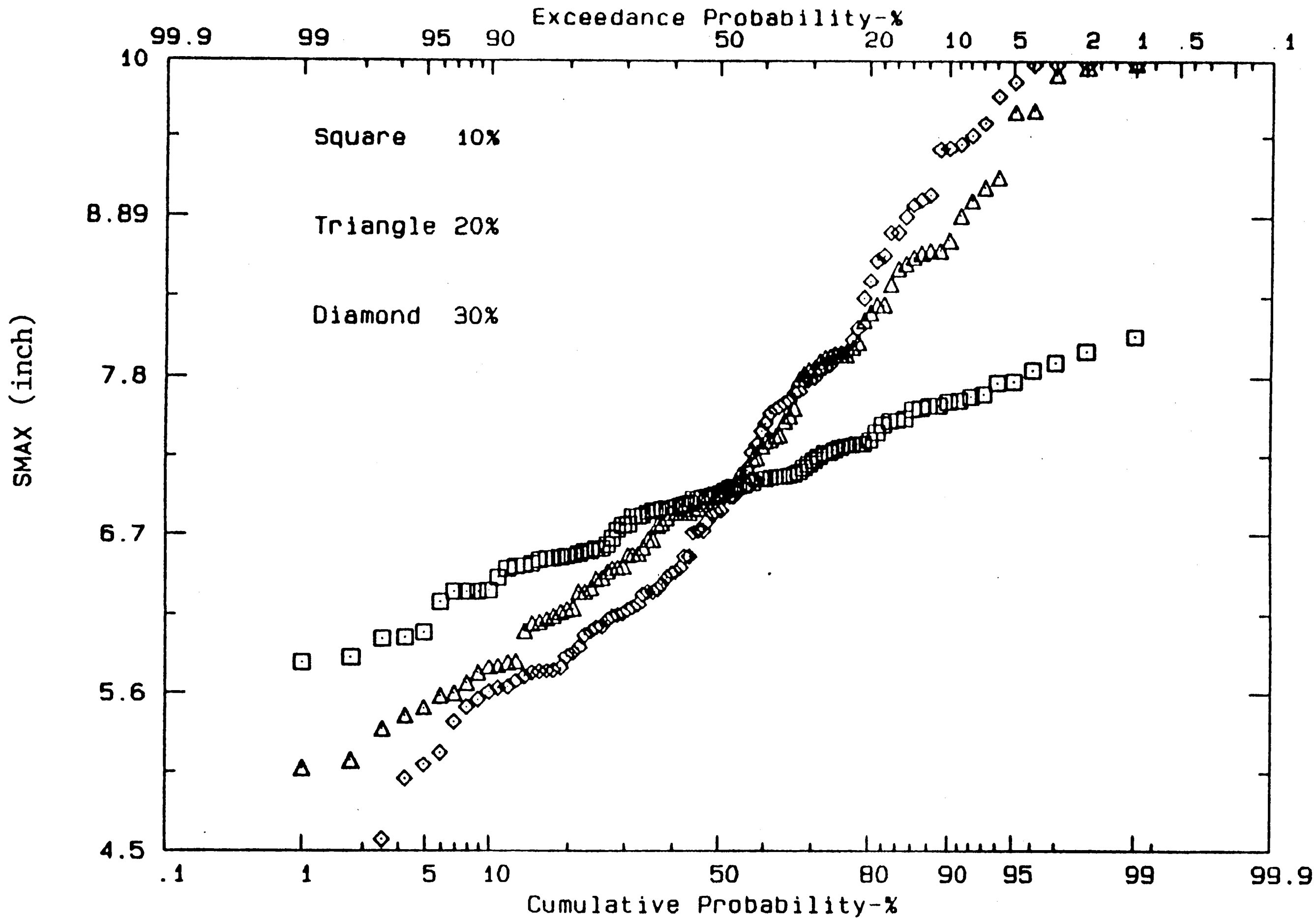

Figure 4. Distribution of SMAX Due to Lognormal Error in Rainfall (OLS)

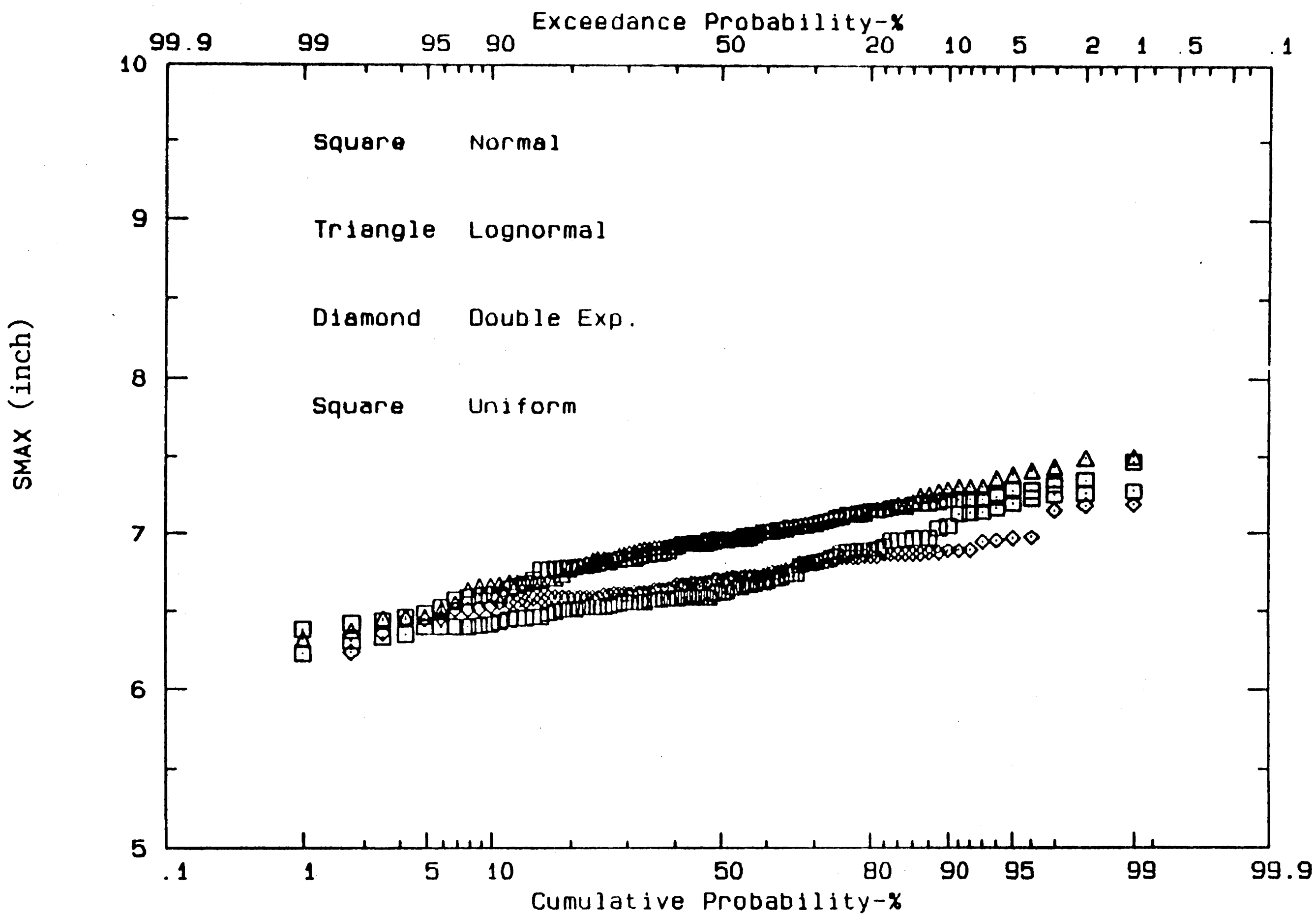

Figure 5. Distribution of SMAX Due to 10% Different Error in Flow (MLE)

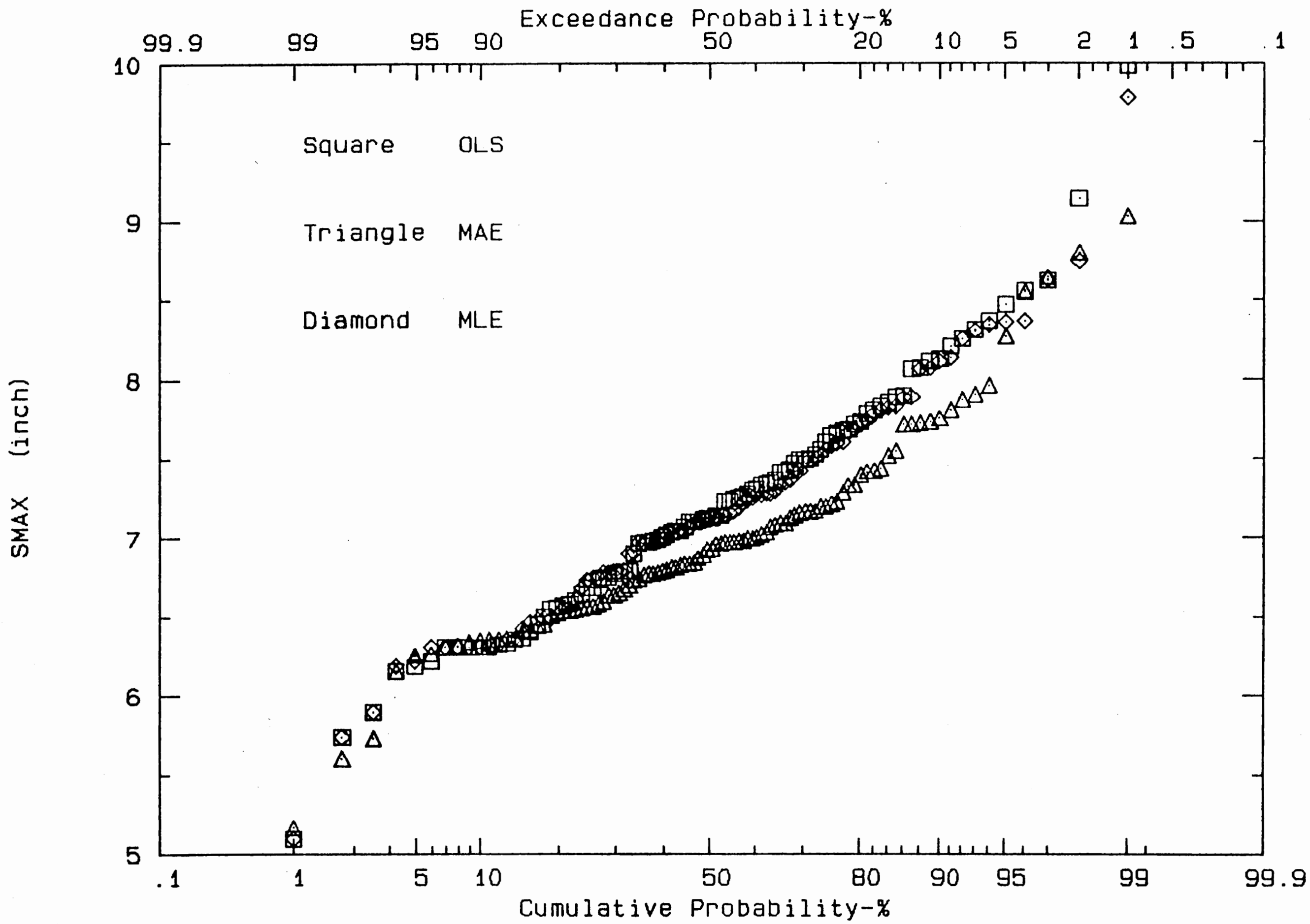

Figure 6. Distribution of SMAX Due to 20% Double Exponential Error in Rainfall

represented by a lognormal distribution.

Parameter REMX

Based on Kolmogorov-Smirnov test at 5% significance level both the normal and lognormal distributions can be assigned (Tables 17 through 21). Figures 7 through 10 depict the distribution of REMX. Irrespective of error distributions, D2 values were consistently lower than D1 in most of the samples. Therefore, Lognormal distribution was selected to approximate the parameter REMX.

Parameter SC1

D1 and D2 values for SC1 (Tables 17 through 21) are similar for all the error models. Both the normal and lognormal distributions can either be accepted or rejected. Figures 11 through 15 show the distribution of SC1. The lognormal distribution is assigned to SC1 because of its interaction with parameter SMAX.

Concluding Remarks

The parameters SMAX, REMX and SC1 can be approximated by either normal or lognormal distribution. However, the lognormal distribution seems to be a better approximation. Since SMAX and SC1 were found to be correlated, a bivariate lognormal distribution would be appropriate to preserve their correlation structure. The distribution of the error model had no impact on the resulting distribution of the parameters. It was also observed that the estimation techniques had very little effect on the distribution of parameters. The error variance did not influence the form of the parameter distribution but changed its distributional parameters depending on degree of error variations.

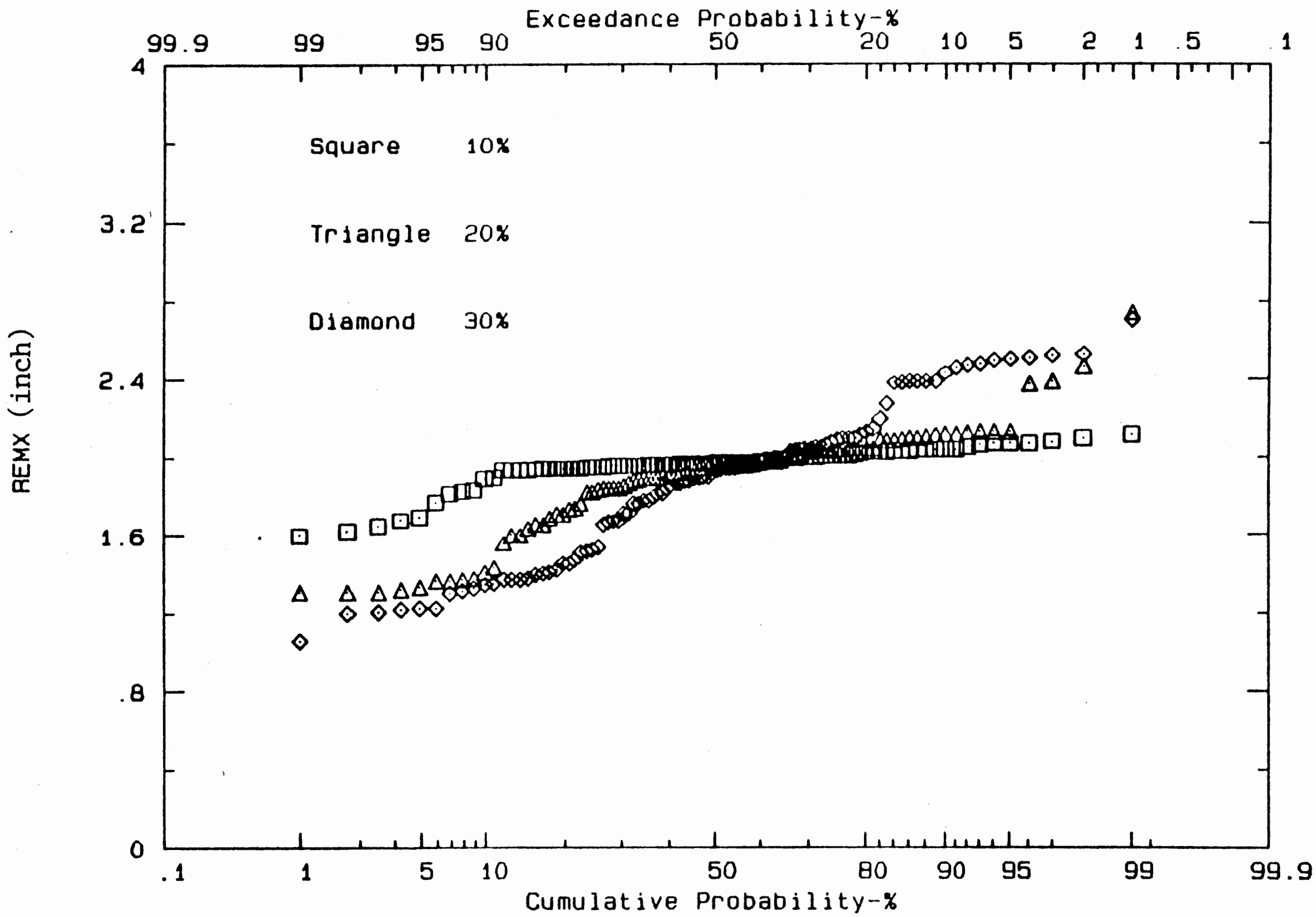

Figure 7. Distribution of REMX Due to Uniform Error in Flow (MLE)

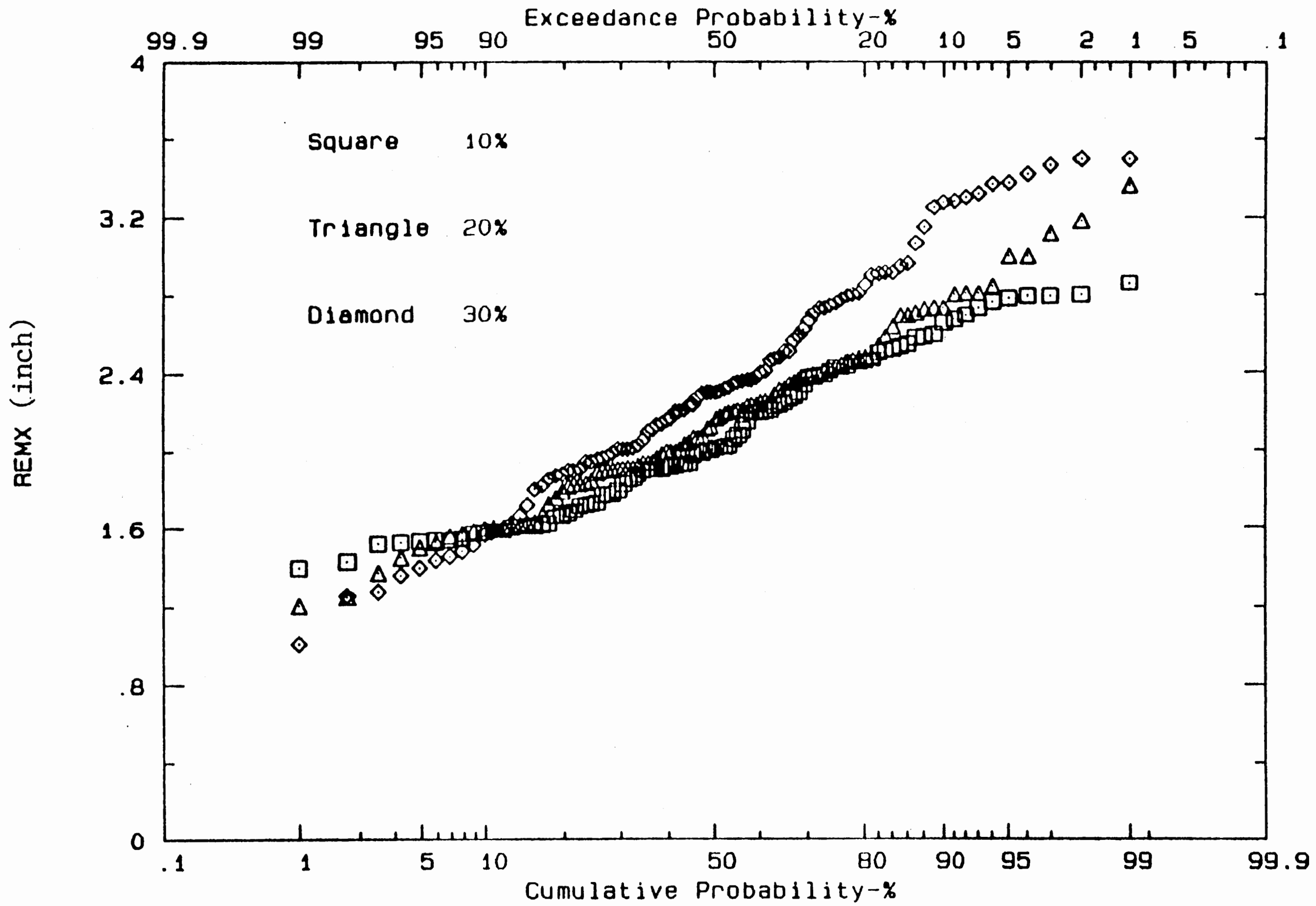

Figure 8. Distribution of REMX Due to Double Exponential Error in Rainfall (MAE)

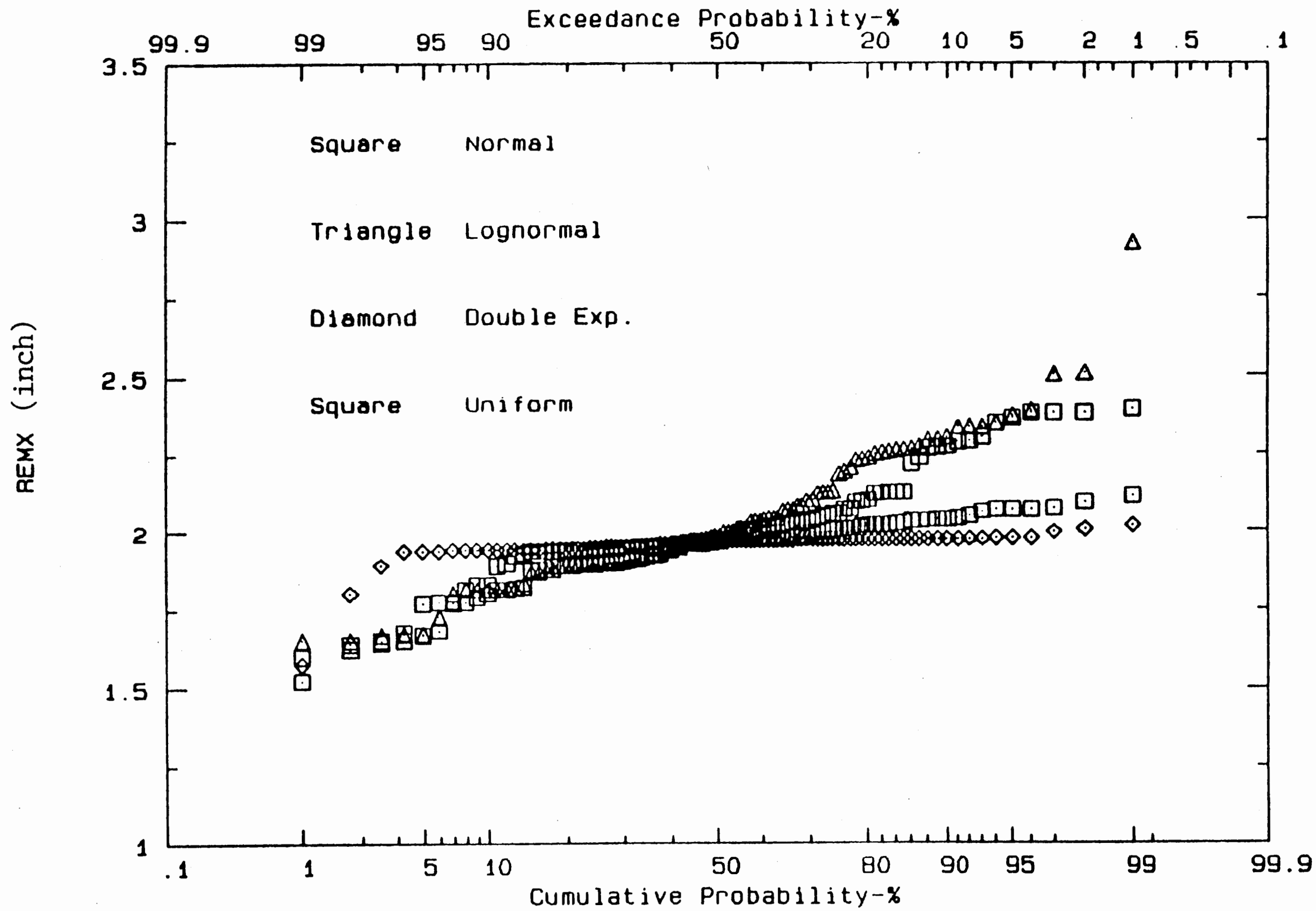

Figure 9. Distribution of REMX Due to 10% Different Error in Flow (OLS)

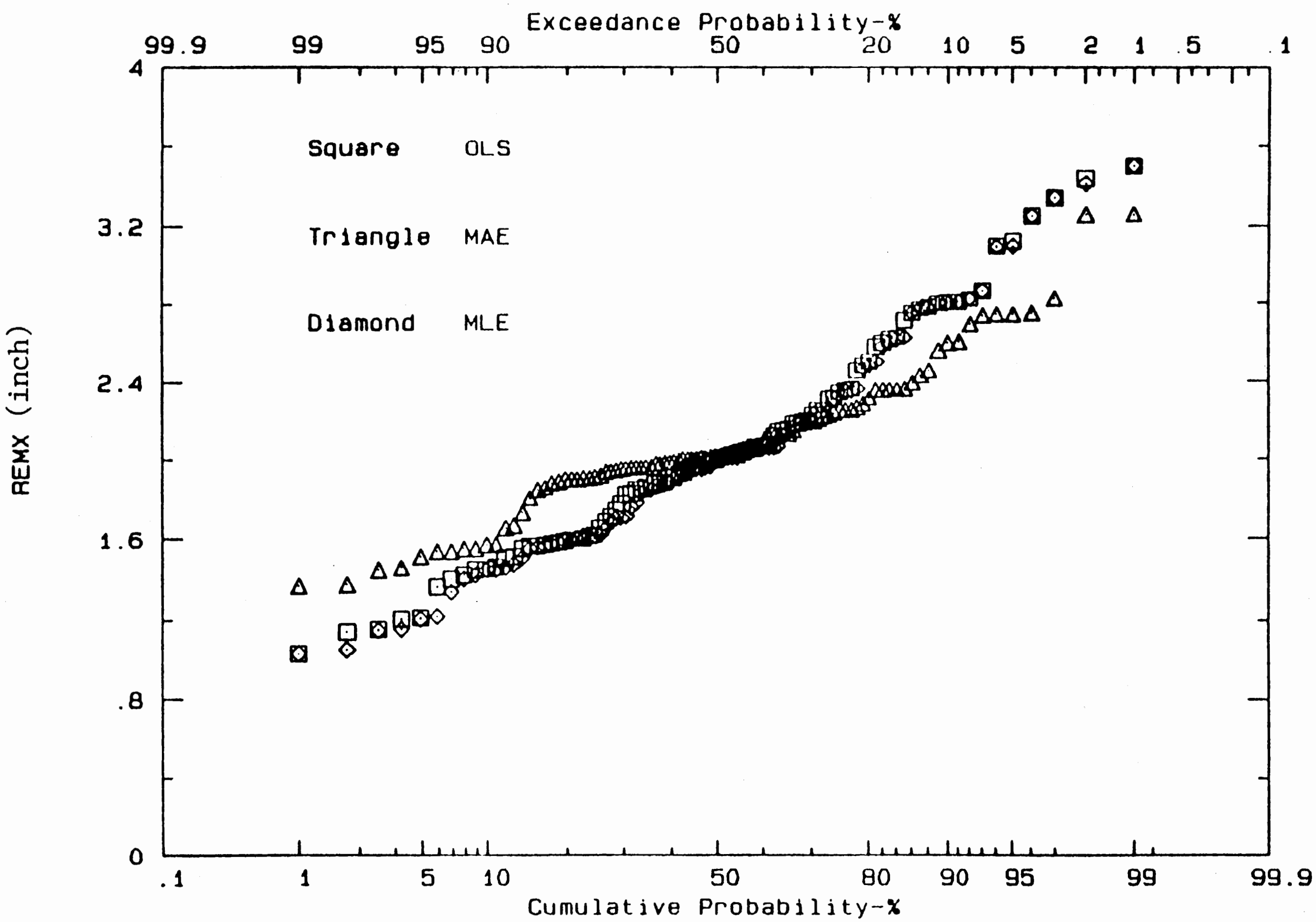

Figure 10. Distribution of REMX Due to 10% Normal Error in Rainfall

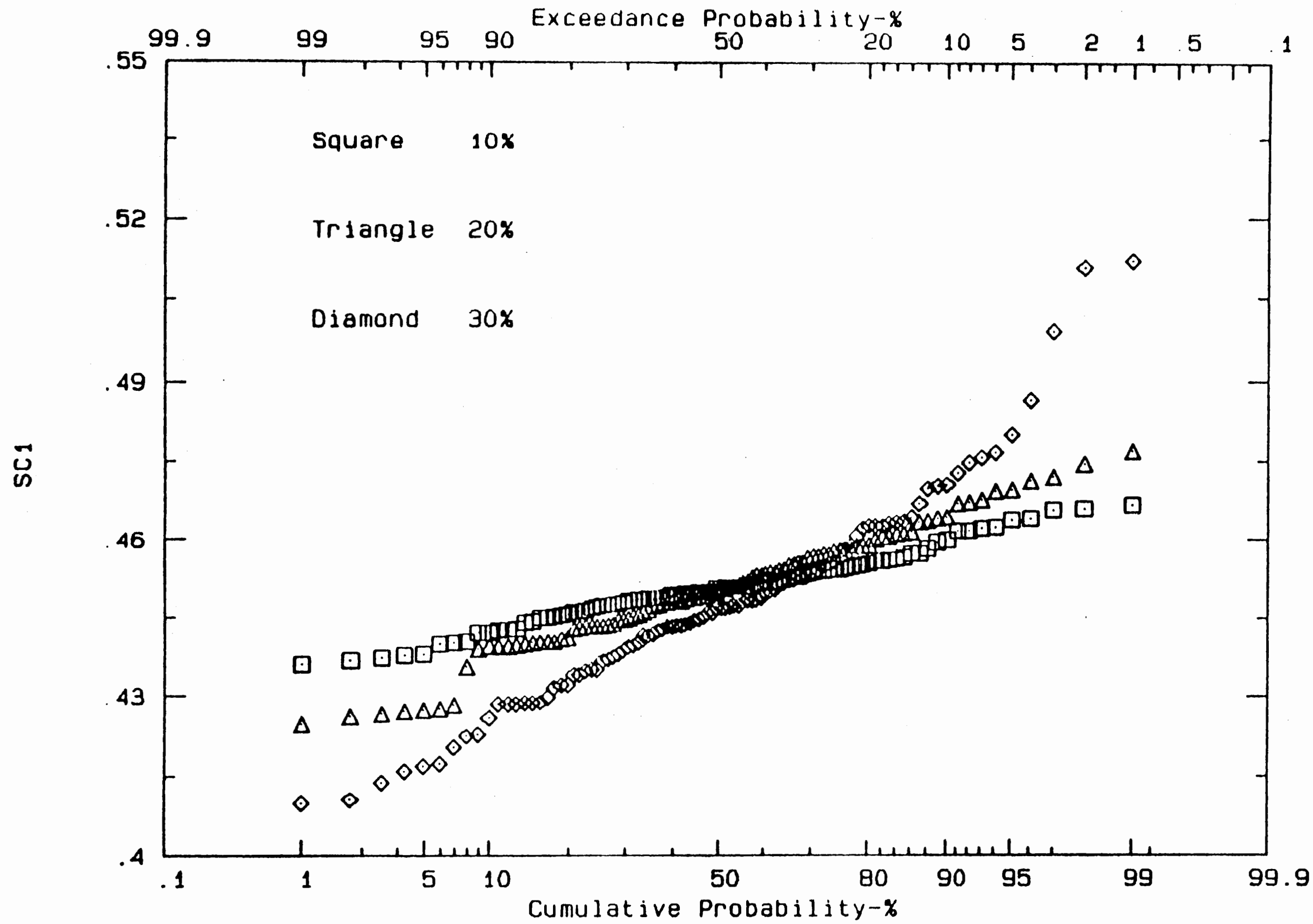

Figure 11. Distribution of SC1 Due to Lognormal Error in Flow (OLS)

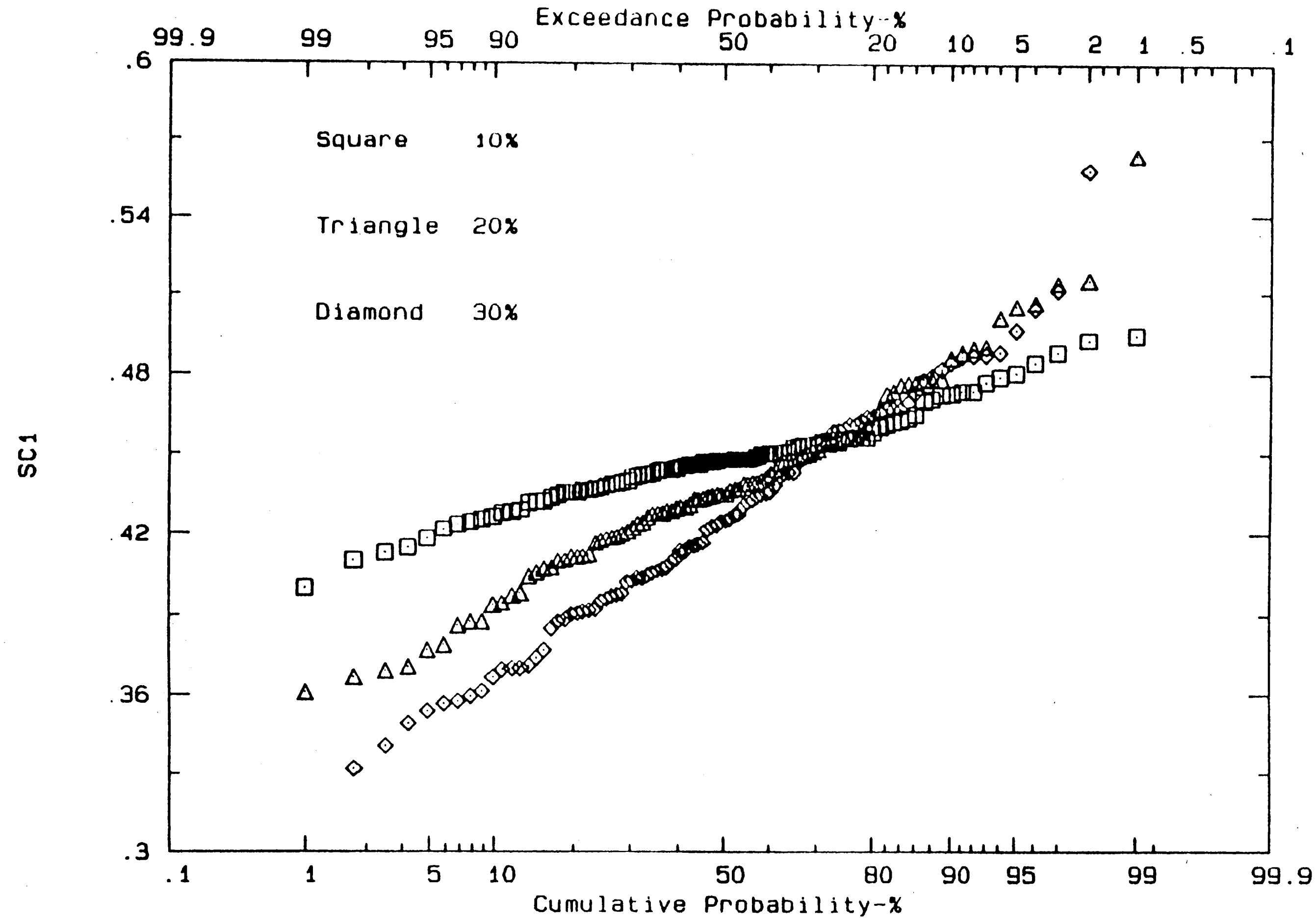

Figure 12. Distribution of SC1 Due to Uniform Error in Rainfall (MLE)

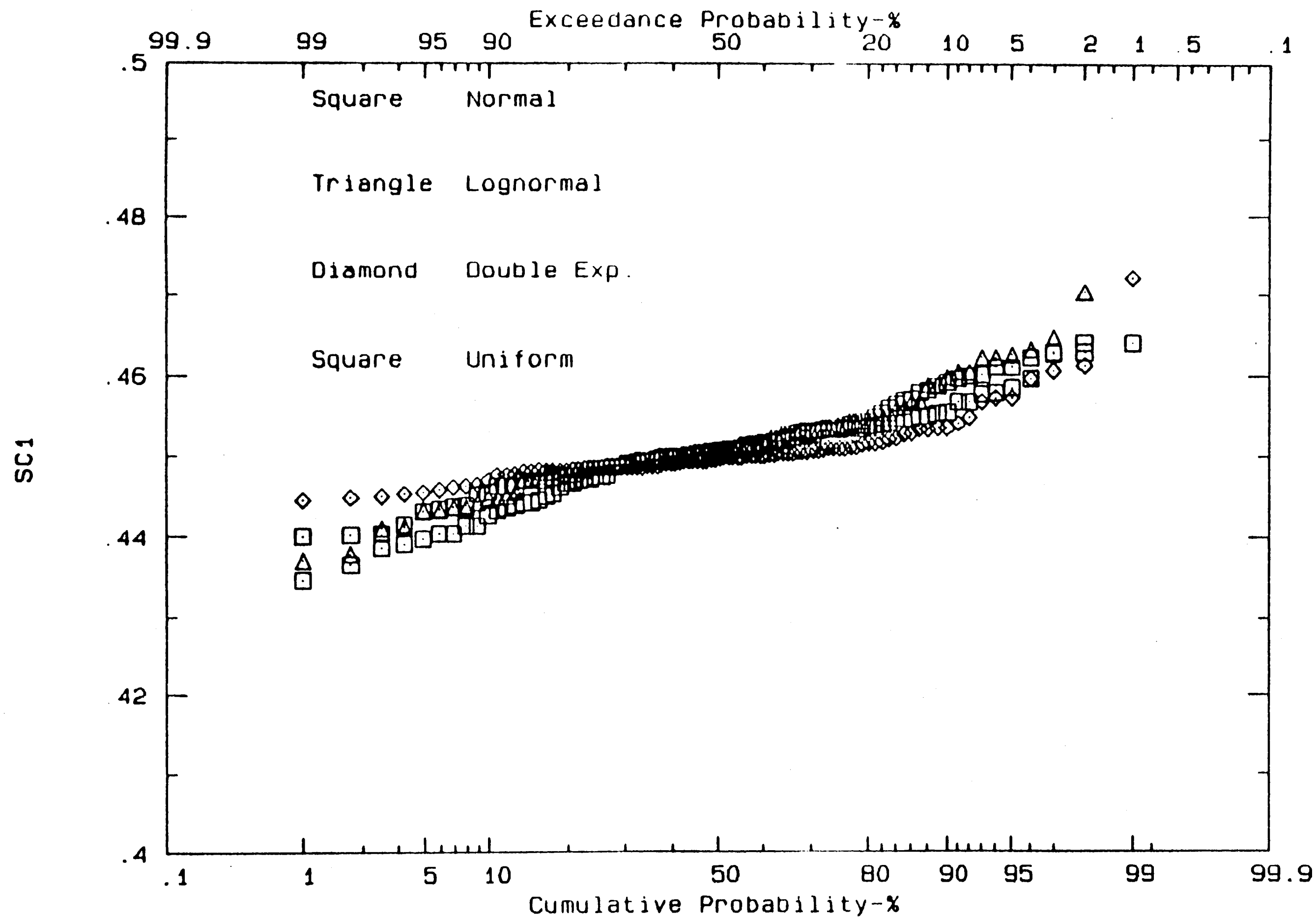

Figure 13. Distribution of SC1 Due to 10% Different Error in Flow (MAE)

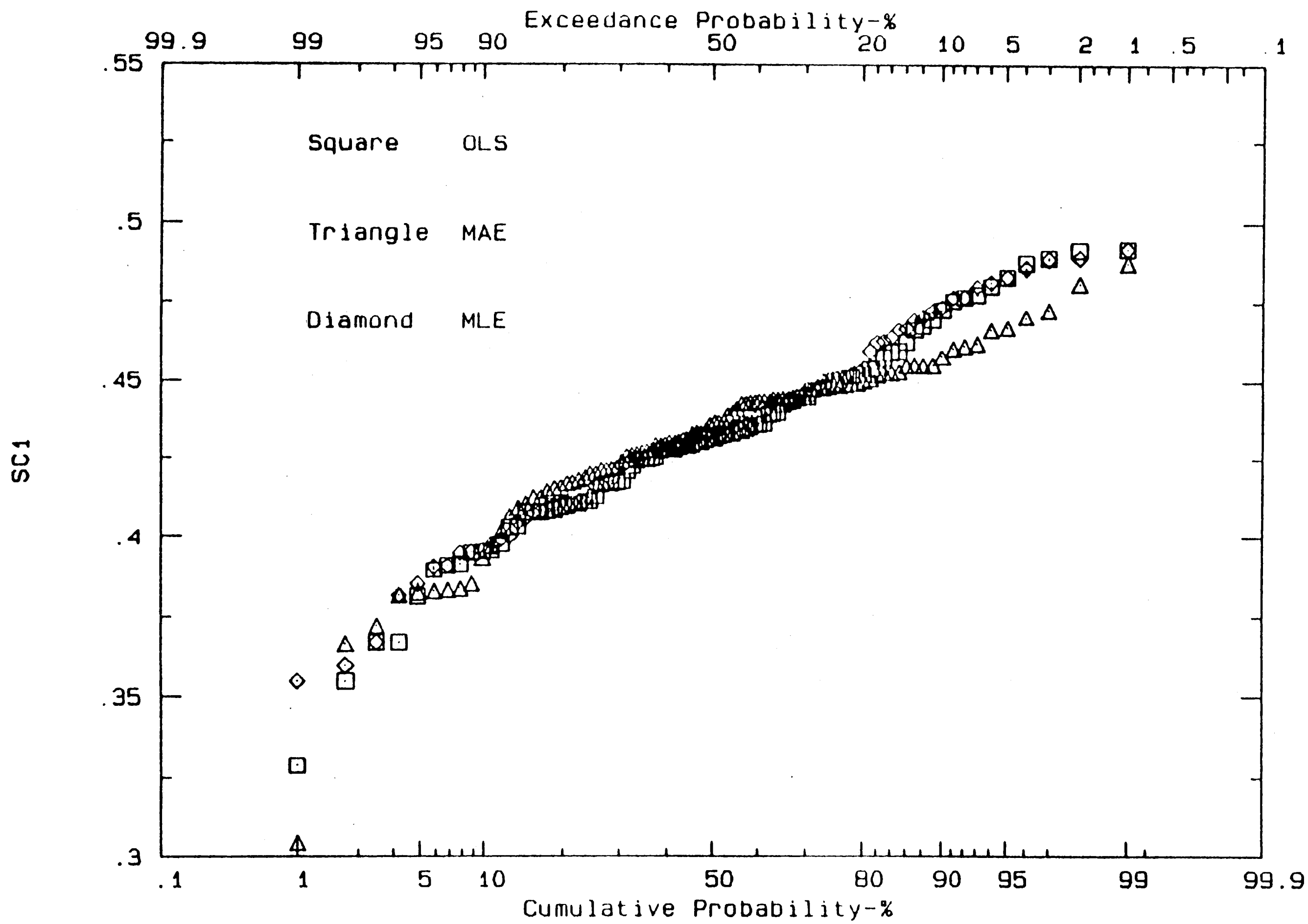

Figure 14. Distribution of SC1 Due to 20% Normal Error in Rainfall

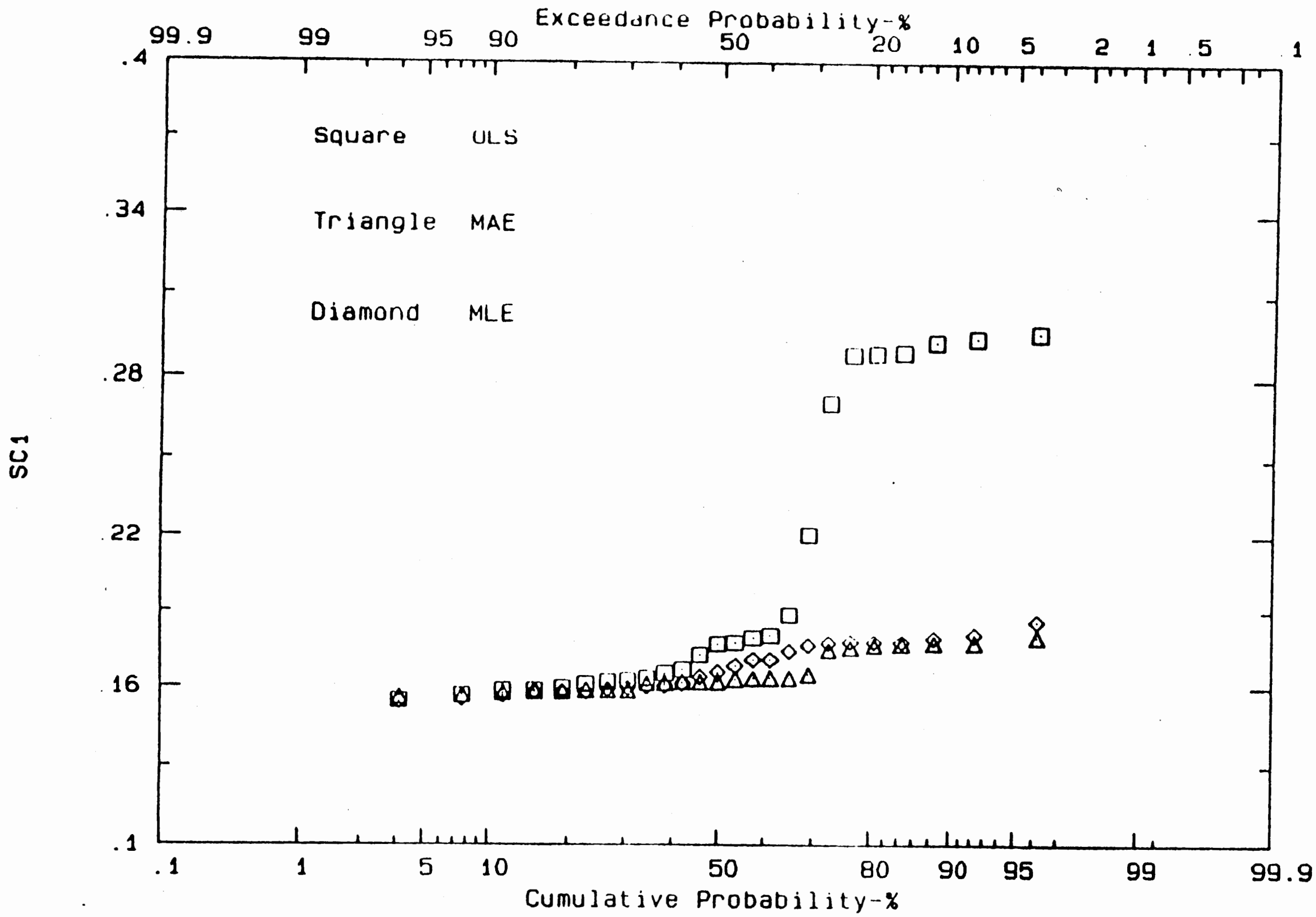

Figure 15. Distribution of SC1 Due to 10% Correlated Error (Correlation = 0.5) in Rainfall

Uncertainty Analysis

The uncertainty analysis in this section was based on 100 Monte Carlo simulations of the PRMS model. For each iteration, parameters were selected from lognormal distributions. Distributions were specified by their true values as mean with a 20% coefficient of variations (CV). A correlation coefficient of -0.5 was assumed between SMAX and SC1. Listing of the computer program for bivariate generation of lognormally distributed pairs of SMAX and SC1 is given in the Appendix C (Matalas, 1967 and Haan, 1977). REMX was drawn from a univariate lognormal distribution. Generated parameters and mean annual runoff values were also included in the Appendix C.

Table 22 shows the summary statistics of the generated data. It is evident that the distributional properties of the parameters were maintained during the generation process. Based on average parameter values, the mean annual runoff was 2.06 cfs days. However considering the parameters as random variables resulted in mean annual runoff of 2.12 cfs days with a standard deviation of 1.07 cfs days (Table 22). Thus 20% variations in the parameters was translated to a 36% coefficient of variation of model output. The result of this error analysis is in agreement with previous research. O'Neill et al. (1980) showed prediction errors were 10 times greater than parameters error.

Both normal and lognormal distributions were tested to describe the cumulative probability distribution of the mean annual runoff. Table 23 shows the maximum deviations of the mean annual runoff from both the above distributions. The critical value of the Kolmogorov-Smirnov test statistic, K is 0.14 at 5% significance level. The mean annual runoff is best described by a normal distribution. Figure 16 shows the probability distribution of the annual runoff values. Now probabilities can be assigned to various runoff estimates. A 90%

TABLE 22

STATISTICAL SUMMARY OF GENERATED DATA

	SMAX (in)	REMX (in)	SC1	MEAN RUNOFF (cfs days)
Minimum	3.940	1.160	0.230	0.560
Maximum	9.700	3.100	0.710	4.190
Mean	6.803	1.993	0.456	2.120
SD	1.263	0.397	0.080	0.764
CV	0.186	0.199	0.166	0.360

ρ(SMAX,SC1) = - 0.46 in generated data

TABLE 23

KOLMOGOROV-SMIRNOV TEST STATISTICS FOR MEAN ANNUAL RUNOFF

DISTRIBUTION	D	K
Normal	0.083	0.140
Lognormal	0.054	0.140

D maximum deviation from the distribution

K the critical value of the K-S test

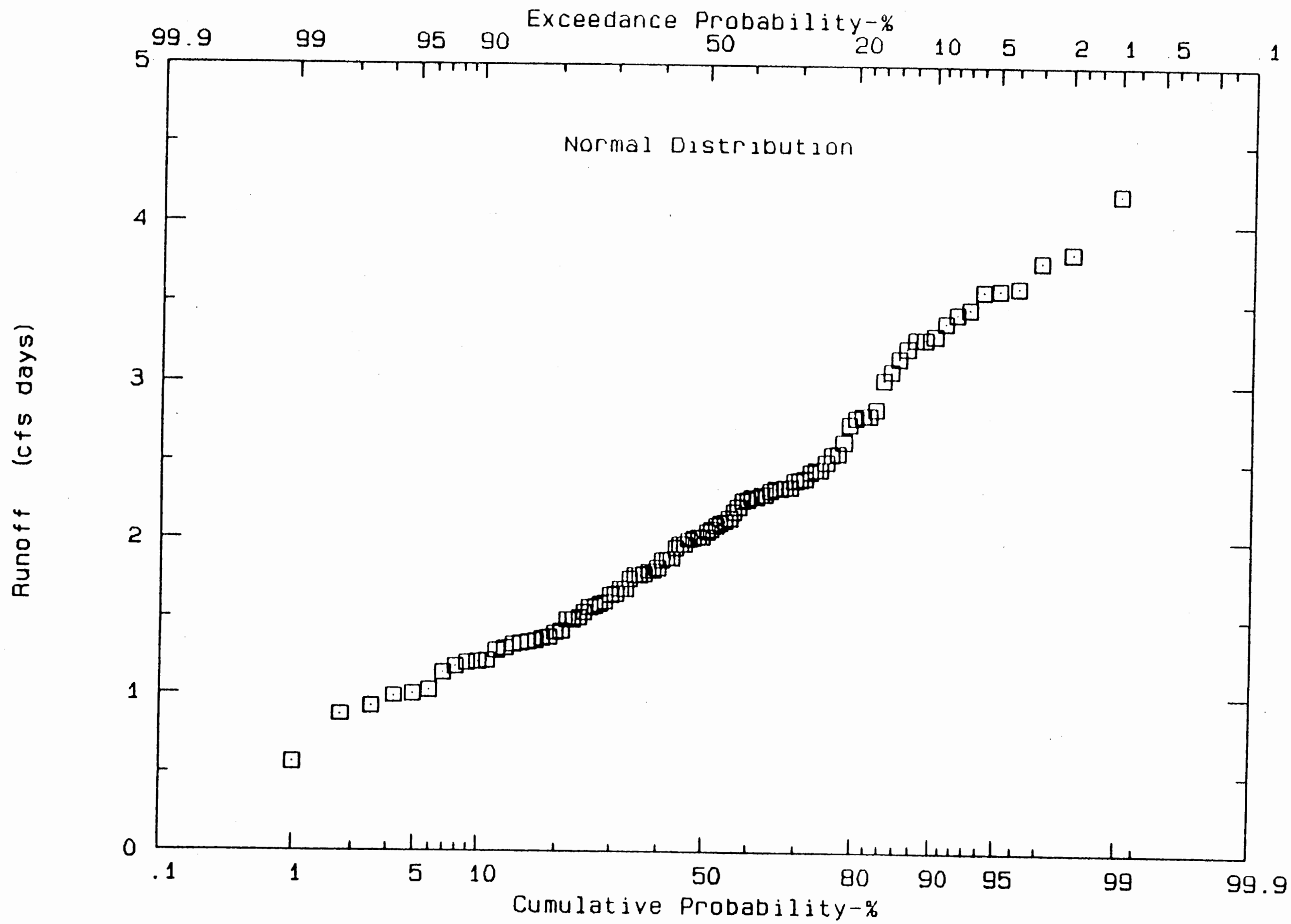

Figure 16. Distribution of the Mean Annual Runoff

confidence interval on the mean flow is between 0.87 cfs days and 3.37 cfs days.

Small errors on the parameters can lead to significant uncertainty in model predictions. In this example, uncertainty in the model parameters was transformed to almost two times more uncertainty in the model output. Thus point estimates of model output is not enough to convey the complete information. At minimum, at least the standard deviation of the model output should be provided. Nevertheless, a probability study would help to describe the complete picture.

CHAPTER VII

SUMMARY AND CONCLUSIONS

Summary

Hydrologic models are not yet developed to the extent that all the parameters are physically measurable. As such, parameters are estimated using observed data, either directly or indirectly that can be considered as random data or data having a random component. Thus parameters may be considered as random variables characterized by probability density functions. Parameter estimation and model application should recognize the probabilistic nature of estimated model parameters.

The objectives of this study were (1) to evaluate the impact of various error distributions in rainfall and streamflow on parameter estimates and (2) to evaluate parameter estimation techniques in the presence of errors in input data.

The USGS Precipitation Runoff Modeling Systems (PRMS) was used in this research. Hydrologic data for four years (1974-1977) from Chickasha R-6 watershed was utilized. The watershed is located in the Washita river basin of Southcentral Oklahoma. The model contains more than 50 parameters. Three of the more sensitive parameters, SMAX, REMX and SC1, were chosen for this investigation.

Representative values for the parameters were found from the initial run of the model. This set of parameters was assumed to be true values for the watershed. Then four years of rainfall and air temperature data were used in PRMS to generate an error-free runoff sequence.

Parameter estimation techniques considered for this study were the ordinary least squares (OLS), method of absolute errors (MAE) and maximum likelihood technique (MLE). All these three methods exactly recovered the true parameter values when both rainfall and runoff data were error-free.

The error models used for this research were normal, lognormal,double exponential, uniform and correlated errors. Errors were introduced separately into each value of rainfall and runoff. Error contaminated data were then used for parameter estimation. There were altogether 105 combinations of Monte Carlo simulations for 5 error models and three parameter estimation techniques.

Errors in precipitation caused more uncertainty in parameters than the error contaminated runoff records. This observation was true for all the error models. Variations in the parameters also increased with an increasing level of errors in the data. For mixed errors, parameter variances were higher than the corresponding level of error in either rainfall or in runoff records.

Parameter variances and biases were used to evaluate the estimation techniques. Performance of an estimation technique was judged based on its ability to produce estimates with lower variance and bias than any other method. MSE was an indicator of this. The method of absolute errors performed better in all the error models except the correlated error situation. MAE is known to perform better in non-normal errors. However, its superiority over OLS in normal error case was perhaps due to low flow sequences. Generally high flow values dominate the OLS objective function. Performance of OLS and MLE were similar for error models other than the correlated error. MLE performed better in correlated error case because its objective function was based on the error structure of the model residuals. Due to long zero flow sequence, the desired degree of correlation could not be achieved in the data. As a result, MAE and OLS also worked well for correlated error model.

Parameters SMAX and SC1 were found to be highly correlated. There were very little interaction between SMAX and REMX and between REMX and SC1. Correlation between SMAX and SC1 results from their functional relationship in the model.

The distribution of the error model did not influence the resulting distribution of the parameters. Parameter SMAX, REMX and SC1 were approximated with lognormal distribution. Since SMAX and SC1 were correlated a bivariate lognormal distribution was used to preserve their correlated structure.

Uncertainty in the model predictions was analyzed using Monte Carlo simulations. Parameters were generated from lognormal distributions maintaining their distributional properties. Then 100 independent model outputs of mean annual flow were generated. The mean annual runoff was found to be normally distributed. Another observation was the higher uncertainty in model predictions than the uncertainty in the model parameters.

Conclusions

Based upon the results of this research the following conclusions can be drawn.

1. Errors in the precipitation records introduced more uncertainty to parameter estimates than errors in runoff data.

2. Parameter uncertainty also increased with increasing level of errors in the data used for optimization.

3. The method of absolute errors was found superior for all the error models excluding the correlated errors. Better estimates were in the context of lower bias and variance.

4. Maximum likelihood technique was superior in the correlated error model.

5. Performances of the ordinary least squares and maximum likelihood techniques were similar in normal as well as non-normal error models.

6. The distribution of the error model had no impact on the resulting distribution of the parameters except for the parameter REMX in the presence of uniform and double exponential errors.

7. Smaller uncertainty in the parameters caused larger uncertainty in the model predictions.

8. Violation of the assumption of uncorrelated errors introduced significant errors in estimated parameters.

Recommendations for Further Research

The following topics are suggested for future investigation.

1. The results of this study may be model specific. Therefore similar studies should be conducted for other rainfall-runoff models.

2. Superiority of the method of absolute errors for non-normal error models is obvious. However for the normal error model, MAE performed better than OLS perhaps due to low flow sequences. For future study, larger watersheds with high flow values should be considered to substantiate this point.

3. More studies on the impact of parameter distributions on estimated flow distributions should be conducted.

BIBLIOGRAPHY

1. Aitken, A. P. 1973. Assessing systematic errors in rainfall-runoff models. Journal of Hydrology 20: 121-136.

2. Amorocho, J. and G. T. Orlob. 1961. An evaluation of the inflow-runoff relations in hydrologic studies. University of California Water Resources Center Report No. 41, Berkeley.

3. Amorocho, J. and B. Epsildora. 1973. Entropy in the assessesment of uncertainty in hydrologic systems and models. Water Resources Research 9(6): 1511-1522.

4. Bard, Y. 1974. Nonlinear Parameter Estimation, Academic Press, NewYork, NY.

5. Beard, L. R. 1967. Optimization techniques for hydrologic engineering. Water Resources Research 3(3): 809-815.

6. Beck, J. V. and K. J. Arnold. 1977. Parameter Estimation In Engineering And Science. John Wiley Inc. NewYork, NY.

7. Beck, M. B. 1987. Water quality modeling : a review of the analysis of uncertainty. Water Resources Research 23(8): 1393- 1442.

8. Bodo, B. and T. E. Unny. 1976. Model uncertainty in flood frequency analysis and frequency-based design. Water Resources Research 12(6): 1109-1117.

9. Boughton, W. F. 1965. A new simulation technique for estimating catchment yield. Water Resources Lab. Report No. 78 , University of New South Wales, Australia.

10. Box, G. E. P. and G. M. Jenkins. 1970. Time Series Analysis : Forcasting and Control. Holden Day, San Francisco, CA, 575pp.

11. Box, G. E. P. and W. J. Hill. 1974. Correcting inhomogeneity of variance with power transformation weighting. Technometrics 16(3): 385-389.

12. Chapman, T. G. 1970. Optimization of a rainfall-runoff model for an arid zone catchment, UNESCO Publication 96: 126-143.

13. Chery, D. L. 1966. Design and tests of a physical watershed model. Journal of Hydrology 4(3): 224-235.

14. Chow, V. T. 1954. The log-probability law and its engineering applications. Proceedings American Society of Civil Engineers 80: 536-1 to 536-25.

15. Chow, V. T., D. R. Maidment and L. W. Mays. 1988. Applied Hydrology. McGraw Hill Book Company Inc., NewYork, NY.

16. Clarke, R. T. 1973. A review of some mathematical models used in hydrology, with observations on their calibration and use. Journal of Hydrology 19(1): 1-20.

17. Clarke, R. T. 1973. Mathematical models in hydrology. FAO Publication No. 19, Rome.

18. Cover, K. A. and T. E. Unny. 1986. Application of computer intensive statistics to parameter uncertainty in streamflow synthesis. Water Resources Bulletin 22(3): 495-507.

19. Davis, D. R., K. C. Kisiel and L. Duckstein. 1972. Bayesian decision theory applied to design in hydrology. Water Resources Research 8(1): 33-41.

20. Davis, R. B. and B. Hutton. 1975. The effects of errors in the independent variables in linear regression. Biometrika 62(2): 383-391.

21. Dawdy, D. R. and T. O'Donnel. 1965. Mathematical models of catchment behavior. Journal of Hydraulics Division ASCE HY4: 123-127.

22. Dawdy, D. R. and T. H. Thompson. 1966. Digital computer simulation in hydrology. Journal of American Water Works Association: 685-688.

23. Dawdy, D. R. and J. M. Bergmann. 1969. Effect of rainfall variability on streamflow simulation. Water Resources Research 5(5): 958-966.

24. Deininger, R. A. 1969. Linear programming for hydrologic analysis. Water Resources Research 5: 1105-1110.

25. Diskin, M. H. and E.Simon. 1977. A procedure for the selection of objective functions for hydrologic simulation models. Journal of Hydrology 34: 129-149.

26. Dooge, J. C. I. 1968. The hydrologic cycle as a closed system. Bull. Int. Sci. Hydrol. 13(1) pp 58-68.

27. Draper, N. R. and H. Smith. 1966. Applied Regression Analysis. John Wiley Inc. NewYork NY.

28. Eagleson, P. W., R. Mejia and F. March. 1966. Computation of optimum realizable unit hydrographs. Water Resources Research 2(4): 755-764.

29. Edwards, D. R. 1988. Incorporating parametric uncertainty into flood estimation methodologies for ungaged watersheds and for watersheds with short records. Unpublished Ph.D. dissertation, Oklahoma State University.

30. Federer, A. C. and D. Lash. 1978. BROOK: A hydrologic simulation model for eastern forests. Water Resources Research Center Report No. 19, University of New Hampshire, Durham, NH.

31. Fisher, W. D. 1961. A note on curve fitting with minimum deviations by linear programming. Journal of the American Statistical Association 56: 359-362.

32. Fourier, J. B. J. Oeuvres De Fourier, edited by G. Darboux, Paris: Gauthier-Villars, 1888-90, Vol 2: 324-325.

33. Freund, J. 1962. Mathematical Statistics. Prentice Hall Inc., Eaglewood Cliffs, NJ.

34. Gardner, R. H., D. D. Huff, R.V. O'Neill, J. B. Mankin, J. Carney and J. Jones. 1980. Application of error analysis to a marsh hydrology model. Water Resources Research 16(4): 659-664.

35. Goldberger, A. S. 1964. Econometrics Theory. John Wiley and Sons, NewYork, NY.

36. Gupta, V. K. and S. Sorooshian. 1983. Uniqueness and observability of conceptual rainfall-runoff model parameters : the percolation process examined. Water Resources Research 19(1): 269-276.

37. Haan, C. T. 1972. Adequacy of hydrologic records for parameter estimation. Journal of Hyd. Div. ASCE, HY8: 1387 -1393.

38. Haan, C. T. 1977. Statistical Methods in Hydrology. The Iowa State University Press, Ames, IA.

39. Haan, C. T. 1989. Parameter uncertainty in hydrologic modeling. Transaction of ASAE 32(1): 137-146.

40. Haan, C. T. and D. M. Allen. 1972. Comparison of multiple regression and principal component regression for predicting water yields in Kentucky. Water Resources Research 8(6): 1593-1596.

41. Hamon, W. R. 1961. Estimating potential evapotranspiration. Proceedings of the American Society of Civil Engineers 87(HY3): 107-120.

42. Ibbitt, R. P. 1972. Effect of random data error on the parameter values for a conceptual model. Water Resources Research 8(1): 70-78.

43. Ibbitt, R. P. and T. O'Donnell. 1971. Fitting methods for conceptual catchment models. Journal of Hyd. Div. ASCE. 97(HY9): 1331-1342.

44. Jackson, D. R. and G. Aron. 1971. Parameter estimation in hydrology : the state of the art. Water Resources Bulletin 7(3): 457-472.

45. Johnson, J. 1972. Econometric Methods. McGraw Hill Inc., NewYork, NY.

46. Johnston, P. R. and D. Pilgrim. 1976. Parameter optimization for watershed models. Water Resources Research 12(3): 477-486.

47. Koch, Roy W. and Gary M. Smillie. 1986. Bias in hydrologic prediction using log-transformed regression models. Water Resources Bulletin 23(5): 717-723.

48. Kuczera, G. 1982. On the relationship between the reliability of parameter estimates and hydrologic time series data used in calibration. Water Resources Research 18(1): 146-154.

49. Kuczera, G. 1982. Combining site specific and regional information : an empirical bays approach. Water Resources Research 18(2): 306-314.

50. Kuczera, G. 1982. Improved inference of catchment model parameters. Report 1/1982. Civil Engineering Dept. Monash University, Clayton, Victoria.

51. Kuczera, G. 1983. Improved parameter inference in catchment models. 1.evaluating parameter uncertainty. Water Resources Research 19(5): 1151-1162.

52. Leavesley, G. H., R. W. Lichty, B. M. Troutman and L. G. Saindon. 1983. Precipitation-Runoff Modeling System: User's Manual, U.S. Geological Survey Water Resources Investigation 83-4238, 206pp.

53. Linsley, R. K. 1982. Rainfall-runoff models : An overview, pp 3-22. In V.P.Singh (Ed.) Rainfall-Runoff Relationship. Water Resources Publication.

54. Little, K. W. and D. T. Lauria. 1989. Water quality model calibration: A comparison of input and output error criteria. Water Resources Bulletin 25(4): 755-764.

55. Lynch, J. A. and W. E. Sopper. 1974. Watershed behavior under controlled simulated rainfall. Technical Report: Institute for Research on Land and Water Resources, Pennsylvania State University, University Park.

56. Magette, W. L., V. O. Stanholz and J.G.Carr. 1976. Estimating selected parameters for the Kentucky watershed model from watershed characteristics. Water Resources Research 12(3): 472-476.

57. Mamisao, J. P. 1952. Development of an agricultural watershed by simulation. M.S. Thesis, Iowa State College, Ames.

58. Mandeville, A. N., P. E. O'Connell, J.V.Sutcliffe and J.E.Nash. 1970. River flow forcasting through conceptual models. Part II, The Ray catchment at Grendon Underwood. Journal of Hydrology 11: 109-128.

59. Matalas, N. C. 1967. Mathematical assessment of synthetic hydrology. Water Resources Research 3(4): 937-945.

60. Mays, L. W. and S. L. Coles. 1980. Optimization of unit hydrograph determination. Journal of American Society of Civil Engineers 106(HY1): 85-97.

61. Mood, A. M., F. A. Graybill and D. C. Boes. 1974. Introduction To The Theory Of Statistics. McGraw Hill Inc., NewYork, NY.

62. Moore, R. J. and R. T. Clarke. 1981. A distribution function approach to rainfall-runoff modeling. Water Resources Research 17(5): 1367 - 1382.

63. Murray, F. W. 1967. On the computation of saturation vapor pressure. Journal of Applied Meteorology 6: 203-204.

64. Nash, J. E. and J. V. Sutcliffe. 1970. River flow forcasting through conceptual models, Part I : A discussion of principles. Journal of Hydrology 10(3): 282-290.

65. O'Neill, R. V., R. H. Gardner and J. B. Mankin. 1980. Analysis of parameter error in a nonlinear model. Ecological Modeling 8(1980): 297-311.

66. Pickup, G. 1977. Testing the efficiencies of algorithms and strategies for automatic calibration of rainfall-runoff models. Hydrologic Science Bulletin 22(2): 257-274.

67. Plinston, D. T. 1972. Parameter sensitivity and interdependence in hydrologic models. Institute of Hydrology, Wallingfor, Berkislir.

68. Riggs, H. C. 1985. Streamflow Characteristics. Elsevier Science Publishing Company Inc. NewYork, NY.

69. Riley, J. P., D. G. Chadwick and E. K. Israelson. 1967. Application of an electronic analogue computer for the simulation of hydrologic events on a southwest watershed. Utah State University, Utah.

70. Rosenbrock, H. H. 1960. An automatic method of finding the greatest or least value of a function. The Computer Journal 3: 175-184.

71. Sage, A. D. and J. L. Melsa. 1971. Estimation Theory With Application To Communication And Control. McGraw Hill Book Company,NewYork.

72. Sorooshian, S. and J. A. Dracup. 1978. Consideration of stochastic properties in parameter estimation of hydrologic rainfall-runoff models. Technical Report UCLA-ENG 7852, University of California, Los Angeles.

73. Sorooshian, S. 1980. Parameter estimation of rainfall-runoff models with heteroscedastic streamflow errors : the noninformative data case. Journal Of Hydrology 52: 127-138.

74. Sorooshian, S. and John A. Dracup. 1980. Stochastic parameter estimation procedures for hydrologic rainfall-runoff models : correlated and heteroscedastic error cases. Water Resources Research 16(2): 430-442.

75. Sorooshian, S., V. K. Gupta. and J. L. Fulton. 1983. Evaluation of maximum likelihood parameter estimation techniques for conceptual rainfall-runoff models : influence of calibration data variability and length on model credibility. Water Resources Research 19(1): 251-259.

76. Singh, K. P. 1976. Unit hydrographs, a comparitive study. Water Resources Bulletin 12(2): 381-392.

77. Troutman, B. M. 1982. Analysis of input errors in precipitation-runoff models using regression with errors in the independent variables. Water Resources Research 18(4): 947-964.

78. Troutman, B. M. 1983. Runoff prediction errors and bias in parameter estimation induced by spatial variability of precipitation. Water Resources Research 19(3): 791-810.

79. Troutman, B. M. 1985. Errors and parameter estimation in precipitation-runoff modeling, 1.Theory. Water Resources Research 21(8): 1195-1213.

80. Troutman, B. M. 1985. Errors and parameter estimation in precipitation-runoff modeling, 2.Case Study. Water Resources Research 21(8): 1214-1222.

81. Tung, Y. K. and L. W. Mays. 1981. Reducing hydrologic parameter uncertainty. ASCE Journal of the Water Resources Planning And Management Division 107(WR1): 245-262.

82. Vicens, G. J., I. Rodriguez-Iturbe and J.C.Schaake,Jr. 1975. Bayesian generation of synthetic streamflows. Water Resources Research 11(6): 827-838.

83. U. S. Department of Agriculture. 1966. Hydrologic data for experimental agricultural watersheds. Miscellaneous Publication 1226.

84. U. S. Department of Agriculture. 1974. Hydrologic data for experimental agricultural watersheds. Miscellaneous Publication 1437.

85. U. S. Department of Agriculture. 1975. Hydrologic data for experimental agricultural watersheds. Miscellaneous Publication 1446.

86. U. S. Department of Agriculture. 1976. Hydrologic data for experimental agricultural watersheds. Miscellaneous Publication 1451.

87. U. S. Department of Agriculture. 1977. Hydrologic data for experimental agricultural watersheds. Miscellaneous Publication 1454.

88. USDA - SCS. 1978. Soil survey of Grady County, Oklahoma.

89. Warwick, John J. 1989. Interplay between parameter uncertainty and model aggregation error. Water Resources Bulletin 25(2): 275-283.
sll

90. Williams, B. J. and W. W. G. Yeh. 1983. Parameter estimation in rainfall-runoff models. Journal of Hydrology 63: 373-393.

91. Wolfe, P. M. and P. C. Koelling. 1983. Basic Engineering and Scientific Programs for IBM PC. Robert J. Brandy Company.

92. Wood, Eric F. 1976. An analysis of the effects of parameter uncertainty in deterministic hydrologic models. Water Resources Research 12(1): 925-932.

93. Wood, E. F. and I. Rodriguez-Iturbe. 1975. A Bayesian approach to analysing uncertainty among flood frequency models. Water Resources Research 11(6): 839-843.

94. Woolshiser, D. A. 1972. Uncertainties in deterministic models. Proceedings international symposium on uncertainties in hydrologic and water resources systems, Tucson, AZ.

95. Yeh, Williams G. 1986. Review of parameter identification procedures in groundwater hydrology : the inverse problem. Water Resources Research 22(2): 95-108.

96. Yen, B. C. 1970. Risks in hydrologic design of engineering projects. Journal Of Hydraulic Division, ASCE 96(HY4): 959-966.

97. Zahner, R. 1967. Refinement in empirical functions for realistic soil-moisture regimes under forest cover, in Sopper, W. E. and H. W. Lull eds., International symposium of forest hydrology, NewYork Pergamon Press: 261-274.

APPENDIXES

APPENDIX A

COMPUTER PROGRAMS FOR GENERATION OF RANDOM ERRORS

```
10 REM ********************************************************************
20 REM            GENERATION OF NORMAL RANDOM ERROR
30 REM ********************************************************************
40 REM INPUT NAME OF INPUT AND OUTPUT FILE
50 REM INPUT STANDARD DEVIATION OF ERROR ,S.
60 REM OUTPUT FILE IS COMPATIBLE TO PRMS DAILY MODE
70 DIM D1(500),D9(500),D10(500),D11(500),D2(500),D3(500)
80 DIM D4(500),D5(500),D6(500),D7(500),D8(500)
90 RANDOMIZE
100 PRINT "ENTER THE NAME OF INPUTFILE::" :INPUT IFN$
110 M$=IFN$
120 OPEN "I",#1,M$
130 PRINT "ENTER THE NAME OF OUTPUT FILE ::", :INPUT OFN$
140 N$=OFN$ : OPEN "O",#2,N$
150 PRINT "STANDARD DEVIATION FOR ERROR:", :INPUT S
160 J=0
170 REM READ INPUT DATA FROM FILE AND ADD ERROR
180 IF EOF(1) THEN 780
190 INPUT#1,D1,D2,D3,D4,D5,D6,D7,D8,D9,D10,D11
200 J=J+1
210 IF D4=0 THEN 250
220 M=D4 : SD= S*D4
230 GOSUB 850
240 D4= RNOR : M=0 : SD=0
250 IF D5=0 THEN 290
260 M=D5 : SD= S*D5
270 GOSUB 850
280 D5=RNOR : M=0 : SD=0
290 IF D6=0 THEN 330
300 M=D6 : SD= S*D6
310 GOSUB 850
320 D6=RNOR : M=0 : SD=0
330 IF D7=0 THEN 370
340 M=D7 : SD= S*D7
350 GOSUB 850
360 D7=RNOR :M=0 :SD=0
370 IF D8=0 THEN 410
380 M=D8 :SD= S*D8
390 GOSUB 850
400 D8=RNOR :M=0 : SD=0
410 IF D9=0 THEN 450
420 M=D9 : SD= S*D9
430 GOSUB 850
440 D9=RNOR :M=0 :SD=0
450 IF D10=0 THEN 490
460 M=D10 : SD= S*D10
470 GOSUB 850
480 D10=RNOR :M=0 :SD=0
490 IF D11=0 THEN 550
500 M=D11 : SD= S*D11
510 GOSUB 850
520 D11=RNOR :M=0 :SD=0
530 PRINT
```

```
540 REM  WRITE DATA TO OUTPUTFILE IN PRMS FORMAT
550 ON D3 GOTO 560,560,560,600
560 PRINT #2,USING "3CKSA.FLOW        ####";D1;
570 PRINT #2,USING "##";D2,D3;
580 PRINT #2,USING "####.##";D4,D5,D6,D7,D8,D9,D10,D11
590 GOTO 770
600 ON D2 GOTO 700,610,700,740,700,740,700,700, 740,700,740,700
610 IF D1=1976 THEN 660
620 PRINT #2,USING "3CKSA.FLOW        ####";D1;
630 PRINT #2,USING "##";D2,D3;
640 PRINT #2,USING "####.##";D4,D5,D6,D7
650 GOTO 770
660 PRINT #2,USING "3CKSA.FLOW        ####";D1;
670 PRINT #2,USING "##";D2,D3;
680 PRINT #2,USING "####.##";D4,D5,D6,D7,D8
690 GOTO 770
700 PRINT #2,USING "3CKSA.FLOW        ####";D1;
710 PRINT #2,USING "##";D2,D3;
720 PRINT #2,USING "####.##";D4,D5,D6,D7,D8,D9,D10
730 GOTO 770
740 PRINT #2,USING "3CKSA.FLOW        ####";D1;
750 PRINT #2,USING "##";D2,D3;
760 PRINT #2,USING "####.##";D4,D5,D6,D7,D8,D9
770 GOTO 180
780 CLOSE #1
790 CLOSE #2
800 END
810 REM ******************************************************************
820 REM
830 REM      SUBROUTINE : NORMAL DISTRIBUTION GENERATOR
840 REM
850 REM ******************************************************************
860 IF NRN=1 THEN 960
870 R1=2*RND-1
880 R2=2*RND-1
890 S=R1^2+R2^2
900 IF S>=1 THEN 870
910 RNN1=R1*SQR((-2*LOG(S))/S)
920 RNN2=R2*SQR((-2*LOG(S))/S)
930 RNOR=M+RNN1*SD
940 NRN=NRN+1
950 RETURN
960 RNOR=M+RNN2*SD
970 NRN=0
980 RETURN
```

```
10 REM ******************************************************************
20 REM          GENERATION OF LOGNORMAL RANDOM ERROR
30 REM ******************************************************************
40 REM PROGRAM READS FROM A FILE AND ADD RANDOM ERROR
50 REM INPUT INPUT AND OUTPUT FILE NAMES, SD OF ERROR, S
60 REM OUTPUT FILE FORMAT IS COMPATIBLE TO PRMS DAILY MODE
70 DIM D1(500),D9(500),D10(500),D11(500),D2(500),D3(500),
80 DIM D4(500),D5(500),D6(500),D7(500),D8(500)
90 PRINT "ENTER THE NAME OF INPUTFILE::" :INPUT IFN$
100 M$="b:"+IFN$
110 OPEN "I",#1,M$
120 PRINT "ENTER THE NAME OF OUTPUT FILE ::",:INPUT OFN$
130 N$="b:"+OFN$ : OPEN "O",#2,N$
140 PRINT "STANDARD DEVIATION OF ERROR:",: INPUT S
150 J=0
160 REM READ DATA AND ADD LOGNORMAL ERROR
170 IF EOF(1) THEN 780
180 INPUT#1,D1,D2,D3,D4,D5,D6,D7,D8,D9,D10,D11
190 J=J+1
200 RANDOMIZE
210 IF D4=0 THEN 250
220 M=.5*(LOG(D4^2/(S^2+1))) : SD=(LOG(S^2+1))^.5
230 GOSUB 840
240 D4= EXP(RNOR) : M=0 : SD=0
250 IF D5=0 THEN 290
260 M= .5*(LOG(D5^2/(S^2+1))) : SD=(LOG(S^2+1))^.5
270 GOSUB 840
280 D5=EXP(RNOR): M=0 : SD=0
290 IF D6=0 THEN 330
300 M=.5*(LOG(D6^2/(S^2+1))) : SD=(LOG(S^2+1))^.5
310 GOSUB 840
320 D6=EXP(RNOR) : M=0 : SD=0
330 IF D7=0 THEN 370
340 M=.5*(LOG(D7^2/(S^2+1))) : SD=(LOG(S^2+1))^.5
350 GOSUB 840
360 D7=EXP(RNOR) : M=0 :SD=0
370 IF D8=0 THEN 410
380 M=.5*(LOG(D8^2/(S^2+1))) :SD=(LOG(S^2+1))^.5
390 GOSUB 840
400 D8=EXP(RNOR) :M=0 : SD=0
410 IF D9=0 THEN 450
420 M=.5*(LOG(D9^2/(S^2+1))) : SD=(LOG(S^2+1))^.5
430 GOSUB 840
440 D9=EXP(RNOR) :M=0 :SD=0
450 IF D10=0 THEN 490
460 M=.5*(LOG(D10^2/(S^2+1))) : SD=(LOG(S^2+1))^.5
470 GOSUB 840
480 D10=EXP(RNOR) :M=0 :SD=0
490 IF D11=0 THEN 550
500 M=.5*(LOG(D11^2/(S^2+1))) : SD=(LOG(S^2+1))^.5
510 GOSUB 840
520 D11=EXP(RNOR) :M=0 :SD=0
530 PRINT
```

```
540 REM PRINT DATA TO OUTPUT FILE IN PRMS FORMAT
550 ON D3 GOTO 560,560,560,600
560 PRINT #2,USING "3CKSA.FLOW       ####";D1;
570 PRINT #2,USING "##";D2,D3;
580 PRINT #2,USING "####.##";D4,D5,D6,D7,D8,D9,D10,D11
590 GOTO 770
600 ON D2 GOTO 700,610,700,740,700,740,700,700,740, 700,740,700
610 IF D1=1976 THEN 660
620 PRINT #2,USING "3CKSA.FLOW       ####";D1;
630 PRINT #2,USING "##";D2,D3;
640 PRINT #2,USING "####.##";D4,D5,D6,D7
650 GOTO 770
660 PRINT #2,USING "3CKSA.FLOW       ####";D1;
670 PRINT #2,USING "##";D2,D3;
680 PRINT #2,USING "####.##";D4,D5,D6,D7,D8
690 GOTO 770
700 PRINT #2,USING "3CKSA.FLOW       ####";D1;
710 PRINT #2,USING "##";D2,D3;
720 PRINT #2,USING "####.##";D4,D5,D6,D7,D8,D9,D10
730 GOTO 770
740 PRINT #2,USING "3CKSA.FLOW       ####";D1;
750 PRINT #2,USING "##";D2,D3;
760 PRINT #2,USING "####.##";D4,D5,D6,D7,D8,D9
770 GOTO 170
780 CLOSE #1
790 CLOSE #2
800 END
810 REM ********************************************************************
820 REM     SUBROUTINE : NORMAL DISTRIBUTION GENERATOR
830 REM ********************************************************************
840 REM SUBROUTINE FOR NORMAL DISTRIBUTION
850 IF NRN=1 THEN 950
860 R1=2*RND-1
870 R2=2*RND-1
880 S=R1^2+R2^2
890 IF S>=1 THEN 860
900 RNN1=R1*SQR((-2*LOG(S))/S)
910 RNN2=R2*SQR((-2*LOG(S))/S)
920 RNOR=M+RNN1*SD
930 NRN=NRN+1
940 RETURN
950 RNOR=M+RNN2*SD
960 NRN=0
970 RETURN
```

```
10 REM **********************************************************************
20 REM   GENERATION OF DOUBLE EXPONENTIAL RANDOM ERROR
30 REM **********************************************************************
40 REM PROGRAM READS FROM A FILE AND ADD RANDOM ERROR
50 REM INPUT INPUT AND OUTPUT FILE NAMES, SD OF ERROR, S.
60 REM OUTPUT FILE FORMAT IS COMPATIBLE TO PRMS DAILY MODE
70 DIM D1(500),D9(500),D10(500),D11(500),D2(500),D3(500),
80 DIM D4(500),D5(500),D6(500),D7(500),D8(500)
90 RANDOMIZE
100 PRINT "ENTER THE NAME OF INPUTFILE::",:INPUT IFN$
110 M$=IFN$
120 OPEN "I",#1,M$
130 PRINT "ENTER THE NAME OF OUTPUT FILE ::",:INPUT OFN$
140 N$=OFN$ : OPEN "O",#2,N$
150 PRINT "STANDARD DEVIATION OF ERROR:",: INPUT S
160 J=0
170 REM READ DATA AND ADD DOUBLE EXPONENTIAL ERROR
180 IF EOF(1) THEN 860
190 INPUT#1,D1,D2,D3,D4,D5,D6,D7,D8,D9,D10,D11
200 J=J+1
210 IF D4=0 THEN 260
220 B= S*D4/SQR(2)
230 GOSUB 920
240 D4=DR+DE : B=0 : DE=0
250 IF D4 < 0 THEN D4=0
260 IF D5=0 THEN 310
270 B= S*D5/SQR(2)
280 GOSUB 920
290 D5=D5+DE : B=0 : DE=0
300 IF D5 < 0 THEN D5=0
310 IF D6=0 THEN 360
320 B= S*D6/SQR(2)
330 GOSUB 920
340 D6=D6+DE : B=0 : DE=0
350 IF D6 < 0 THEN D6=0
360 IF D7=0 THEN 410
370 B= S*D7/SQR(2)
380 GOSUB 920
390 D7=D7+DE : B=0 : DE=0
400 IF D7 < 0 THEN D7=0
410 IF D8=0 THEN 460
420 B= S*D8/SQR(2)
430 GOSUB 920
440 D8=D8+DE : B=0 : DE=0
450 IF D8 < 0 THEN D8=0
460 IF D9=0 THEN 510
470 B= S*D9/SQR(2)
480 GOSUB 920
490 D9=D9+DE : B=0 : DE=0
500 IF D9 < 0 THEN D9=0
510 IF D10=0 THEN 560
520 B= S*D10/SQR(2)
530 GOSUB 920
```

```
540 D10=D10+DE : B=0 : DE=0
550 IF D10 < 0 THEN D10=0
560 IF D11=0 THEN 630
570 B= S*D11/SQR(2)
580 GOSUB 920
590 D11=D11+DE : B=0 : DE=0
600 IF D11 < 0 THEN D11=0
610 PRINT
620 REM WRITE TO OUTPUT FILE IN PRMS FORMAT
630 ON D3 GOTO 640,640,640,680
640 PRINT #2,USING "3CKSA.FLOW        ####";D1;
650 PRINT #2,USING "##";D2,D3;
660 PRINT #2,USING "###.###";D4,D5,D6,D7,D8,D9,D10,D11
670 GOTO 850
680 ON D2 GOTO 780,690,780,820,780,820,780,780,820, 780,820,780
690 IF D1=1976 THEN 740
700 PRINT #2,USING "3CKSA.FLOW        ####";D1;
710 PRINT #2,USING "##";D2,D3;
720 PRINT #2,USING "###.###";D4,D5,D6,D7
730 GOTO 850
740 PRINT #2,USING "3CKSA.FLOW        ####";D1;
750 PRINT #2,USING "##";D2,D3;
760 PRINT #2,USING "###.###";D4,D5,D6,D7,D8
770 GOTO 850
780 PRINT #2,USING "3CKSA.FLOW        ####";D1;
790 PRINT #2,USING "##";D2,D3;
800 PRINT #2,USING "###.###";D4,D5,D6,D7,D8,D9,D10
810 GOTO 850
820 PRINT #2,USING "3CKSA.FLOW ####";D1;
830 PRINT #2,USING "##";D2,D3;
840 PRINT #2,USING "###.###";D4,D5,D6,D7,D8,D9
850 GOTO 180
860 CLOSE #1
870 CLOSE #2
880 END
890 REM ******************************************************************
900 REM                    SUBROUTINE:
910 REM      DOUBLE EXPONENTIAL DUSTRIBUTION GENERATOR
920 REM ******************************************************************
930 M=RND
940 DE=-B*(LOG(2*M))
950 RETURN
```

```
10 REM *********************************************************************
20 REM        GENERATION OF UNIFORM RANDOM ERROR
30 REM *********************************************************************
40 REM PROGRAM READS FROM A FILE AND ADD RANDOM ERROR
50 REM INPUT INPUT AND OUTPUT FILE NAMES, SD OF ERROR, S.
60 REM OUTPUT FILE FORMAT IS COMPATIBLE TO PRMS DAILY MODE
70 DIM D1(500),D9(500),D10(500),D11(500),D2(500),D3(500),
80 DIM D4(500),D5(500),D6(500),D7(500),D8(500)
90 RANDOMIZE
100 PRINT "ENTER THE NAME OF INPUTFILE::" :INPUT IFN$
110 M$=IFN$
120 OPEN "I",#1,M$
130 PRINT "ENTER THE NAME OF OUTPUT FILE ::", :INPUT OFN$
140 N$=OFN$ : OPEN "O",#2,N$
150 PRINT "STANDARD DEVIATION OF ERROR:",: INPUT S
160 J=0
170 REM READ DATA AND ADD RANDOM ERROR
180 IF EOF(1) THEN 860
190 INPUT#1,D1,D2,D3,D4,D5,D6,D7,D8,D9,D10,D11
200 J=J+1
210 IF D4=0 THEN 260
220 B= S*D4*SQR(3) : A=-B
230 GOSUB 890
240 D4=DR+DE : B=0 : A=0
250 IF D4 < 0 THEN D4=0
260 IF D5=0 THEN 310
270 B= S*D5*SQR(3) : A=-B
280 GOSUB 890
290 D5=D5+DE : B=0 : A=0
300 IF D5 < 0 THEN D5=0
310 IF D6=0 THEN 360
320 B= S*D6*SQR(3) : A=-B
330 GOSUB 890
340 D6=D6+DE : B=0 : A=0
350 IF D6 < 0 THEN D6=0
360 IF D7=0 THEN 410
370 B= S*D7*SQR(3) : A=-B
380 GOSUB 890
390 D7=D7+DE : B=0 : A=0
400 IF D7 < 0 THEN D7=0
410 IF D8=0 THEN 460
420 B= S*D8*SQR(3) : A=-B
430 GOSUB 890
440 D8=D8+DE : B=0 : A=0
450 IF D8 < 0 THEN D8=0
460 IF D9=0 THEN 510
470 B= S*D9*SQR(3) : A=-B
480 GOSUB 890
490 D9=D9+DE : B=0 : A=0
500 IF D9 < 0 THEN D9=0
510 IF D10=0 THEN 560
520 B= S*D10*SQR(3) : A=-B
530 GOSUB 890
```

```
540 D10=D10+DE : B=0 : A=0
550 IF D10 < 0 THEN D10=0
560 IF D11=0 THEN 630
570 B= S*D11*SQR(3) : A=-B
580 GOSUB 890
590 D11=D11+DE : B=0 : A=0
600 IF D11 < 0 THEN D11=0
610 PRINT
620 REM WRITE TO OUTPUT FILE IN PRMS FORMAT
630 ON D3 GOTO 640,640,640,680
640 PRINT #2,USING "3CKSA.FLOW        ####";D1;
650 PRINT #2,USING "##";D2,D3;
660 PRINT #2,USING "###.###";D4,D5,D6,D7,D8,D9,D10,D11
670 GOTO 850
680 ON D2 GOTO 780,690,780,820,780,820,780,780,820, 780,820,780
690 IF D1=1976 THEN 740
700 PRINT #2,USING "3CKSA.FLOW        ####";D1;
710 PRINT #2,USING "##";D2,D3;
720 PRINT #2,USING "###.###";D4,D5,D6,D7
730 GOTO 850
740 PRINT #2,USING "3CKSA.FLOW        ####";D1;
750 PRINT #2,USING "##";D2,D3;
760 PRINT #2,USING "###.###";D4,D5,D6,D7,D8
770 GOTO 850
780 PRINT #2,USING "3CKSA.FLOW        ####";D1;
790 PRINT #2,USING "##";D2,D3;
800 PRINT #2,USING "###.###";D4,D5,D6,D7,D8,D9,D10
810 GOTO 850
820 PRINT #2,USING "3CKSA.FLOW        ####";D1;
830 PRINT #2,USING "##";D2,D3;
840 PRINT #2,USING "###.###";D4,D5,D6,D7,D8,D9
850 GOTO 180
860 CLOSE #1
870 CLOSE #2
880 END
890 REM ********************************************************************
900 REM      SUBROUTINE: UNIFORM DISTRIBUTION GENERATOR
910 REM ********************************************************************
920 M=RND
930 DE =A+(B-A)*M
940 RETURN
```

```
10 REM *******************************************************************
20 REM        GENERATION OF CORRELATED NORMAL ERROR
30 REM *******************************************************************
40 REM PROGRAM READS FROM A FILE AND ADD RANDOM ERROR
50 REM INPUT FILE NAMES,SD OF ERROR, S,AND CORRELATION, R.
60 REM OUTPUT FILE FORMAT IS COMPATIBLE TO PRMS DAILY MODE
70 DIM D1(500),D9(500),D10(500),D11(500),D2(500),D3(500),
80 DIM D4(500),D5(500),D6(500),D7(500),D8(500),E(500)
90 RANDOMIZE
100 E11=.001
110 PRINT "ENTER THE NAME OF INPUTFILE::",: INPUT IFN$
120 M$=IFN$
130 OPEN "I",#1,M$
140 PRINT "ENTER THE NAME OF OUTPUT FILE ::", :INPUT OFN$
150 N$=OFN$ : OPEN "O",#2,N$
160 PRINT "STANDARD DEVIATION OF ERROR:", : INPUT S
170 PRINT "CORRELATION OF ERROR :", : INPUT R
180 J=0
190 IF EOF(1) THEN 880
200 REM READ DATA AND ADD CORRELATED RANDOM ERROR
210 INPUT#1,D1,D2,D3,D4,D5,D6,D7,D8,D9,D10,D11
220 J=J+1
230 IF D4=0 THEN 270
240 M=D4 : SD= S*D4
250 GOSUB 910
260 E4=  R*E11+RNOR : D4=D4+E4 : M=0 : SD=0 : GOTO 280
270 E4=0
280 IF D5=0 THEN 320
290 M=D5 : SD= S*D5
300 GOSUB 910
310 E5= R*E4+RNOR : D5=D5+E5 : M=0 : SD=0 : GOTO 330
320 E5=0
330 IF D6=0 THEN 370
340 M=D6 : SD= S*D6
350 GOSUB 910
360 E6= R*E5+RNOR : D6=D6+E6 : M=0 : SD=0 : GOTO 380
370 E6=0
380 IF D7=0 THEN 420
390 M=D7 : SD= S*D7
400 GOSUB 910
410 E7= R*E6+RNOR : D7=D7+E7 : M=0 : SD=0 : GOTO 430
420 E7=0
430 IF D8=0 THEN 470
440 M=D8 :SD= S*D8
450 GOSUB 910
460 E8= R*E7+RNOR : D8=D8+E8 : M=0 : SD=0 : GOTO 480
470 E8=0
480 IF D9=0 THEN 520
490 M=D9 : SD= S*D9
500 GOSUB 910
510 E9= R*E8+RNOR : D9=D9+E9 : M=0 : SD=0 : GOTO 530
520 E9=0
530 IF D10=0 THEN 570
```

```
540 M=D10 : SD= S*D10
550 GOSUB 910
560 E10= R*E9+RNOR :D10=D10+E10 :SD=0 : M=0 :GOTO 580
570 E10=0
580 IF D11=0 THEN 630
590 M=D11 : SD= S*D11
600 GOSUB 910
610 E11= R*E10+RNOR :D11=D11+E11 : M=0 :SD=0 :GOTO 650
620 PRINT
630 E11=0
640 REM WRITE TO OUTPUT FILE IN PRMS FORMAT
650 ON D3 GOTO 660,660,660,700
660 PRINT #2,USING "3CKSA.FLOW        ####";D1;
670 PRINT #2,USING "##";D2,D3;
680 PRINT #2,USING "####.##";D4,D5,D6,D7,D8,D9,D10,D11
690 GOTO 870
700 ON D2 GOTO 800,710,800,840,800,840,800,800,840,800, 840, 800
710 IF D1=1976 THEN 760
720 PRINT #2,USING "3CKSA.FLOW        ####";D1;
730 PRINT #2,USING "##";D2,D3;
740 PRINT #2,USING "####.##";D4,D5,D6,D7
750 GOTO 870
760 PRINT #2,USING "3CKSA.FLOW        ####";D1;
770 PRINT #2,USING "##";D2,D3;
780 PRINT #2,USING "####.##";D4,D5,D6,D7,D8
790 GOTO 870
800 PRINT #2,USING "3CKSA.FLOW        ####";D1;
810 PRINT #2,USING "##";D2,D3;
820 PRINT #2,USING "####.##";D4,D5,D6,D7,D8,D9,D10
830 GOTO 870
840 PRINT #2,USING "3CKSA.FLOW        ####";D1;
850 PRINT #2,USING "##";D2,D3;
860 PRINT #2,USING "####.##";D4,D5,D6,D7,D8,D9
870 GOTO 190
880 CLOSE #1
890 CLOSE #2
900 END
910 REM *******************************************************************
920 REM      SUBROUTINE : NORMAL DISTRIBUTION GENERATOR
930 REM *******************************************************************
940 IF NRN=1 THEN 1040
950 R1=2*RND-1
960 R2=2*RND-1
970 S=R1^2+R2^2
980 IF S>=1 THEN 950
990 RNN1=R1*SQR((-2*LOG(S))/S)
1000 RNN2=R2*SQR((-2*LOG(S))/S)
1010 RNOR=M+RNN1*SD
1020 NRN=NRN+1
1030 RETURN
1040 RNOR=M+RNN2*SD
1050 NRN=0
1060 RETURN
```

APPENDIX B

GENERATED PARAMETER FILES FROM MONTE CARLO ANALYSIS

Generated parameter files are stored in a 5.25"x5.25" floppy disc. Each file contains 100 set of independent parameter values. There are altogether 105 files for different error combinations and estimation techniques. File names are identified by the code name for error distribution, percentage of error standard deviation, source of error and estimation technique respectively. Each file has the extension "DAT".

Error distribution code:

N - Normal.
L - Lognormal
DE - Double Exponential
UF - Uniform
AR - Autoregressive

Error standard deviation code:

10 - 10% of error free value
20 - 20% of error free value
30 - 30% of error free value

Source of error code:

P - Precipitation
Q - Streamflow

Estimation technique code:

L - Least squares.
A - Absolute errors
M - Maximum likelihood

For example a file "DE10QA.DAT" indicates that double exponential (DE) error of 10% (10) in streamflow (Q) was used to estimate the parameters by absolute error estimation technique (A).

APPENDIX C

COMPUTER PROGRAM FOR GENERATION OF BIVARIATE LOGNORMAL DATA

```
10 REM ****************************************************************
20 REM GENERATION OF BIVARIATE CORRELATED LOGNORMALLY
30 REM           DISTRIBUTED RANDOM VARIABLES.
40 REM ****************************************************************
50 REM N IS NO OF OBSERVATIONS TO BE GENERATED.
60 REM MEAN1 AND MEAN2,MEANS OF VARIABLE 1 AND 2.
70 REM SD1,SD2  ARE SD OF VARIABLES 1 AND 2.
80 REM R IS CORRELATION BETWEEN THE VARIABLES.
90 DIM XRND(200), X(2,200), Z(2,200),V1(200),V2(200)
100 PRINT "INPUT MEAN1 AND MEAN2"
110 INPUT MEAN1,MEAN2
120 PRINT "SD1 AND SD2 FOR ORIGINAL DATA"
130 INPUT SD1,SD2
140 PRINT "INPUT THE CORRELATION COEFFICIENT"
150 INPUT R
160 SD1=LOG((SD1/MEAN1)^2+1)
170 SD1=SQR(SD1)
180 SD2=LOG((SD2/MEAN2)^2+1)
190 SD2=SQR(SD2)
200 MEAN1=LOG(MEAN1)-SD1^2/2
210 MEAN2=LOG(MEAN2)-SD2^2/2
220 A1=EXP(SD1^2)-1
230 A2=EXP(SD2^2)-1
240 A3=SD1*SD2
250 R=LOG(1+R*SQR(A1*A2))/A3
260 PRINT "THE NUMBER OF OBSERVATIONS"
270 INPUT N
280 REM EIGENVALUES
290 L1=1+R
300 L2=1-R
310 REM THE A MATRIX
320 A(1,1)=1/SQR(2)
330 A(2,1)=A(1,1)
340 A(1,2)=A(1,1)
350 A(2,2)=-A(1,1)
360 REM GENERATION OF Z VALUES
370 MEAN=0
380 SD=SQR(L1)
390 GOSUB 640
400 FOR I=1 TO N
410 Z(1,I)=XRND(I)
420 NEXT I
430 SD=SQR(L2)
440 GOSUB 630
450 FOR I=1 TO N
460 Z(2,I)=XRND(I)
470 NEXT I
480 REM TRANSFORMATION TO X VALUES
490 FOR I=1 TO N
500 X(1,I)=Z(1,I)*A(1,1)+Z(2,I)*A(1,2)
510 V1(I)=EXP(X(1,I)*SD1+MEAN1)
520 X(2,I)=Z(1,I)*A(1,2)+Z(2,I)*A(2,2)
530 V2(I)=EXP(X(2,I)*SD2+MEAN2)
```

```
540 NEXT I
550 PRINT "ENTER DISK FILE NAME TO STORE DATA"
560 INPUT F$
570 OPEN "O",#1,F$
580 FOR I=1 TO N
590 PRINT #1, I,V1(I),V2(I)
600 NEXT I
610 CLOSE #1
620 END
630 REM ***SUBROUTINE: NORMAL DISTRIBUTION GENERATOR***
640 RANDOMIZE TIMER
650 FOR I=1 TO N
660 IF NRN=1 THEN 770
670 R1=2*RND-1
680 R2=2*RND-1
690 S=R1^2+R2^2
700 IF S>=1 THEN 670
710 RNN1=R1*SQR((-2*LOG(S))/S)
720 RNN2=R2*SQR((-2*LOG(S))/S)
730 XRND(I)=MEAN+RNN1*SD
740 NRN=NRN+1
750 IF I>=N THEN 810
760 GOTO 800
770 XRND(I)=MEAN+RNN2*SD
780 NRN=0
790 IF I>=N THEN 810
800 NEXT I
810 RETURN
```

Generated Parameters and Mean Annual Runoff

SMAX (in)	REMX (in)	SC1	Runoff (cfs days)
6.42	1.40	0.44	1.95
5.60	1.40	0.49	2.14
8.88	1.56	0.37	1.78
6.53	1.89	0.41	1.56
4.62	2.13	0.47	1.34
6.25	2.96	0.47	2.00
7.53	1.49	0.51	3.28
7.81	2.01	0.44	2.28
8.30	2.01	0.53	3.61
7.36	1.82	0.55	3.59
5.58	2.06	0.47	3.47
7.32	1.99	0.49	2.74
7.15	1.83	0.32	0.87
6.42	1.98	0.48	2.28
8.05	1.56	0.35	1.33
5.30	2.74	0.48	1.75
5.50	1.62	0.48	1.88
5.35	1.71	0.53	2.32
7.06	2.14	0.48	2.50
5.88	1.54	0.41	1.41
7.53	1.82	0.53	3.44
9.22	1.71	0.33	1.35
5.03	1.60	0.51	1.97
9.70	2.13	0.23	0.56
8.09	2.21	0.43	2.18
7.86	2.08	0.36	1.36
6.16	2.08	0.49	2.26
8.87	3.02	0.38	1.68
9.01	2.46	0.33	1.22
7.52	1.55	0.61	4.19
6.63	2.70	0.48	2.29
7.80	1.64	0.43	2.21
8.72	2.35	0.33	1.18
8.30	1.69	0.42	2.25
6.41	1.59	0.37	1.21
6.86	1.58	0.47	2.38
7.81	2.42	0.43	2.09
7.48	1.73	0.39	1.64
8.01	1.89	0.42	2.11
6.59	2.23	0.41	1.57
6.20	1.69	0.49	2.34
8.20	1.92	0.35	1.32
5.84	1.83	0.44	1.60
4.05	1.61	0.46	1.03
7.55	2.00	0.46	2.44
4.58	1.66	0.47	1.37
4.88	1.68	0.68	3.38
5.01	1.98	0.65	3.22

Generated Parameters and Mean Annual Runoff

SMAX (in)	REMX (in)	SC1	Runoff (cfs days)
5.00	1.96	0.63	3.02
8.36	1.41	0.44	2.63
5.82	1.72	0.49	2.12
7.95	2.02	0.46	2.56
7.25	2.05	0.47	2.45
6.77	2.63	0.37	1.20
5.83	2.01	0.49	2.07
3.94	2.05	0.47	1.00
5.64	2.13	0.38	0.99
6.69	2.54	0.42	1.65
5.22	1.96	0.71	3.82
7.13	2.00	0.48	2.55
8.76	1.69	0.45	2.84
4.98	2.55	0.48	1.59
6.13	1.16	0.58	3.58
6.72	1.83	0.43	1.82
7.49	1.43	0.55	3.77
4.43	2.56	0.53	1.80
8.96	2.24	0.42	2.33
6.03	2.40	0.40	1.30
6.63	1.94	0.44	1.88
6.73	1.88	0.45	2.02
9.07	1.70	0.46	3.08
9.47	1.36	0.45	3.30
7.19	1.91	0.52	3.16
6.09	2.11	0.41	1.40
7.23	1.59	0.41	1.77
6.56	2.07	0.43	1.77
5.41	2.27	0.49	1.87
6.59	1.38	0.44	2.01
5.89	2.14	0.52	2.45
6.88	1.98	0.44	1.97
7.36	2.02	0.40	1.68
7.69	2.33	0.45	2.29
6.90	2.32	0.40	1.53
6.09	2.28	0.58	3.28
6.33	2.90	0.37	1.14
7.83	1.46	0.41	2.02
6.21	2.33	0.39	1.28
6.78	2.46	0.43	1.80
7.72	3.10	0.44	2.05
5.74	1.81	0.43	1.50
7.99	1.77	0.44	2.40
7.41	2.35	0.46	2.33
5.37	2.51	0.45	1.48
7.43	2.50	0.50	2.78
6.61	1.89	0.40	1.48

Generated Parameters and Mean Annual Runoff

SMAX (in)	REMX (in)	SC1	Runoff (cfs days)
5.85	1.88	0.36	0.92
6.59	2.40	0.52	2.80
6.42	1.90	0.49	2.39
5.72	1.52	0.51	2.34
6.62	2.21	0.52	2.80

VITA 2

Apurba K. Borah

Candidate for the Degree of

Doctor of Philosophy

Thesis: ERROR DISTRIBUTIONS AND PARAMETER ESTIMATION IN RAINFALL-RUNOFF MODELING

Major Field: Agricultural Engineering

Biographical:

Personal Data: Born in Biswanath Charali, Assam, India, January 29, 1960, son of Mr. and Mrs. Premananda Borah.

Education: Graduated from Charali Higher Secondary and Multipurpose School, Sonitpur, Assam, India in 1975; passed Pre-University (Science) Examination from the Biswanath College, Gauhati University, India in 1977; received Bachelor of Engineering (Hons.) Degree in Agricultural Engineering from the University of Udaipur, Rajasthan, India in 1982; received Master of Engineering Degree in Agricultural Engineering from the Asian Institute of Technology, Bangkok, Thailand in 1985; attended the Graduate School, University of Manitoba, Winnepeg, Canada from September 1985 to December 1986; completed the requirements for the Doctor of Philosophy Degree at Oklahoma State University, Stillwater in May, 1990.

Professional Experience: Assistant Engineer, Assam State Minor Irrigation Development Corporation, Assam, India, December, 1982, to December, 1983; Graduate Research Assistant, Agricultural Engineering Department, University of Manitoba, 1985 to 1986; Graduate Research Assistant, Agricultural Engineering Department, Oklahoma State University, 1987 to 1989.

Professional Organizations: Member of the Alpha Epsilon, the Honor Society of Agricultural Engineers ; American Society of Agricultural Engineers.