COMPLETION TECHNICAL REPORT
06

COMPARATIVE EVALUATION OF ESTIMATORS OF SOME FLOOD FREQUENCY MODELS USING MONTE CARLO SIMULATION

by

VIJAY P. SINGH
KISHORE ARORA

for

U. S. GEOLOGICAL SURVEY
Department of Interior
Reston, VA 22092

LOUISIANA WATER RESOURCES RESEARCH INSTITUTE
Louisiana State University
Baton Rouge, LA 70803

DISCLAIMER

This completion report is published with funds provided in part by the U. S. Geological Survey, Department of Interior, as authorized by the Water Resources Act of 1984.

A COMPARATIVE EVALUATION OF THE ESTIMATORS OF SOME FLOOD FREQUENCY MODELS USING MONTE CARLO SIMULATION

Kishore Arora and Vijay P. Singh
Department of Civil Engineering
Louisiana State University
Baton Rouge, LA 70803-6405

February 1987

ACKNOWLEDGEMENTS

This study was supported in part by the U.S. Department of Interior, Geological Survey, through the Louisiana Water Resources Research Institute, under the project, "A Multivariate Stochastic Analysis of Flood Magnitude, Duration and Volume," The work reported here also constituted an M.S. thesis of the senior author. Helpful suggestions of Drs. J. F. Cruise, H. Schneider and M. Fiorentino are gratefully acknowledged.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS

AM	Annual maximum series
AFS	Annual flood series
BIAS(.)	Bias of the estimator
CV	Coefficient of variation
CMOM	Bias-corrected moment estimator
EFF(.)	Relative efficiency of the estimator
E(.)	Expectation of the variate
EV1	Extreme value type 1 (Gumbel)
ENT	Entropy
GEV	Generalized extreme value
ICM	Incomplete means
LEAS	Least squares
L	Likelihood function
LL	Log-likelihood function
LN2	Two parameter log normal
LN3	Three parameter log normal
LP3	Log Pearson type 3
MIX	Method of mixed moments
MMD	Direct (Untransformed) method of moment
MMI	Indirect (log-transformed) method of moment
MOM	Method of moment
MLE	Maximum likelihood estimation
MSE	Mean square error
POT	Peak over threshold (partial duration)
PWM	Probability weight moments
RMSE	Root mean square error

SE(.)	Standard error of estimator (= standard deviation)
STD(.)	Standard deviation of the variate
SKEW	Skewness coefficient
TCEV	Two component extreme value
VAR(.)	Variance of the variate
WAK	Five parameter Wakeby
μ	Mean
β	Coefficient of variation
γ	Skewness coefficient of real data
γ_y	Skewness coefficient of log-transformed data
λ	Kurtosis coefficient
μ_r'	r-th moment of variate about the origin
Γ	Gamma function
ψ	psi or digamma function
g	Sample skewness coefficient estimate
$\hat{x}$	Estimator of x
s_x	Sample standard deviation estimate
x_T	T-year return period quantile
σ_x	Standard derivation of variate
$\hat{\theta}$	Estimator of θ

ABSTRACT

Performance of estimators of several commonly used flood frequency models is evaluated in terms of the statistical criteria of bias (BIAS), standard error (SE), and mean square error (MSE). The procedure is based on Monte Carlo simulation. The models considered are the Gumbel's extreme value type 1 (EV1), log Pearson type 3 (LP3), and the two component extreme value (TCEV) distributions. The estimators used are based on the methods of moments (MOM), maximum likelihood estimation (MLE), probability weighted moments (PWM), least squares (LEAS), incomplete means (ICM), mixed moments (MIX), and entropy (ENT).

The performance of the EV1 estimators is evaluated for random as well as serially correlated samples. MLE provides efficient quantile estimates even for small samples, closely followed by ENT. The PWM quantile estimates are unbiased for random samples and least biased for serially correlated samples. The performance of the estimators worsens for the serially correlated case, though they perform much more closely in this case. A new correction is derived, based on simulation results, to correct the bias in the MOM quantile estimates.

The LP3 estimators are evaluated over the range of coefficients of variation and skewness characteristic of the real flood occurrence. The performance of the estimator recommended by the U.S. Water Resources Council, and other competing estimators is significantly inferior to that of MIX and another estimator based on the moments of the variate in real space. A new methodology is developed to solve for the MIX parameter estimates directly, without having to resort to iterative procedure.

The TCEV distribution has been shown by Rossi, et al. (1984) to possess many of the properties characteristic of the real world floods. The ENT estimator is derived for this model. It performs similar to MLE estimator and has the advantage of being relatively easier to solve.

Chapter 1

INTRODUCTION

The approaches to estimate the design flood fall under two categories--deterministic and probabilistic. The deterministic approach essentially utilizes the rainfall-runoff relation to estimate the flood corresponding to a specified design storm. While the probabilistic approach utilizes the flood probability model in place of the rainfall-runoff relation, and associates probability levels with the estimated flood events. The flood phenomena is the hydrologic response at the watershed scale. The probabilistic approach recognizes that the hydrologic systems at the watershed scale are poorly identified in terms of the generally well understood local hydrologic processes, and suffer from scanty data bases. Thus, as a logical outcome, this approach employs stochastic lumped models whose parameters are no longer physically measurable and so must be estimated statistically from the sample of the lumped response--the flood record.

Flood frequency analysis encompasses the techniques utilized to estimate the magnitude of extreme flood events corresponding to specified probability levels. It is the probabilistic approach to flood estimation. The knowledge of magnitude of extreme flood events is important in such areas as flood plain management and design of hydraulic structures. The increased water resources development and many programs of national scope involving large outlays of capital, such as bridge and drainage design for national highway systems, flood protection works, and flood insurance programs call for hydrologic input of frequency of occurrence of extreme events (Benson, 1968). Thus, the flood frequency analysis forms the basis for the engineering design of many projects and the economic analysis of flood-related programs.

The magnitude of a flood event is commonly referred to as T-year flood, where T usually called the return period or the recurrence interval, is a measure of the probability level of the event. The T-year flood event is that which can be expected to be exceeded once on an average of every T years. The estimation of the T-year flood typically involves inferences based on n years of flood records. The occurrence of extreme flood events is believed to be a stochastic process, since they tend to occur in an apparently random manner. Thus, one of the fundamental hypotheses of flood frequency analysis is that extreme flood events are random variables. This can be contrasted with the so-called PMF or SPF approach which treats the extreme events as deterministic in nature, and specifies upper bounds on the occurrence of such events. Yevjevich (1972) has characterized the difference between the use of the PMF and the probabilistic flood frequency approach as the difference between expediency and truth.

Due to the probabilistic assumption of flood frequency analysis, it is important that the sequence of flood records used for inferences be representative of a random sample. The popular approach for abstracting a random sample is to compose the sample of instantaneous annual maximum discharges recorded during successive water years. This so-called annual maximum (AM) series, while seemingly consisting of occurrences of independent and identically distributed random variables (random sample), utilizes a very small fraction of the flood records available. Another approach is to construct the sample of all flood peaks exceeding a certain threshold discharge. The resulting series is called the partial duration series, or peak over threshold (POT) series. Various models have been developed for the analysis of both AM and POT series (Cunnane, 1986). However, the POT analysis has not achieved the popularity of the AM models, primarily because the random sample criterion is hard to justify for partial duration series, and the POT models are more complicated than the AM models. This study concerns the AM series models only. When the AM series is being analyzed, the probability of the T-year flood being exceeded in any single water year is 1/T. In the case of POT series, this probability is, in general, a function of many other factors besides T.

In many instances, the problem of flood frequency analysis boils down to the estimation of the T-year return period flood event Q_T, given the annual maximum series of flood discharges derived from n years of record, where n usually is much less than T. In these cases, a probability model is used to fit the random sample, which is then extrapolated to the probability level, or equivalently, return period of interest.

The true probability distribution of floods is obviously unknown and the aim of fitting a probability model to the AM series is to obtain a good estimate $\hat{Q}_T$ of the T-year flood Q_T (unknown). The question of how to obtain the best estimate of Q_T has received considerable attention. $\hat{Q}_T$ is subject to variability on account of both model error and sampling error. If T is greater than n, the error in the T-year flood estimate $\hat{Q}_T$ can be very large and the associated design losses quite considerable. Since the decisions based on the use of the T-year flood estimate have economic implications, the choice of this estimate should ideally be based on the economic loss function which expresses the loss incurred in terms of increased cost due to the mis-specification of a design flood by a particular estimate (Slack, et al., 1975). However, the form of the loss function is application-specific, and reasonable loss function types are yet to be found (Cunnane, 1986). It is, therefore, necessary to use surrogate loss functions to evaluate the performance of competing estimators for identifying the best estimator (Stedinger, 1980). One such surrogate loss function is the symmetric quadratic (Slack, et al., 1975; Landwehr, et al., 1980) which is proportional to the mean square error (MSE) of the quantity being estimated. Another surrogate loss function is the linear form which is proportional to the bias (BIAS). In all recent studies evaluating the performance of various competing flood quantile estimators, MSE (also equalling the sum of the variance and the square of the bias) and its components standard error and BIAS have been used as the criteria for selecting the most efficient (or best) estimator. The most efficient estimate is that which minimizes the MSE (or equivalently, the root MSE, equalling the square root of the MSE).

While the true or the so-called parent distribution of the real-world annual floods is unknown, many of its characteristics derived from a great many real world AM series have been investigated with particular emphasis on the behavior of the distribution's right-hand tail, since this portion of the distribution more directly affects the bias in the extreme flood quantile estimator. As the sample skewness is a statistic particularly sensitive to the behavior of the right tail of the distribution, the analysis of the skewness of the observed AM series should be especially useful. Matalas, et al. (1975) found that within a given geographical region, the skewness of the observed AM series exhibited a very high variability (standard deviation) about its regional mean. They also found that this high degree of variability could not be accounted for by the commonly used two- or three-parameter distributions. This phenomenon has since come to be known as the condition of separation. This condition has also been noted in Italian AM data by Rossi (1984) and in British data by Beran, et al. (1986). It was shown by Wallis, et al. (1977) that systematic changes in skewness over time or the mixing of different values of skewness within the region could cause separation. However, it is yet to be proved whether separation indeed occurs due to these reasons, or it is an inherent property of the parent distribution. It appears from the current literature that, at present, the condition of separation is thought to be an inherent property of the parent distribution by many hydrologists. Landwehr, et al. (1978) observed that high kurtosis is a necessary but not sufficient condition for separation. Shen, et al.(1980) and Ochoa, et al. (1980) drew a distinction between the so-called light-tailed distributions and the Paretian-tailed (or heavy-tailed) distributions. Ochoa, et al. (1980), based on the analysis of 407 AM series in the U.S., observed that the frequency with which the Paretian-tailed distribution provided a better fit to the data suggested that such distributions occurred very frequently in nature.

It would, therefore, appear from the evidence thus far that the true distribution of floods is heavily tailed with skewness greater than 1.14 (Gumbel skew), is highly kurtotic, and has coefficient of variation in the range of 0.3 to 0.8 (for U.S. floods; Landwehr, et al., 1978). The choice of a "good" parent distribution that mimics these characteristics is an important step in flood frequency analysis. In fact, the performance of various competing estimators can be realistically evaluated and compared by performing sampling experiments from the parent distribution.

Flood quantile estimators can be based on several distributions and also on various estimation methods for each distribution. Obviously, we can also have estimators based on the parent distribution that is being used for sampling experiments. Does this mean that an estimator based on a "good" parent distribution is the best estimator? Interestingly enough, the answer is no. In fact, the parent distribution, if directly used for flood quantile estimation yields estimates that are accurate (small bias) but rather imprecise (large variability) due to the small sample size. It has been argued, therefore, (Landwehr, 1980; Kuczera, 1982) that preference should be given to more robust flood quantile estimators derived from a distribution more parsimonious (fewer number of parameters) in the number of parameters, as long as they are more efficient (smaller MSE) and resistant over a wide parameter range. Then the loss incurred from the large bias resulting from the more parsimonious model would be more than offset by the reduced estimator variance. Nevertheless, it is worth noting that when the size of the sample used in the estimation is large (as in regionalization techniques), the variance of the estimator will no longer be the main contribution to its MSE. Then the most efficient estimator may be drived from the parent distribution.

Two empirical distributions similar to one another have been suggested as the flood parent able to account for the condition of separation without having to invoke spatial mixing of the distribution parameters. These include the 5-parameter Wakeby distribution suggested by H. A. Thomas, Jr., and discussed by Houghton (1977, 1978), and Landwehr, et al. (1978), and the Lambda distribution, introduced by Joiner and Rosenblatt (1971) and discussed by Stedinger (1977). Recently, a theoretical-type model based on the compound Poisson process is shown to be a viable parent distribution accounting for the condition of separation. This is the 4-parameter two component extreme value (TCEV) distribution suggested by Rossi, et al. (1984).

The techniques of flood frequency analysis utilizing AM series can be classified in one of the following analyses: at-site only, at-site/regional, and regional only (Cunnane, 1986).

In at-site analysis a probability model is selected and parameters of the model are estimated by using a suitable estimation technique. Subsequently, the T-year flood quantile can be estimated from the fitted model. Many probabilistic models have been used in flood frequency analysis; notable among these are the log Pearson 3 distribution, recommended by the U.S. Water Resources Council (1967, 1975, 1977, 1981), the three parameter log normal distribution (Burges, et al., 1975), and Gumbel's extreme value type 1 (EV1) distribution (Gumbel, 1958). The British Flood Studies Report (1975) recommended the usage of the general extreme value (distribution (3 parameters). EV1 distribution is a special case of the general extreme value distribution. The at-site analysis is simple in application and can yield reasonable estimates. However, the at-site analysis is almost always based on a rather short flood record (small n). Thus, the resulting flood quantile estimates can be highly variable. This has prompted exploration of regional estimation methods which pool together information from a number of sites in similar hydro-meteorological regions.

The at-site/regional analysis includes three important procedures: index flood method, empirical Bayes, and the two component extreme value method. All types of at-site/regional analysis methods implicitly or explicitly make assumptions about the regional distribution of annual floods. In the index flood method, the annual flood series are normalized by the at-site means. Then it is assumed that the distribution of the standardized annual flood is identical at all sites in the region. The two component extreme value distribution procedure is a modification of the index flood method, in which the standardized annual flood series that are assumed to follow a two component extreme value distribution are normalized using the at-site estimates of the two parameters of the extreme value distribution. An alternative regionalization approach which does not rely on normalized flood series is the empirical Bayes method. This method assumes that the parameters governing the distribution of annual floods at a site in a region come from a specified superpopulation, which has unknown parameters. The parameters are inferred from the flood data themselves, or from relationships between flood peak characteristics and physiographic and climatic factors. Kuczera (1982a, 1982b) suggests several estimators of flood quantiles based on this approach.

The "regional only" approach involves finding a regression relationship between the at-site flood means and the catchment characteristics (such as area, slope) in the region. The regressed relation is used to estimate the at-site mean at the ungauged site which is then used to estimate the T-year flood quantile from the regional flood frequency model.

The objective of this work is to evaluate and compare the performance of estimation methods of some commonly used probability models. The models considered are the Gumbel's (EV1), log Pearson 3 (LP3), and the two component extreme value (TCEV) distributions. Use of EV1 and LP3 distributions have been very common in flood analysis. The TCEV distribution, a 4-parameter model, is of recent origin (Rossi, et al., 1984). The distribution emerges as a mixture of two EV1 distributions, and can account for the outliers in flood series. It offers physical justification to the stochastic process of floods, at least partially, and appears suitable for regional analysis.

Several estimation methods exist for EV1 and LP3 models. Some of these methods have been in use for a long time and are termed here as classical methods. Several other methods have been proposed relatively recently along with various claims about their performance.

It is important to evaluate the performance of all available estimators of a model, especially for small samples, for which the variability of estimators is quite large and so is the marked difference in their performance. To minimize design losses, one would like to use the most efficient estimator.

Approximate formulae can be derived for asymptotic standard error of several of the estimators. But one is chiefly interested in the sampling properties of the estimators for rather small sample sizes (≤ 50) not covered by the asymptotic formulae. The sampling distribution of the estimator is generally intractable in the sample range of interest. Monte Carlo sampling experiments, therefore, offer an attractive procedure for evaluating and comparing the performance of estimators. Cunnane (1986) points out that the simulation experiments have been reported in most recent work on flood frequency analysis. Thus, the use of simulation has become a standard technique to evaluate the performance of competing estimators.

Using random sampling experiments, the estimators of EV1 and LP3 distributions are assessed in an at-site framework, while the TCEV estimator is investigated in a regional framework.

Chapter 2

LITERATURE REVIEW

The field of flood frequency analysis has rapidly developed since the late 1960's. Mos and Kirby (1985) have traced the principal aspects of American practice since the 1860's, while at the other end, Greis (1983) reviewed the developments in flood frequency analysis in the early 1980's. Lettenmaier (1985) has reviewed recent developments in regional flood frequency analysis. Recently Cunnane (1986) has outlined the selected developments of flood frequency analysis under a number of different headings such as AM model development, regional flood estimates, parameter estimation, and simulation. In the following survey, an attempt has been made to review the literature that is directly or indirectly relevant to the present work.

2.1 The Probability Models and Estimators

A number of empirical models have been suggested for flood frequency analysis. One of the oldest among them to find hydrological acceptance is the Gumbel's extreme value type 1 (EV1) distribution (Gumbel, 1958). According to Lettenmaier (1982), the EV1 distribution has a certain intuitive appeal. It provides partial physical justification to the stochastic processes controlling floods (it can be seen to arise as the asymptotic distribution of extremes, where the process generating the extreme falls into the class of distributions having exponential tails). From a statistical standpoint, the EV1 is attractive, since the distribution has only two parameters which can be estimated rather easily, and the inverse explicit form of the cumulative distribution function permits easy estimation of quantiles. The EV1 distribution has a constant skew (= 1.14). Matalas, et al. (1975) found, in a comprehensive study of the U.S. flood records, that for most regions considered, the regional mean skew coefficient was within one standard deviation of the EV1 value.

Several methods have been proposed for estimating the parameters and quantiles of the EV1 model. While the methods of moment (MOM) and maximum likelihood estimation (MLE) have been in use for a long time, several new methods of estimation have been proposed fairly recently. The method of probability weighted moments (PWM), introduced by Greenwood, et al. (1979), was used by Landwehr and Matalas, et al. (1979) to estimate Gumbel parameters and quantiles. Based on the principle (Jaynes, 1961, 1982) of maximum entropy, Jowitt (1979) suggested entropy estimators (ENT). The least squares estimators (LEAS) were used by Chow (1953), and Lowery and Nash (1970), and have since been applied to the EV1 distribution by Jain and Singh (1986). Houghton (1978) introduced the method of incomplete means (ICM), which is easy to apply for EV1 distribution (Jain and Singh, 1986). Fiorentino and Gabriele (1984) derived a bias-corrected MLE estimator and compared it to other estimators through sampling experiments. Lettenmaier and Burges (1982) took a critical look at Gumbel's fitting method reported in many textbooks (Linsley, et al., 1975). Greis and Wood (1981) proposed a regional procedure based on the EV1/PWM.

Some of the other distributions that have received attention in flood frequency analysis are the general extreme value (GEV), log Pearson 3 (LP3), log normal 2 and 3 (LN2, LN3), the five parameter Wakeby (WAK), and the recently developed two component extreme value (TCEV). These distributions are popular because they are all capable of modeling the positively skewed, thick-tailed distributions from which the AM floods are believed to arise.

The GEV distribution was recommended for the British rivers by the Flood Studies Report (NERC, 1975a). The EV1 distribution emerges as a special case of the GEV distribution (Hall, 1984). Hosking, et al. (1985) proposed a GEV/PWM regionalization procedure and compared it with other procedures using sampling experiments. Nevertheless, the GEV distribution fails to account for the separation phenomenon (Cunnane, 1986).

The U.S. Water Resources Council (USWRC) (Bulletin 15, 1967), based on a comprehensive study reported in Benson (1968), recommended the use of the LP3 distribution as a base method of flood frequency analysis for U.S. floods. It was noted (Benson, 1968) that the choice was not fully conclusive from a statistical standpoint, but rather was made to fulfill the demonstrated need for standardization and uniformity, especially among the federal agencies. The method of estimation proposed in Bulletin 15 (1967), and referred to in this study as the indirect method of moment (MMI), is basically the method of moments applied to the log-transformed data in which the mean, variance, and skewness estimates of the log-data are equated to the population values, and the resulting equations solved to estimate the parameters. The sample skewness estimate used in parameter estimation is a statistic with peculiar properties. It is highly variable, is considerably downward biased, and algebraically bounded (Wallis, et al., 1974; Kirby, 1974; Lettenmaier and Burges, 1980). Hence, the estimation method proposed by the USWRC is highly susceptible to these problems in the sample skewness estimate. To circumvent this problem, the skewness estimate was modified in USWRC

Bulletin 17A (1977) as a sum of the sample skewness estimate and regional skewness coefficient weighted by a function of the sample size. The USWRC (Bulletin 17B, 1981) has again modified the skewness estimate and re-expressed it as a weighted sum for all sample sizes.

The endorsement of the LP3 distribution by the U.S. Water Resources Council initiated considerable research about the advisability of using the LP3 distribution, and in general, boosted research in other directions of flood frequency analysis (Greis, 1983). Bobee (1975) derived the density function of the distribution and investigated its properties. He also derived a new method of parameter estimation, referred to here as the direct method of moments (MMD), and recommended its usage in preference to MMI of USWRC. Bobee showed that the LP3 density function was capable of assuming diverse shapes ranging from reverse J to U, and could have upper or lower bounds depending upon the parameter values. A comparison of different estimators of the LP3 was made by Bobee and Robitaille (1977) based on annual flood series. They found that MMD provided a better fit to the data than MMI. Condie (1977) derived the MLE estimator and the asymptotic standard error of its quantile estimate for the LP3 distribution. Nozdryn-Plotnicki and Watt (1979) carried out sampling experiments and tested the performance of several estimation methods of LP3 distribution over the range of parameter values encountered for Canadian rivers.

In a series of papers, Rao (1980a, 1980b, 1981, 1983, 1986) investigated very comprehensively the properties and estimators of the LP3 and several other three parameter distributions (1981). Noting the problems with estimation of the skewness from a sample, he argued that the MMD and MMI utilizing the skew estimate, are likely to result in poor estimation. To obviate the problem, he suggested a new estimation method, referred to by him as the method of mixed moments (MIX). MIX combines the first two moments of the real and the log-transformed data, and thus avoids the use of skewness. Rao carried out sampling experiments and, over the range tested by him, found MIX to be superior to the other methods (MMD, MMI, MLE). Hoshi and Burges (1981) derived the asymptotic sampling covariance structure of the MMD parameter and quantile estimates. Lettenmaier and Burges (1980) derived the bias correction factors of the skewness estimator of the LP3 and found them to be a function of not only the sample size but also the coefficient of variation and skewness. Singh and Singh (1985) derived a new estimator based on the principle of maximum entropy. The estimator is referred to here as the entropy estimator (ENT).

The two component extreme value distribution (TCEV) has been proposed recently by Rossi, et al. (1984). The distribution could account for the condition of separation in Italian AM series, and as such, is a viable contender for the parent distribution of annual floods. Rossi, et al. (1984) proposed a regional algorithm using the TCEV/MLE.

2.2 Sampling Experiments in Flood Analysis

Monte Carlo based sampling experiments have come to be of increasing use in flood frequency research. Wallis, et al. (1974) carried out extensive sampling experiments on commonly used sample statistics. They found pronounced bias, skew, and bounds in the sampling distributions of the statistics. For instance, skewness was found to be significantly down-biased for all distributions investigated by them. Kirby (1974) later proved that the sample skew estimate are algebraically bounded, and the bounds are independent of the distribution used. Sampling experiments have become a standard tool in assessing the performance of various competing estimators under a range of underlying populations. In a classic investigation, Matalas, et al. (1975) observed the condition of separation in AM series, and found on the basis of sampling experiments that the commonly used distributions could not account for separation. Cunnane (1986) summarizes the simulation experiments reported in a host of studies on various facets of flood frequency analysis.

Chapter 3

METHODOLOGY

The frequency analysis procedure involves selecting a probabilistic model and estimating its parameters from a suitable estimator. Many probability models have been employed for flood analysis. At the same time, several estimators have been proposed for each model. There is no general agreement on the superiority of a probability model. Investigators have fitted several models to various data and have found that no model fits uniformly better under all situations (Beard, 1974; Boughton, 1980; Houghton, 1978; Jain, 1986).

Likewise, there is no general agreement on the estimator to be used. Traditionally, the use of the method of moments has been justified on the grounds of computational simplicity. However, a number of estimators (e.g., maximum likelihood) which were previously computationally intractable, have now become feasible, due largely to enhanced computational facilities available these days.

It is emphasized here that choosing the estimator of the model can be as important as selecting the model itself, for several estimators of a model can lead to parameter and quantile estimates widely differing in variability (MSE) and bias. Hence, it is important to study the performance of various estimators of each probability model, especially for small samples encountered in practice. The quantile estimates derived from at-site analysis are highly variable. This variability increases as the sample size decreases and the return period, T, becomes large. In this situation, it becomes important to search for the most efficient estimator. This would ensure that the best estimator extracting maximum useful information from the data sample is utilized in hydrologic analysis. It is hoped that an estimator of a model performing well in its own population would behave robustly with respect to an unknown population, as would be the case in practice.

The objective of this study is to evaluate the performance of estimators of several commonly employed probability models in hydrology. The estimators being random variables, their properties can be characterized by statistical performance indices such as bias and mean square error. In this study, the performance indices of the estimators are computed by using random sampling experiments: A large number of synthetic samples are generated from population of a probability model, and parameters and quantiles estimated from each estimator. These estimates are used to calculate the performance indices of estimators.

The simulation approach offers an objective and practical procedure for evaluating the performance of estimators for small samples usually encountered in hydrology. A theoretical approach to the problem would be to derive analytically the sampling distribution of each estimator and evaluate the performance of the estimator from this distribution. This, in general, is a mathematically very complex task and, hence, not a viable approach. The simulation approach is clearly superior to using the real world data because the real world data has an unknown parent distribution that introduces model errors, and is also prone to measurement errors.

Besides the methods of moments and maximum likelihood estimation, several new estimators have been developed over the past few years. The method of probability weighted moments, introduced by Greenwood, et al. (1979), and is suitable for distributions with explicit inverse forms.

Based on the principle of maximum entropy (Jaynes, 1961), Jowitt (1979), and Singh and Singh (1985), suggested entropy estimators. The least squares estimators were used by Chow (1953), Lowery and Nash (1970), and Jain and Singh (1986), among others. Houghton (1978) introduced the method of incomplete means. Rao (1980b) developed the method of mixed moments in the context of the log Pearson type 3 distribution. These estimators will be used, wherever applicable, and their statistical properties calculated by the simulation approach.

The probability models proposed to be analyzed are Gumbel (EV1), log Pearson type 3 (LP3) and two component extreme value (TCEV) distributions. These are two, three and four parameter models respectively. The EV1 is the oldest and a widely used model in engineering practice, while the TCEV is the most recently developed model.

3.1 Performance Criteria of Estimators

The estimators are random variables. Hence, their performance has to be assessed statistically. Let, $\hat{\theta}$ be an estimator for parameters or quantiles, and $E(\cdot)$ denote statistical expectation. The performance statistics of $\hat{\theta}$ are defined as:

$$\text{Bias:} \quad BIAS(\hat{\theta}) = E(\hat{\theta}) - \theta \tag{3.1}$$

$$\text{Variance:} \quad STD^2(\hat{\theta}) = E[\hat{\theta} - E(\hat{\theta})]^2 \tag{3.2}$$

$$\text{Mean Square Error:} \quad MSE(\hat{\theta}) = E(\hat{\theta} - \theta)^2 \tag{3.3}$$

$$= BIAS^2 + STD^2$$

In many applications, the performance statistics are normalized by dividing them by the population values (e.g., in LP3 analysis in this work).

The expectation and variance of the estimators could be calculated by following the procedure below.

3.2 Experimental Design

It is desired to estimate the statistics of $\hat{\theta}$ through random sampling experiments based on Monte Carlo simulation. The following procedure is suggested:

1. From the model under consideration, generate N Monte Carlo samples of size n.
2. Estimate $\hat{\theta}$ for each of the N samples using the available estimators.

$$E(\hat{\theta}) \cong \sum_{i=1}^{N} \hat{\theta}_i / N \tag{3.4}$$

$$STD^2(\hat{\theta}) \cong \sum_{i=1}^{N} [\hat{\theta}_i - E(\hat{\theta})]^2/(N-1) \tag{3.5}$$

Having estimated the expectation and variance, the BIAS and MSE of $\hat{\theta}$ can be computed from (3.1) and (3.3).

3.3 Probability Models

3.3.1 Gumbel Distribution (EV1)

The cumulative distribution function (cdf) is given as:

$$F(x) = \exp[-\exp\{-a(x-b)\}] \tag{3.6}$$

where 1/a and b are the scale and location parameters respectively. The EV1 distribution represents the asymptotic distribution of maximum of independent and identically distributed (iid) flood peaks, when they are assumed to follow exponential distribution, and their occurrence approaches infinity, and hence provides partial physical justification to the flood phenomena.

3.3.2 Log Pearson Type 3 (LP3)

If Y(= ln X) is distributed as a Pearson Type 3, then X is distributed as an LP3 variable. The probability density function (pdf) of X is given as:

$$f(x) = \frac{1}{|a| x \Gamma b} \left[\frac{\ln x - c}{a}\right]^{b-1} \exp\left[- \left(\frac{\ln x - c}{a}\right)\right] \qquad (3.7)$$

where a, b and c are the scale, shape and location parameters respectively. It was recommended by the U.S. Water Resources Council (1981) for flood frequency analysis.

3.3.3 Two Component Extreme Value Distribution (TCEV)

The pdf is given as

$$f(x) = \left[\frac{\Lambda_1}{\theta_1} \exp(- x/\theta_1) + \frac{\Lambda_2}{\theta_2} \exp(- x/\theta_2)\right]$$

$$\exp[- \Lambda_1 \exp(- x/\theta_1) - \Lambda_2 \exp(- x/\theta_2)], \ x > 0$$

$$= \exp(- \Lambda_1 - \Lambda_2), \ x = 0 \qquad (3.8)$$

where $\Lambda_1 > 0$, $\Lambda_2 \geq 0$, $\theta_2 \geq \theta_1 \geq 0$ are parameters. This distribution, being a four-parameter distribution, is suitable for regional analysis (Rossi, et al., 1984).

3.4 Estimators

The following estimators have been reported in the literature from time to time:

1. Classical
   - a. Method of Moments (MOM)
   - b. Maximum Likelihood (MLE)
   - c. Least Squares (LEAS)
2. Recent
   - a. Probability Weights Moments (PWM)
   - b. Entropy (ENT)
   - c. Incomplete Means (ICM)
   - d. Mixed Moments (MIX)

Each of these methods is based on a different philosophy and it is not possible to compare their performance without undertaking detailed simulation studies.

Chapter 4

GUMBEL'S EXTREME VALUE 1 (EV1) DISTRIBUTION

4.1 Introduction

The EV1 or Gumbel distribution has been widely used and still continues to be employed, particularly in developing countries. It is a rather simple distribution involving only two parameters and lends itself well to parameter estimation, particularly for small samples. The explicit inverse form of the distribution facilitates straightforward quantile estimation, once the parameters have been estimated.

The cumulative distribution function (cdf) of an extreme value type 1 (EV1) or Gumbel random variable x is given by:

$$F(x) = \exp[- \exp\{- a(x-b)\}] \qquad (4.1)$$

where 1/a and b are respectively the scale and location parameters of the distribution. The methods of moments, least squares, and maximum likelihood estimation have been the classical methods to estimate the parameters a and b and consequently to estimate the quantiles. However, several relatively new methods have been developed and successfully applied over the past few years. The method of probability weighted moments, introduced by Greenwood, et al. (1979), was used by Landwehr and Matalas (1979) to estimate Gumbel parameters and quantiles. Jowitt (1979) suggested entropy estimators. The least squares estimators were used by Chow (1953), and Lowery and Nash (1970) and have since been applied to the EV1 distribution by Jain and Singh (1986). Houghton (1978) introduced the method of incomplete means, which is easy to apply for EV1 distribution. Similarly, the method of mixed moments used by

Rao (1980, 1983) is straightforward and yields closed form algebraic expressions for parameter estimators (Jain, 1986).

Here, the performance of the aforementioned methods has been evaluated and compared using Monte Carlo simulation. This work is an extension of the one by Landwehr and Matalas (1979), where the authors used sampling experiments to compare the method of probability weighted moments with the method of moments and maximum likelihood estimation in two cases: purely random samples and serially correlated samples. The present work is made comprehensive by including all methods that are apparently available to date.

Additionally, this work also addressed the question of bias correction for method of moments-quantile estimator. The moment estimator has been widely used, owing to its simplicity. However, as investigated by Matalas, et al. (1979) and Lettenmaier, et al. (1982) among others, and also corroborated by this work, this estimator yields biased estimates of the quantile.

Usually, selection of the best estimator is governed by the type of the loss function which is a measure of the loss resulting from over or under-design. In certain situations of design, the loss function is minimized by least-biased estimator. Moreover, the bias corrected moment estimator of quantile, if not accompanied by a significant worsening of the mean square error (MSE), can prove to be useful in regional estimation procedures where, due to the large sample size, the variance of the estimator is no longer the main contribution to its MSE.

Sampling experiments were used to arrive at a practically unbiased moment-quantile estimator. This bias corrected estimator yielded nearly unbiased estimates of the quantiles, even for samples of small size.

Moreover, it did not entail any practical worsening of the MSE. Indeed, as is shown by simulation, the MSE values obtained from the bias-corrected estimator were close to those from the uncorrected estimator.

4.2 Analysis

The inverse form of (1) is given by

$$x(F) = b - \frac{\ln(-\ln F)}{a} \tag{4.2}$$

where x(F) denotes the quantile of cumulative non-exceedence probability F. It is noted that the return period of x(F) equals 1/(1-F). For sample sizes n = 5, 10, 15, 20, 30, 50, 100, 1000, the parameters a and b were estimated by the methods of moments (MOM), maximum likelihood estimation (MLE), probability weighted moments (PWM), entropy (ENT), mixed moments (MIX), least squares (LEAS), and incomplete means (ICM). The quantiles x(F), for F = 0.05, . . . , 0.99, were then calculated from (4.2).

The performance statistics of various estimators were estimated through Monte Carlo sampling experiments, i.e., through generation of a large number of synthetic samples for: (1) purely random process (independent and identically distributed Gumbel random variables), and (2) serially correlated process (with the first order serial correlation coefficient of 0.5). These estimates are expected to be very close approximations to the theoretical values owing to the large value of N (= 50,000 for n = 5, . . . ,100, and, = 10,000 for n = 1000).

The mean square error of all methods relative to that of MLE was compared using the relative efficiency defined as:

$$EFF(\hat{\theta}) = \frac{MSE(\hat{\theta}|MLE)}{MSE(\hat{\theta}|\text{other method})} \tag{4.3}$$

A value of $EFF(\hat{\theta}) < 1$ implies that the method under consideration is less efficient (i.e., has higher mean square error) compared to MLE and vice versa.

4.3 Estimation of Parameters and Quantiles

The parameter estimation equations for various methods are summarized as follows:

(1) MOM: $\hat{a} = \pi/(\sqrt{6} \cdot \sigma_x)$

$$\hat{b} = \bar{x} - 0.57721/\hat{a} \qquad (4.4)$$

where σ_x is standard deviation.

(2) MLE: $$\hat{a} = \frac{\Sigma \exp(-\hat{a}x_i)}{\bar{x} \cdot \Sigma \exp(-\hat{a}x_i) - \Sigma x_i \exp(-\hat{a}x_i)}$$

$$\hat{b} = \frac{1}{\hat{a}} \cdot \ln \left[\frac{n}{\Sigma \exp(-\hat{a}x_i)}\right] \qquad (4.5)$$

(3) PWM: $$\hat{a} = \frac{\ln(2)}{\bar{x} - 2 \cdot \Sigma x_i(n-i)/n(n-1)}$$

$$\hat{b} = x - 0.57721/\hat{a} \qquad (4.6)$$

(4) ENT: $$\hat{a} = \frac{0.55721 + \ln[n/\Sigma \exp(-\hat{a}x_i)]}{\bar{x}}$$

$\hat{b}$ = same formula as for MLE (4.7)

(5) LEAS: $$\hat{a} = \frac{n \cdot \Sigma z_i x_i - \Sigma x_i \Sigma z_i}{(\Sigma x_i)^2 - n \cdot \Sigma x_i^2}$$

$$\hat{b} = \bar{x} + \Sigma z_i/(n \cdot \hat{a}) \qquad (4.8)$$

where $z_i = \ln[-\ln(F_p(x_i)]$ is obtained from the plotting position formula which defines the cumulative probability of non-exceedance $F_p(x_i)$ for each data point x_i.

(6) MIX: $\hat{a} = 1.28255/\sigma_x$

$$\hat{b} = \frac{1}{\hat{a}} \cdot \ln[1 + \hat{a} \cdot \bar{x} + \frac{\hat{a}^2}{2} \cdot \frac{\Sigma x_i^2}{n}] \qquad (4.9)$$

(8) ICM: $$\bar{x}_i = \hat{b} - \frac{n}{\hat{a}(n-n_i)} \cdot [J \ln J - J^2 \ln J/2$$

$$+ J^3 \ln J/6 - J^4 \ln j/24 + \ldots], \; i = 1, 2 \qquad (4.10)$$

where $J = \ln(n/n_i)$, n is the sample size, and n_i is the number of observations on which the incomplete mean $\bar{x}_i$ is calculated.

All the summations above were performed over i = 1 to n and the sample statistics were calculated as:

$$\bar{x} = \Sigma \; x_i/n \qquad (4.11)$$

$$\sigma_x^2 = \Sigma \; (x_i - \bar{x})^2/(n-1) \qquad (4.12)$$

Landwehr and Matalas (1979) performed their calculations with the biased expression of σ_x^2, i.e., with n replacing n-1 in (4.12).

Having estimated parameters a and b, the respective quantiles can be calculated from (4.2) as:

$$\hat{x} = \hat{b} - \ln(- \ln F)/\hat{a} \qquad (4.13)$$

4.4 Comparison of Performance Statistics of the Parameter and Quantiles Estimators

4.4.1 Case 1: Purely Random Process

The results of this case are shown in Tables 4.1-4.2.

4.4.1.1 Parameter Estimates

The MIX and ICM can be prima-facie rejected as unreliable estimators of Gumbel parameters. MIX, while performing reasonably efficiently for estimation of 'a', failed in providing even moderately biased estimate of 'b', and thus resulted in highly inefficient estimate of 'b'. ICM failed to provide satisfactory estimate for either 'a' or

TABLE 4.1

PERFORMANCE STATISTICS OF PARAMETER ESTIMATES : RANDOM SAMPLES

METHOD	SAMPLE SIZE	BIAS(A)	STD(A)	EFF.(A)	BIAS(B)	STD(B)	EFF.(B)
MOM	5	-0.360	0.778	1.139	-0.051	0.485	1.017
MLE		-0.434	0.806	1.000	-0.081	0.485	1.000
PWM		-0.233	0.701	1.533	0.000	0.480	1.049
ENT		-0.415	0.790	1.052	-0.083	0.485	0.997
LEA		-0.141	0.650	1.893	-0.019	0.484	1.032
MIX		-0.360	0.778	1.139	-0.834	0.497	0.257
ICM		-5.274	111.876	0.000	-0.468	0.874	0.246
MOM	10	-0.156	0.381	0.881	-0.028	0.341	0.991
MLE		-0.167	0.349	1.000	-0.040	0.338	1.000
PWM		-0.092	0.353	1.126	-0.001	0.336	1.024
ENT		-0.159	0.348	1.024	-0.040	0.338	0.996
LEA		-0.026	0.341	1.279	0.003	0.341	0.994
MIX		-0.156	0.381	0.881	-0.884	0.371	0.126
ICM		-0.532	1.489	0.060	-0.065	0.930	0.133
MOM	15	-0.103	0.295	0.782	-0.020	0.278	0.981
MLE		-0.102	0.256	1.000	-0.026	0.275	1.000
PWM		-0.058	0.268	1.010	-0.001	0.274	1.015
ENT		-0.097	0.258	1.004	-0.027	0.275	0.996
LEA		-0.005	0.268	1.059	0.007	0.279	0.979
MIX		-0.103	0.295	0.782	-0.902	0.309	0.084
ICM		-0.280	0.623	0.163	-0.005	0.770	0.129
MOM	20	-0.076	0.249	0.739	-0.015	0.239	0.974
MLE		-0.074	0.211	1.000	-0.019	0.236	1.000
PWM		-0.041	0.225	0.953	-0.000	0.235	1.010
ENT		-0.069	0.213	0.995	-0.019	0.236	0.996
LEA		0.004	0.231	0.938	0.010	0.241	0.966
MIX		-0.076	0.249	0.739	-0.911	0.269	0.062
ICM		-0.191	0.482	0.186	0.007	0.650	0.132
MOM	30	-0.051	0.200	0.688	-0.010	0.196	0.968
MLE		-0.047	0.165	1.000	-0.013	0.192	1.000
PWM		-0.026	0.180	0.888	-0.000	0.192	1.005
ENT		-0.044	0.168	0.979	-0.013	0.193	0.995
LEA		0.009	0.190	0.812	0.010	0.197	0.957
MIX		-0.051	0.200	0.688	-0.921	0.225	0.041
ICM		-0.121	0.365	0.199	0.005	0.502	0.147

TABLE 4.1 (CONTINUED)

METHOD	SAMPLE SIZE	BIAS(A)	STD(A)	EFF.(A)	BIAS(B)	STD(B)	EFF.(B)
MOM	50	-0.030	0.156	0.657	-0.007	0.151	0.960
MLE		-0.027	0.126	1.000	-0.008	0.148	1.000
PWM		-0.015	0.139	0.846	-0.001	0.148	1.002
ENT		-0.025	0.129	0.968	-0.008	0.148	0.995
LEA		0.012	0.151	0.722	0.008	0.152	0.947
MIX		-0.030	0.156	0.657	-0.929	0.177	0.024
ICM		-0.069	0.269	0.215	0.003	0.368	0.161
MOM	75	-0.020	0.129	0.652	-0.004	0.124	0.955
MLE		-0.017	0.104	1.000	-0.005	0.121	1.000
PWM		-0.009	0.115	0.835	-0.000	0.121	0.999
ENT		-0.016	0.106	0.967	-0.005	0.121	0.995
LEA		0.011	0.126	0.689	0.008	0.124	0.942
MIX		-0.020	0.129	0.652	-0.932	0.146	0.016
ICM		-0.043	0.214	0.231	0.003	0.294	0.169
MOM	100	-0.014	0.113	0.653	-0.003	0.107	0.956
MLE		-0.012	0.091	1.000	-0.004	0.105	1.000
PWM		-0.007	0.101	0.832	-0.000	0.105	1.000
ENT		-0.012	0.093	0.966	-0.004	0.105	0.994
LEA		0.011	0.112	0.678	0.007	0.108	0.943
MIX		-0.014	0.113	0.653	-0.934	0.129	0.012
ICM		-0.032	0.184	0.243	0.002	0.252	0.174
MOM	1000	-0.000	0.048	0.798	-0.000	0.034	0.956
MLE		-0.000	0.043	1.000	-0.001	0.033	1.000
PWM		0.000	0.045	0.928	-0.000	0.033	0.999
ENT		-0.000	0.043	1.000	-0.001	0.034	0.992
LEA		0.004	0.048	0.811	0.001	0.034	0.952
MIX		-0.000	0.048	0.794	-0.941	0.051	0.001
ICM		-0.002	0.065	0.439	-0.001	0.077	0.189

TABLE 4.2 - BIAS IN QUANTILE ESTIMATES ; RANDOM SAMPLES

(F = PROB. OF NON EXCEEDENCE, X = QUANTILE)

METHOD	SAMPLE SIZE	F = 0.001	0.010	0.020	0.050	0.100	0.250	0.500	0.750	0.900	0.950	0.980	0.990	0.999
		X = -1.933	-1.527	-1.364	-1.097	-0.834	-0.327	0.367	1.246	2.250	2.970	3.902	4.600	6.907
CMOM	5	-0.014	-0.009	-0.006	-0.005	-0.005	-0.004	-0.001	-0.001	0.012	0.016	0.020	0.021	0.024
MOM		-0.231	-0.189	-0.174	-0.150	-0.127	-0.081	-0.017	0.061	0.161	0.229	0.315	0.378	0.586
MLE		-0.390	-0.322	-0.297	-0.255	-0.214	-0.134	-0.022	0.116	0.283	0.400	0.549	0.660	1.025
PWM		-0.010	-0.005	-0.002	-0.001	-0.001	-0.001	0.002	0.002	0.014	0.018	0.020	0.022	0.024
ENT		-0.367	-0.304	-0.280	-0.242	-0.204	-0.131	-0.029	0.098	0.252	0.360	0.497	0.599	0.934
LEA		0.135	0.105	0.093	0.073	0.051	0.007	-0.049	-0.124	-0.198	-0.256	-0.334	-0.392	-0.589
MIX		-1.008	-0.972	-0.958	-0.934	-0.909	-0.863	-0.801	-0.714	-0.618	-0.552	-0.468	-0.405	-0.198
ICM		-1.181	-1.032	-0.972	-0.874	-0.776	-0.588	-0.333	-0.010	0.366	0.635	0.982	1.240	2.092
CMOM	10	-0.009	-0.004	-0.001	-0.001	-0.001	-0.002	0.000	-0.001	0.010	0.013	0.014	0.015	0.014
MOM		-0.129	-0.104	-0.094	-0.081	-0.069	-0.045	-0.010	0.032	0.090	0.129	0.175	0.210	0.320
MLE		-0.198	-0.161	-0.147	-0.126	-0.106	-0.066	-0.010	0.058	0.146	0.206	0.281	0.336	0.516
PWM		-0.012	-0.007	-0.004	-0.003	-0.003	-0.003	0.000	0.001	0.013	0.017	0.020	0.021	0.024
ENT		-0.182	-0.148	-0.135	-0.117	-0.099	-0.064	-0.014	0.046	0.126	0.179	0.246	0.295	0.456
LEA		0.135	0.109	0.100	0.083	0.064	0.026	-0.023	-0.088	-0.150	-0.200	-0.266	-0.317	-0.487
MIX		-0.978	-0.959	-0.951	-0.938	-0.925	-0.900	-0.867	-0.815	-0.762	-0.726	-0.680	-0.647	-0.537
ICM		-0.224	-0.187	-0.173	-0.152	-0.132	-0.092	-0.036	0.033	0.123	0.183	0.258	0.314	0.496
CMOM	15	-0.009	-0.004	-0.001	0.000	-0.001	-0.001	0.001	0.000	0.011	0.014	0.016	0.016	0.015
MOM		-0.093	-0.074	-0.066	-0.057	-0.049	-0.032	-0.006	0.023	0.067	0.095	0.128	0.153	0.231
MLE		-0.134	-0.108	-0.097	-0.084	-0.070	-0.044	-0.006	0.038	0.101	0.141	0.192	0.228	0.348
PWM		-0.012	-0.007	-0.004	-0.003	-0.003	-0.002	0.001	0.001	0.013	0.017	0.020	0.022	0.024
ENT		-0.122	-0.099	-0.089	-0.077	-0.065	-0.042	-0.009	0.030	0.087	0.123	0.167	0.200	0.305
LEA		0.115	0.095	0.088	0.074	0.058	0.026	-0.014	-0.068	-0.117	-0.158	-0.213	-0.255	-0.396
MIX		-0.969	-0.955	-0.950	-0.940	-0.930	-0.913	-0.890	-0.850	-0.812	-0.786	-0.753	-0.730	-0.652
ICM		-0.080	-0.061	-0.052	-0.043	-0.035	-0.017	0.008	0.038	0.083	0.111	0.145	0.170	0.249
CMOM	20	-0.007	-0.002	0.000	0.001	0.000	-0.001	0.001	-0.001	0.009	0.012	0.013	0.013	0.010
MOM		-0.073	-0.057	-0.050	-0.043	-0.037	-0.024	-0.004	0.018	0.054	0.076	0.102	0.121	0.180
MLE		-0.103	-0.082	-0.073	-0.063	-0.053	-0.033	-0.004	0.030	0.079	0.111	0.149	0.177	0.268
PWM		-0.011	-0.006	-0.003	-0.002	-0.002	-0.002	0.001	0.001	0.013	0.017	0.020	0.021	0.023
ENT		-0.093	-0.074	-0.066	-0.057	-0.049	-0.032	-0.006	0.023	0.068	0.096	0.130	0.154	0.233
LEA		0.102	0.085	0.079	0.068	0.053	0.026	-0.009	-0.056	-0.098	-0.133	-0.181	-0.218	-0.341
MIX		-0.962	-0.952	-0.948	-0.941	-0.933	-0.919	-0.902	-0.869	-0.839	-0.819	-0.794	-0.776	-0.717
ICM		-0.041	-0.028	-0.022	-0.016	-0.011	-0.001	0.015	0.033	0.064	0.082	0.103	0.118	0.166

TABLE 4.2 (CONTINUED)

METHOD	SAMPLE SIZE	F = 0.001	0.010	0.020	0.050	0.100	0.250	0.500	0.750	0.900	0.950	0.980	0.990	0.999
		X = -1.933	-1.527	-1.364	-1.097	-0.834	-0.327	0.367	1.246	2.250	2.970	3.902	4.600	6.907
CMOM	30	-0.006	-0.002	0.001	0.001	0.001	0.000	0.001	-0.001	0.008	0.011	0.011	0.011	0.008
MOM		-0.053	-0.041	-0.035	-0.030	-0.026	-0.017	-0.002	0.013	0.042	0.058	0.076	0.090	0.131
MLE		-0.072	-0.056	-0.049	-0.042	-0.035	-0.022	-0.002	0.020	0.057	0.080	0.106	0.125	0.185
PWM		-0.011	-0.006	-0.003	-0.002	-0.002	-0.002	0.001	0.001	0.013	0.017	0.020	0.021	0.024
ENT		-0.065	-0.051	-0.044	-0.038	-0.032	-0.021	-0.003	0.016	0.049	0.069	0.092	0.109	0.161
LEA		0.082	0.070	0.065	0.056	0.045	0.023	-0.004	-0.043	-0.074	-0.101	-0.139	-0.168	-0.266
MIX		-0.957	-0.950	-0.947	-0.942	-0.936	-0.926	-0.915	-0.888	-0.866	-0.852	-0.834	-0.821	-0.781
ICM		-0.030	-0.020	-0.015	-0.010	-0.007	-0.000	0.011	0.022	0.046	0.060	0.074	0.085	0.116
CMOM	50	-0.007	-0.003	0.000	0.001	0.000	-0.001	0.001	0.000	0.010	0.012	0.014	0.014	0.012
MOM		-0.038	-0.028	-0.023	-0.020	-0.017	-0.012	-0.001	0.008	0.030	0.042	0.054	0.063	0.089
MLE		-0.048	-0.037	-0.031	-0.027	-0.023	-0.015	-0.001	0.012	0.040	0.055	0.072	0.084	0.122
PWM		-0.012	-0.007	-0.004	-0.003	-0.003	-0.002	0.001	0.001	0.013	0.017	0.020	0.022	0.024
ENT		-0.044	-0.033	-0.028	-0.024	-0.021	-0.014	-0.002	0.010	0.035	0.048	0.063	0.074	0.106
LEA		0.058	0.050	0.048	0.042	0.033	0.017	-0.002	-0.030	-0.050	-0.069	-0.096	-0.117	-0.188
MIX		-0.952	-0.948	-0.946	-0.943	-0.939	-0.932	-0.925	-0.905	-0.889	-0.879	-0.867	-0.859	-0.834
ICM		-0.022	-0.014	-0.010	-0.007	-0.005	-0.001	0.007	0.013	0.033	0.042	0.052	0.058	0.078
CMOM	75	-0.008	-0.004	-0.001	0.000	0.000	-0.001	0.002	0.001	0.012	0.015	0.017	0.017	0.018
MOM		-0.029	-0.021	-0.016	-0.014	-0.012	-0.008	0.000	0.007	0.026	0.035	0.045	0.051	0.071
MLE		-0.036	-0.026	-0.022	-0.018	-0.015	-0.010	0.000	0.010	0.032	0.044	0.057	0.065	0.092
PWM		-0.012	-0.006	-0.003	-0.002	-0.002	-0.002	0.002	0.002	0.014	0.019	0.022	0.024	0.027
ENT		-0.033	-0.024	-0.020	-0.017	-0.014	-0.009	-0.000	0.008	0.029	0.039	0.051	0.058	0.081
LEA		0.043	0.038	0.037	0.033	0.026	0.014	0.000	-0.021	-0.033	-0.047	-0.067	-0.082	-0.136
MIX		-0.950	-0.947	-0.945	-0.943	-0.939	-0.935	-0.930	-0.912	-0.899	-0.892	-0.883	-0.877	-0.859
ICM		-0.016	-0.009	-0.005	-0.003	-0.002	0.001	0.006	0.011	0.027	0.034	0.041	0.046	0.058
CMOM	100	-0.009	0.004	-0.001	0.000	0.001	-0.001	0.001	0.001	0.012	0.015	0.017	0.018	0.017
MOM		-0.025	-0.017	-0.013	-0.011	-0.009	-0.006	0.000	0.005	0.023	0.031	0.039	0.045	0.060
MLE		-0.030	-0.021	-0.017	-0.014	-0.012	-0.008	0.001	0.008	0.028	0.038	0.048	0.055	0.076
PWM		-0.012	-0.006	-0.003	-0.002	-0.002	-0.002	0.002	0.002	0.014	0.019	0.022	0.024	0.027
ENT		-0.028	-0.020	-0.016	-0.013	-0.011	-0.007	0.000	0.006	0.025	0.034	0.043	0.049	0.067
LEA		0.034	0.031	0.031	0.027	0.022	0.012	0.001	-0.017	-0.025	-0.035	-0.051	-0.064	-0.108
MIX		-0.948	-0.946	-0.945	-0.943	-0.940	-0.936	-0.932	-0.916	-0.905	-0.899	-0.892	-0.887	-0.872
ICM		-0.016	-0.009	-0.006	-0.004	-0.003	-0.001	0.005	0.008	0.024	0.030	0.037	0.041	0.052
MOM	1000	-0.003	-0.003	-0.002	-0.002	-0.002	-0.001	-0.000	0.002	0.002	0.003	0.003	0.004	0.006
MLE		-0.004	-0.004	-0.003	-0.003	-0.003	-0.001	-0.000	0.002	0.003	0.004	0.005	0.005	0.009
PWM		-0.002	-0.002	-0.002	-0.002	-0.001	-0.000	-0.000	0.001	0.002	0.002	0.002	0.002	0.004
ENT		-0.004	-0.003	-0.003	-0.003	-0.002	-0.001	-0.000	0.002	0.003	0.004	0.004	0.005	0.009
LEA		0.007	0.006	0.005	0.004	0.004	0.003	0.000	-0.002	-0.006	-0.008	-0.012	-0.014	-0.021
MIX		-0.944	-0.943	-0.943	-0.943	-0.942	-0.941	-0.940	-0.940	-0.939	-0.939	-0.938	-0.938	-0.932
ICM		-0.003	-0.003	-0.003	-0.002	-0.002	-0.001	-0.000	0.001	0.002	0.002	0.003	0.003	0.006

TABLE 4.3 - STD. DEVIATION OF QUANTILE ESTIMATES , RANDOM SAMPLES

(F = PROB. OF NON EXCEEDENCE, X = QUANTILE)

METHOD	SAMPLE SIZE	F =	0.001	0.010	0.020	0.050	0.100	0.250	0.500	0.750	0.900	0.950	0.980	0.990	0.999
		X =	-1.933	-1.527	-1.364	-1.097	-0.834	-0.327	0.367	1.246	2.250	2.970	3.902	4.600	6.907
CMOM	5		0.968	0.804	0.741	0.656	0.585	0.493	0.529	0.777	1.191	1.505	1.919	2.211	3.223
MOM			0.875	0.731	0.682	0.614	0.553	0.484	0.530	0.752	1.122	1.406	1.780	2.057	2.960
MLE			0.722	0.618	0.586	0.539	0.501	0.471	0.530	0.716	1.017	1.252	1.562	1.797	2.533
PWM			0.927	0.770	0.711	0.631	0.566	0.483	0.526	0.767	1.167	1.470	1.869	2.150	3.123
ENT			0.738	0.627	0.594	0.544	0.504	0.471	0.533	0.728	1.041	1.285	1.605	1.848	2.609
LEA			1.036	0.854	0.783	0.682	0.602	0.493	0.536	0.821	1.282	1.627	2.072	2.389	3.506
MIX			0.493	0.371	0.333	0.296	0.297	0.395	0.620	0.971	1.387	1.685	2.059	2.327	3.264
ICM			1.573	1.400	1.334	1.229	1.130	0.961	0.805	0.798	1.059	1.339	1.746	2.068	3.151
CMOM	10		0.680	0.558	0.511	0.455	0.405	0.342	0.373	0.549	0.855	1.083	1.380	1.587	2.278
MOM			0.646	0.532	0.491	0.443	0.396	0.341	0.374	0.540	0.826	1.044	1.326	1.534	2.178
MLE			0.516	0.427	0.403	0.373	0.346	0.325	0.373	0.512	0.746	0.926	1.158	1.332	1.839
PWM			0.624	0.512	0.469	0.422	0.380	0.332	0.373	0.539	0.825	1.038	1.316	1.508	2.144
ENT			0.526	0.435	0.408	0.377	0.348	0.325	0.376	0.521	0.764	0.949	1.189	1.368	1.893
LEA			0.748	0.612	0.559	0.486	0.430	0.350	0.376	0.571	0.912	1.161	1.486	1.699	2.474
MIX			0.351	0.262	0.234	0.210	0.214	0.294	0.460	0.734	1.050	1.274	1.546	1.734	2.418
ICM			1.883	1.665	1.580	1.444	1.313	1.074	0.788	0.612	0.854	1.176	1.648	2.010	3.230
CMOM	15		0.565	0.459	0.419	0.376	0.333	0.281	0.305	0.448	0.709	0.902	1.150	1.321	1.871
MOM			0.545	0.443	0.405	0.368	0.327	0.279	0.305	0.443	0.693	0.880	1.121	1.296	1.818
MLE			0.427	0.347	0.322	0.303	0.280	0.263	0.304	0.419	0.623	0.777	0.974	1.120	1.514
PWM			0.507	0.411	0.374	0.340	0.305	0.269	0.304	0.439	0.681	0.859	1.088	1.246	1.740
ENT			0.435	0.353	0.326	0.305	0.281	0.263	0.306	0.426	0.637	0.796	0.998	1.148	1.558
LEA			0.620	0.504	0.458	0.399	0.353	0.286	0.306	0.462	0.751	0.961	1.231	1.403	2.023
MIX			0.291	0.215	0.192	0.172	0.176	0.243	0.382	0.621	0.893	1.083	1.308	1.457	2.023
ICM			1.609	1.417	1.341	1.222	1.108	0.898	0.643	0.484	0.716	1.007	1.422	1.735	2.784
CMOM	20		0.497	0.401	0.363	0.328	0.289	0.243	0.264	0.388	0.625	0.798	1.018	1.169	1.635
MOM			0.484	0.390	0.354	0.323	0.286	0.242	0.262	0.382	0.609	0.778	0.993	1.148	1.592
MLE			0.376	0.301	0.275	0.262	0.241	0.226	0.262	0.360	0.547	0.686	0.861	0.990	1.314
PWM			0.442	0.354	0.320	0.293	0.263	0.231	0.262	0.376	0.594	0.752	0.953	1.091	1.498
ENT			0.383	0.306	0.279	0.264	0.242	0.226	0.263	0.366	0.559	0.702	0.882	1.014	1.352
LEA			0.546	0.441	0.399	0.349	0.308	0.249	0.263	0.395	0.654	0.841	1.078	1.225	1.750
MIX			0.255	0.188	0.167	0.149	0.151	0.211	0.331	0.550	0.795	0.964	1.159	1.282	1.776
ICM			1.381	1.213	1.145	1.043	0.945	0.761	0.541	0.408	0.629	0.887	1.251	1.522	2.421

TABLE 4.3 (CONTINUED)

METHOD	SAMPLE SIZE	F =	0.001	0.010	0.020	0.050	0.100	0.250	0.500	0.750	0.900	0.950	0.980	0.990	0.999
		X =	-1.933	-1.527	-1.364	-1.097	-0.834	-0.327	0.367	1.246	2.250	2.970	3.902	4.600	6.907
CMOM	30		0.420	0.333	0.300	0.273	0.240	0.200	0.215	0.316	0.523	0.673	0.861	0.987	1.348
MOM			0.411	0.326	0.293	0.269	0.237	0.198	0.215	0.314	0.517	0.665	0.850	0.983	1.333
MLE			0.318	0.246	0.219	0.214	0.197	0.184	0.214	0.295	0.464	0.587	0.739	0.849	1.089
PWM			0.368	0.288	0.257	0.238	0.213	0.188	0.215	0.308	0.500	0.638	0.809	0.924	1.232
ENT			0.323	0.250	0.223	0.215	0.197	0.184	0.216	0.300	0.474	0.600	0.755	0.869	1.120
LEA			0.456	0.364	0.327	0.287	0.254	0.204	0.215	0.322	0.549	0.708	0.909	1.030	1.441
MIX			0.213	0.155	0.137	0.122	0.124	0.175	0.273	0.470	0.683	0.829	0.988	1.082	1.488
ICM			1.086	0.949	0.894	0.813	0.738	0.590	0.418	0.326	0.530	0.747	1.043	1.263	1.962
CMOM	50		0.340	0.262	0.231	0.214	0.188	0.156	0.168	0.243	0.425	0.552	0.708	0.810	1.057
MOM			0.337	0.259	0.228	0.213	0.187	0.154	0.166	0.242	0.422	0.550	0.705	0.813	1.056
MLE			0.264	0.193	0.165	0.167	0.152	0.142	0.165	0.226	0.380	0.489	0.619	0.711	0.860
PWM			0.299	0.224	0.194	0.185	0.165	0.145	0.166	0.236	0.407	0.525	0.667	0.762	0.964
ENT			0.267	0.196	0.168	0.168	0.153	0.142	0.167	0.230	0.388	0.498	0.630	0.724	0.882
LEA			0.367	0.285	0.251	0.225	0.198	0.158	0.166	0.246	0.441	0.576	0.740	0.835	1.123
MIX			0.169	0.122	0.107	0.095	0.096	0.137	0.210	0.386	0.569	0.691	0.812	0.870	1.181
ICM			0.812	0.703	0.660	0.601	0.544	0.433	0.307	0.246	0.431	0.605	0.836	1.003	1.503
CMOM	75		0.293	0.218	0.188	0.178	0.156	0.128	0.137	0.197	0.366	0.480	0.618	0.706	0.878
MOM			0.290	0.216	0.185	0.177	0.155	0.127	0.136	0.196	0.363	0.478	0.615	0.708	0.874
MLE			0.231	0.161	0.131	0.138	0.126	0.117	0.136	0.183	0.331	0.432	0.547	0.629	0.711
PWM			0.257	0.186	0.155	0.153	0.136	0.119	0.136	0.191	0.350	0.456	0.583	0.665	0.795
ENT			0.233	0.164	0.133	0.139	0.126	0.117	0.137	0.186	0.337	0.438	0.557	0.638	0.731
LEA			0.311	0.235	0.203	0.187	0.162	0.130	0.136	0.198	0.377	0.495	0.637	0.716	0.916
MIX			0.141	0.101	0.089	0.078	0.079	0.113	0.167	0.335	0.498	0.604	0.699	0.733	0.977
ICM			0.655	0.562	0.525	0.480	0.434	0.345	0.245	0.198	0.369	0.519	0.711	0.848	1.220
CMOM	100		0.264	0.191	0.161	0.156	0.136	0.111	0.118	0.168	0.332	0.439	0.566	0.646	0.768
MOM			0.263	0.190	0.158	0.155	0.135	0.111	0.118	0.170	0.331	0.439	0.566	0.650	0.773
MLE			0.212	0.141	0.111	0.122	0.110	0.102	0.118	0.158	0.304	0.401	0.510	0.584	0.631
PWM			0.234	0.161	0.131	0.134	0.119	0.104	0.118	0.165	0.319	0.421	0.539	0.614	0.700
ENT			0.214	0.143	0.113	0.122	0.110	0.102	0.119	0.161	0.309	0.406	0.518	0.592	0.646
LEA			0.278	0.204	0.173	0.162	0.141	0.113	0.118	0.170	0.342	0.452	0.582	0.654	0.801
MIX			0.123	0.088	0.077	0.068	0.068	0.099	0.143	0.306	0.459	0.557	0.636	0.654	0.863
ICM			0.568	0.483	0.450	0.414	0.373	0.295	0.210	0.170	0.336	0.472	0.644	0.765	1.064
MOM	1000		0.095	0.079	0.072	0.061	0.051	0.035	0.037	0.071	0.112	0.109	0.116	0.141	0.271
MLE			0.081	0.068	0.062	0.053	0.044	0.032	0.037	0.068	0.104	0.093	0.086	0.101	0.229
PWM			0.087	0.071	0.065	0.055	0.047	0.032	0.037	0.069	0.108	0.102	0.101	0.120	0.246
ENT			0.082	0.068	0.062	0.052	0.045	0.032	0.037	0.069	0.105	0.095	0.089	0.109	0.231
LEA			0.097	0.080	0.073	0.062	0.051	0.035	0.037	0.071	0.113	0.111	0.118	0.142	0.276
MIX			0.050	0.037	0.025	0.022	0.022	0.041	0.068	0.106	0.104	0.113	0.145	0.172	0.380
ICM			0.181	0.158	0.148	0.132	0.118	0.091	0.064	0.070	0.114	0.123	0.152	0.189	0.351

TABLE 4.4 - EFFICIENCY OF QUANTILE ESTIMATES : RANDOM SAMPLES

(F = PROB. OF NON EXCEEDENCE, X = QUANTILE)

METHOD	SAMPLE SIZE	F = 0.001 X = -1.933	0.010 -1.527	0.020 -1.364	0.050 -1.097	0.100 -0.834	0.250 -0.327	0.500 0.367	0.750 1.246	0.900 2.250	0.950 2.970	0.980 3.902	0.990 4.600	0.999 6.907
MOM	5	0.823	0.852	0.872	0.890	0.923	0.995	1.003	0.923	0.868	0.851	0.839	0.838	0.820
MLE		1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
PWM		0.783	0.818	0.855	0.893	0.929	1.026	1.017	0.895	0.818	0.799	0.785	0.793	0.766
ENT		0.992	0.999	1.001	1.002	1.004	1.003	0.989	0.976	0.971	0.971	0.971	0.971	0.972
LEA		0.617	0.656	0.694	0.757	0.815	0.985	0.972	0.764	0.663	0.637	0.623	0.625	0.591
MIX		0.535	0.448	0.420	0.371	0.325	0.266	0.274	0.362	0.483	0.549	0.615	0.657	0.698
ICM		0.174	0.160	0.159	0.156	0.158	0.189	0.371	0.827	0.887	0.786	0.683	0.630	0.522
MOM	10	0.702	0.710	0.736	0.766	0.811	0.932	0.998	0.909	0.837	0.814	0.794	0.788	0.753
MLE		1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
PWM		0.783	0.796	0.834	0.871	0.908	1.000	1.005	0.914	0.850	0.835	0.820	0.830	0.794
ENT		0.984	0.989	0.996	0.998	1.002	1.003	0.987	0.972	0.964	0.965	0.964	0.964	0.962
LEA		0.529	0.540	0.569	0.639	0.693	0.895	0.985	0.795	0.678	0.648	0.623	0.632	0.574
MIX		0.282	0.211	0.191	0.168	0.145	0.123	0.145	0.221	0.344	0.419	0.498	0.551	0.595
ICM		0.085	0.074	0.073	0.074	0.075	0.095	0.224	0.707	0.777	0.635	0.511	0.456	0.342
MOM	15	0.655	0.655	0.671	0.711	0.762	0.906	0.997	0.898	0.823	0.796	0.775	0.766	0.719
MLE		1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
PWM		0.779	0.784	0.809	0.850	0.893	0.988	1.001	0.918	0.861	0.844	0.832	0.841	0.798
ENT		0.980	0.987	0.991	0.995	1.000	1.002	0.986	0.969	0.964	0.962	0.961	0.961	0.958
LEA		0.504	0.503	0.520	0.598	0.650	0.862	0.990	0.810	0.690	0.658	0.631	0.642	0.568
MIX		0.196	0.138	0.121	0.108	0.093	0.080	0.099	0.159	0.274	0.348	0.433	0.492	0.534
ICM		0.077	0.066	0.063	0.066	0.068	0.088	0.224	0.751	0.767	0.608	0.482	0.430	0.309
MOM	20	0.634	0.624	0.633	0.681	0.734	0.888	0.997	0.892	0.816	0.789	0.766	0.759	0.700
MLE		1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
PWM		0.777	0.777	0.793	0.842	0.884	0.981	0.999	0.920	0.865	0.851	0.839	0.850	0.801
ENT		0.977	0.981	0.989	0.994	0.999	1.002	0.986	0.968	0.961	0.962	0.961	0.962	0.956
LEA		0.493	0.482	0.489	0.574	0.624	0.838	0.991	0.818	0.698	0.666	0.639	0.654	0.566
MIX		0.153	0.103	0.088	0.080	0.068	0.059	0.074	0.123	0.229	0.301	0.387	0.450	0.490
ICM		0.080	0.066	0.062	0.067	0.068	0.090	0.233	0.778	0.764	0.608	0.484	0.434	0.305
MOM	30	0.620	0.591	0.581	0.649	0.702	0.871	0.996	0.883	0.811	0.786	0.764	0.756	0.680
MLE		1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
PWM		0.785	0.769	0.765	0.833	0.875	0.975	0.997	0.921	0.875	0.862	0.851	0.862	0.803

TABLE 4.4 (CONTINUED)

METHOD	SAMPLE SIZE	F =	0.001	0.010	0.020	0.050	0.100	0.250	0.500	0.750	0.900	0.950	0.980	0.990	0.999
		X =	-1.933	-1.527	-1.364	-1.097	-0.834	-0.327	0.367	1.246	2.250	2.970	3.902	4.600	6.907
ENT			0.978	0.977	0.980	0.991	0.998	1.001	0.985	0.966	0.963	0.962	0.963	0.960	0.952
LEA			0.494	0.465	0.455	0.553	0.602	0.820	0.994	0.827	0.712	0.685	0.658	0.676	0.568
MIX			0.111	0.069	0.055	0.053	0.045	0.039	0.050	0.086	0.180	0.248	0.333	0.399	0.432
ICM			0.090	0.071	0.063	0.072	0.073	0.099	0.263	0.819	0.771	0.625	0.509	0.460	0.316
MOM	50		0.623	0.569	0.537	0.624	0.676	0.853	0.995	0.874	0.815	0.798	0.779	0.770	0.671
MLE			1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
PWM			0.803	0.771	0.749	0.829	0.869	0.970	0.995	0.920	0.883	0.879	0.873	0.883	0.810
ENT			0.984	0.977	0.973	0.990	0.996	1.001	0.984	0.966	0.964	0.967	0.969	0.967	0.956
LEA			0.520	0.462	0.432	0.542	0.591	0.807	0.994	0.838	0.742	0.720	0.698	0.721	0.582
MIX			0.077	0.042	0.031	0.032	0.027	0.023	0.030	0.053	0.131	0.194	0.275	0.343	0.361
ICM			0.109	0.078	0.065	0.079	0.080	0.109	0.290	0.843	0.781	0.659	0.554	0.508	0.333
MOM	75		0.648	0.567	0.513	0.618	0.668	0.844	0.996	0.873	0.834	0.822	0.796	0.793	0.668
MLE			1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
PWM			0.828	0.771	0.736	0.830	0.868	0.966	0.995	0.924	0.902	0.903	0.889	0.901	0.812
ENT			0.988	0.972	0.975	0.989	0.995	1.001	0.986	0.966	0.967	0.973	0.967	0.973	0.950
LEA			0.557	0.471	0.417	0.542	0.595	0.806	0.995	0.851	0.773	0.762	0.737	0.768	0.598
MIX			0.059	0.029	0.020	0.022	0.018	0.016	0.021	0.036	0.105	0.163	0.239	0.306	0.304
ICM			0.128	0.084	0.064	0.084	0.085	0.116	0.306	0.854	0.807	0.698	0.597	0.554	0.344
MOM	100		0.658	0.561	0.496	0.624	0.670	0.846	0.995	0.864	0.848	0.835	0.816	0.811	0.673
MLE			1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
PWM			0.834	0.779	0.726	0.839	0.873	0.969	0.994	0.920	0.914	0.911	0.901	0.911	0.825
ENT			0.981	0.973	0.965	0.991	0.996	1.000	0.983	0.962	0.973	0.976	0.972	0.974	0.958
LEA			0.584	0.478	0.407	0.555	0.605	0.812	0.995	0.853	0.794	0.788	0.768	0.797	0.618
MIX			0.050	0.023	0.014	0.017	0.014	0.012	0.016	0.027	0.091	0.145	0.219	0.283	0.269
ICM			0.142	0.087	0.062	0.088	0.088	0.119	0.316	0.857	0.824	0.724	0.630	0.586	0.356
MOM	1000		0.731	0.747	0.743	0.757	0.760	0.833	0.994	0.909	0.858	0.719	0.554	0.513	0.717
MLE			1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
PWM			0.879	0.907	0.894	0.928	0.916	0.966	0.990	0.951	0.924	0.835	0.731	0.707	0.868
ENT			0.988	1.010	0.997	1.008	0.998	1.000	0.980	0.971	0.978	0.947	0.948	0.857	0.984
LEA			0.709	0.722	0.720	0.730	0.751	0.821	0.995	0.903	0.846	0.701	0.531	0.504	0.688
MIX			0.007	0.005	0.004	0.003	0.002	0.001	0.002	0.005	0.012	0.010	0.008	0.011	0.052
ICM			0.202	0.186	0.173	0.159	0.143	0.122	0.333	0.932	0.837	0.573	0.325	0.286	0.426

'b' as reflected by its very high bias and standard deviation of 'a' and high standard deviation of 'b'.

Of the remaining five methods, the bias of 'a' showed the following trend:

MLE > ENT > MOM > PWM > LEAS, for n = 5, 10

MOM > MLE > ENT > PWN > LEAS, for n = 15 - 50

MOM > MLE > ENT > LEAS > PWM, for n = 75, 100

LEAS > others , for n = 1000

The standard deviation of 'a' compared as:

MLE > ENT > MOM > PWM > LEAS, for n = 5

MOM > PWM > MLE > ENT > LEAS, for n = 10

MOM > PWM = LEAS > ENT > MLE, for n = 15

MOM > LEAS > PWM > ENT > MLE, for n > 20

The efficiency of 'a' compared as:

LEAS > PWM > MOM > ENT > MLE, for n = 5

LEAS > PWM > ENT > MLE > MOM, for n = 10, 15

MLE > ENT > PWM > LEAS > MOM, for n > 20

where '>' means that the method on the left hand side of > has a bigger statistic (bias, standard deviation, or efficiency) then the method on the right hand side.

From the above trends, it appears that for rather small samples $(n \leq 15)$, LEAS is the preferred method for estimating 'a'. Table 4.1 also reveals that as n increases, the efficiency of estimating 'a' by ENT remained close to 1.00, while the efficiency from other methods reduced considerably.

In estimating 'b', PWM provided practically unbiased estimates. Analyzing the bias, standard deviation and the efficiency in much the same way as for 'a', it can be easily concluded that PWM provided superior estimates of 'b' for the entire sample range considered.

4.4.1.2 Quantile Estimates

PWM provided unbiased quantile estimates for all n and F. MOM provided estimates with lower bias than MLE and ENT. ENT resulted in slightly less bias than MLE, while LEAS produced more bias than MLE for all n except for n = 5. MIX and ICM again failed to provide satisfactory estimates of quantiles compared to other estimators because their lower bias was deteriorated by high standard deviation and vice versa, thus resulting in low efficiencies of estimates compared to MLE estimates.

The standard deviation of quantile estimates, while decreasing for increasing n, increased for higher non-exceedance probabilities, F, as expected from (4.2). MLE resulted in lowest standard deviation closely followed by ENT. MOM had slightly higher standard deviation than PWM for nearly all n and F. LEAS estimates showed higher standard deviation than MOM, although the difference reduced as n increased.

MLE estimates were most efficient, closely followed by ENT estimates. The PWM estimates proved to be more efficient than MOM estimates, though less efficient than ENT estimates.

4.4.2 Case 2: Serially Correlated Process

Tables 4.5-4.8 summarize the results for this case.

4.4.2.1 Parameter Estimates

When the samples were generated from a serially correlated process but assumed to be random for the purposes of estimation, all the estimating methods produced significantly higher bias and standard deviation than the corresponding random process estimators (case 1). However, LEAS consistently produced least bias of 'a' followed by PWM.

TABLE 4.5

PERFORMANCE STATISTICS OF PARAMETER ESTIMATES : SERIALLY CORRELATED SAMPLES

METHOD	SAMPLE SIZE	BIAS(A)	STD(A)	EFF.(A)	BIAS(B)	STD(B)	EFF.(B)
MOM	5	-0.706	1.051	1.124	-0.146	0.746	1.016
MLE		-0.783	1.090	1.000	-0.167	0.748	1.000
PWM		-0.543	0.949	1.505	-0.101	0.737	1.061
ENT		-0.765	1.068	1.043	-0.167	0.748	1.001
LEA		-0.432	0.869	1.911	-0.121	0.742	1.040
MIX		-0.706	1.051	1.124	-0.778	0.601	0.608
ICM		-8.583	176.487	0.000	-0.537	1.025	0.439
MOM	10	-0.338	0.507	0.884	-0.088	0.556	1.000
MLE		-0.329	0.468	1.000	-0.093	0.555	1.000
PWM		-0.257	0.469	1.146	-0.059	0.550	1.033
ENT		-0.325	0.469	1.006	-0.092	0.555	1.000
LEA		-0.187	0.448	1.390	-0.060	0.552	1.024
MIX		-0.338	0.507	0.884	-0.839	0.507	0.329
ICM		-0.833	2.806	0.038	-0.219	0.890	0.376
MOM	15	-0.227	0.384	0.813	-0.062	0.461	0.996
MLE		-0.210	0.343	1.000	-0.064	0.460	1.000
PWM		-0.170	0.358	1.028	-0.041	0.457	1.024
ENT		-0.208	0.347	0.990	-0.063	0.460	0.998
LEA		-0.119	0.353	1.164	0.038	0.459	1.017
MIX		-0.227	0.384	0.813	-0.867	0.442	0.228
ICM		-0.462	0.757	0.206	-0.121	0.757	0.366
MOM	20	-0.171	0.324	0.778	-0.049	0.401	0.994
MLE		-0.153	0.284	1.000	-0.049	0.400	1.000
PWM		-0.126	0.303	0.967	-0.031	0.398	1.018
ENT		-0.152	0.289	0.979	-0.048	0.400	0.997
LEA		-0.084	0.304	1.050	-0.026	0.400	1.012
MIX		-0.171	0.324	0.778	-0.883	0.396	0.173
ICM		-0.332	0.579	0.235	-0.087	0.666	0.359
MOM	30	-0.115	0.261	0.733	-0.034	0.330	0.990
MLE		-0.099	0.223	1.000	-0.033	0.329	1.000
PWM		-0.083	0.241	0.916	-0.022	0.328	1.012
ENT		-0.098	0.228	0.969	-0.033	0.329	0.997
LEA		-0.051	0.246	0.944	-0.015	0.330	1.003
MIX		-0.115	0.261	0.733	-0.900	0.334	0.118
ICM		-0.213	0.433	0.256	-0.052	0.556	0.350

TABLE 4.5 (CONTINUED)

METHOD	SAMPLE SIZE	BIAS(A)	STD(A)	EFF.(A)	BIAS(B)	STD(B)	EFF.(B)
MOM	50	-0.070	0.200	0.701	-0.022	0.259	0.986
MLE		-0.058	0.168	1.000	-0.021	0.257	1.000
PWM		-0.049	0.184	0.870	-0.014	0.256	1.007
ENT		-0.057	0.172	0.960	-0.020	0.257	0.997
LEA		-0.026	0.193	0.834	-0.007	0.258	0.995
MIX		-0.070	0.200	0.701	-0.915	0.266	0.073
ICM		-0.125	0.319	0.269	-0.030	0.435	0.349
MOM	75	-0.046	0.164	0.687	-0.014	0.211	0.983
MLE		-0.038	0.136	1.000	-0.013	0.209	1.000
PWM		-0.032	0.150	0.846	-0.009	0.209	1.005
ENT		-0.037	0.140	0.954	-0.013	0.209	0.996
LEA		-0.014	0.160	0.777	-0.003	0.211	0.988
MIX		-0.046	0.164	0.687	-0.923	0.221	0.049
ICM		-0.082	0.255	0.279	-0.019	0.354	0.349
MOM	100	-0.034	0.144	0.683	-0.011	0.182	0.982
MLE		-0.027	0.119	1.000	-0.010	0.181	1.000
PWM		-0.023	0.131	0.839	-0.007	0.180	1.004
ENT		-0.027	0.122	0.952	-0.010	0.181	0.996
LEA		-0.008	0.141	0.750	-0.002	0.182	0.986
MIX		-0.034	0.144	0.683	-0.928	0.194	0.036
ICM		-0.060	0.219	0.289	-0.014	0.306	0.349
MOM	1000	-0.003	0.056	0.764	-0.001	0.058	0.982
MLE		-0.002	0.049	1.000	-0.001	0.058	1.000
PWM		-0.002	0.052	0.886	-0.001	0.058	1.001
ENT		-0.002	0.049	0.977	-0.001	0.058	0.995
LEA		0.002	0.056	0.766	0.001	0.058	0.981
MIX		-0.003	0.056	0.766	-0.940	0.070	0.004
ICM		-0.005	0.074	0.426	-0.003	0.095	0.369

TABLE 4.6 - BIAS IN QUANTILE ESTIMATES : SERIALLY CORRELATED SAMPLES

(F = PROB. OF NON EXCEEDENCE, X = QUANTILE)

METHOD	SAMPLE SIZE	F = 0.001	0.010	0.020	0.050	0.100	0.250	0.500	0.750	0.900	0.950	0.980	0.990	0.999
		X = -1.933	-1.527	-1.364	-1.097	-0.834	-0.327	0.367	1.246	2.250	2.970	3.902	4.600	6.907
MOM	5	-0.626	-0.526	-0.486	-0.420	-0.355	-0.228	-0.055	0.163	0.419	0.602	0.836	1.010	1.584
MLE		-0.730	-0.612	-0.565	-0.487	-0.411	-0.263	-0.060	0.195	0.493	0.706	0.980	1.184	1.856
PWM		-0.433	-0.361	-0.333	-0.288	-0.244	-0.158	-0.038	0.110	0.289	0.415	0.576	0.695	1.087
ENT		-0.717	-0.602	-0.556	-0.480	-0.406	-0.261	-0.062	0.187	0.479	0.688	0.956	1.156	1.813
LEA		0.329	0.281	-0.264	-0.237	-0.209	-0.156	-0.082	0.010	0.125	0.204	0.304	0.377	0.619
MIX		-1.259	-1.159	-1.118	-1.051	-0.985	-0.859	-0.688	-0.463	-0.208	-0.027	0.205	0.379	0.952
ICM		-1.481	-1.283	-1.204	-1.073	-0.944	-0.696	-0.358	0.070	0.564	0.919	1.377	1.719	2.847
MOM	10	-0.377	-0.313	-0.289	-0.250	-0.212	-0.137	-0.033	0.097	0.254	0.364	0.504	0.608	0.949
MLE		-0.406	-0.337	-0.311	-0.269	-0.227	-0.146	-0.033	0.106	0.276	0.395	0.546	0.658	1.027
PWM		-0.254	-0.209	-0.193	-0.167	-0.141	-0.092	-0.022	0.063	0.171	0.245	0.339	0.408	0.634
ENT		-0.397	-0.330	-0.305	-0.263	-0.223	-0.144	-0.034	0.102	0.267	0.382	0.530	0.639	0.998
LEA		-0.146	-0.124	-0.116	-0.105	-0.095	-0.075	-0.045	-0.010	0.042	0.074	0.113	0.142	0.235
MIX		-1.126	-1.066	-1.042	-1.002	-0.962	-0.887	-0.785	-0.647	-0.494	-0.386	-0.247	-0.144	0.197
ICM		-0.647	-0.558	-0.522	-0.463	-0.405	-0.292	-0.139	0.055	0.285	0.448	0.656	0.811	1.320
MOM	15	-0.272	-0.224	-0.207	-0.179	-0.151	-0.098	-0.023	0.069	0.185	0.265	0.365	0.440	0.683
MLE		-0.283	-0.232	-0.215	-0.185	-0.157	-0.101	-0.022	0.074	0.195	0.278	0.383	0.461	0.716
PWM		-0.180	-0.147	-0.134	-0.116	-0.099	-0.064	-0.015	0.044	0.123	0.176	0.241	0.290	0.447
ENT		-0.275	-0.226	-0.209	-0.181	-0.153	-0.099	-0.023	0.070	0.187	0.267	0.369	0.444	0.691
LEA		-0.083	-0.070	-0.064	-0.059	-0.054	-0.045	-0.030	-0.014	0.016	0.033	0.053	0.067	0.111
MIX		-1.072	-1.029	-1.012	-0.983	-0.955	-0.901	-0.828	-0.726	-0.616	-0.538	-0.439	-0.365	-0.122
ICM		-0.395	-0.335	-0.312	-0.275	-0.239	-0.168	-0.070	0.052	0.201	0.305	0.437	0.535	0.857
MOM	20	-0.213	-0.175	-0.161	-0.139	-0.118	-0.076	-0.018	0.053	0.146	0.208	0.286	0.344	0.533
MLE		-0.218	-0.178	-0.164	-0.141	-0.120	-0.077	-0.017	0.056	0.151	0.215	0.295	0.355	0.549
PWM		-0.139	-0.113	-0.102	-0.089	-0.075	-0.049	-0.011	0.033	0.096	0.136	0.187	0.224	0.343
ENT		-0.211	-0.173	-0.159	-0.137	-0.116	-0.076	-0.018	0.053	0.144	0.206	0.283	0.340	0.527
LEA		-0.051	-0.042	-0.038	-0.036	-0.034	-0.030	-0.022	-0.016	0.004	0.013	0.022	0.029	0.047
MIX		-1.042	-1.009	-0.996	-0.974	-0.951	-0.909	-0.853	-0.772	-0.685	-0.625	-0.548	-0.490	-0.302
ICM		-0.292	-0.245	-0.228	-0.201	-0.174	-0.122	-0.049	0.041	0.155	0.232	0.330	0.403	0.641
MOM	30	-0.152	-0.123	-0.112	-0.097	-0.083	-0.054	-0.012	0.037	0.105	0.149	0.205	0.245	0.377
MLE		-0.152	-0.124	-0.112	-0.097	-0.082	-0.053	-0.011	0.039	0.107	0.152	0.208	0.249	0.382
PWM		-0.099	-0.079	-0.071	-0.061	-0.052	-0.035	-0.008	0.023	0.069	0.098	0.133	0.159	0.241

TABLE 4.6 (CONTINUED)

METHOD	SAMPLE SIZE	F = 0.001	0.010	0.020	0.050	0.100	0.250	0.500	0.750	0.900	0.950	0.980	0.990	0.999
		X = -1.933	-1.527	-1.364	-1.097	-0.834	-0.327	0.367	1.246	2.250	2.970	3.902	4.600	6.907
ENT		-0.147	-0.119	-0.108	-0.094	-0.080	-0.052	-0.012	0.036	0.101	0.144	0.198	0.237	0.364
LEA		0.023	-0.019	-0.016	-0.015	-0.016	-0.016	-0.015	-0.016	-0.005	-0.002	-0.000	-0.000	-0.001
MIX		-1.012	-0.989	-0.980	-0.964	-0.948	-0.919	-0.880	-0.820	-0.758	-0.715	-0.661	-0.621	-0.490
ICM		-0.187	-0.155	-0.143	-0.125	-0.108	-0.074	-0.026	0.031	0.108	0.160	0.224	0.271	0.424
MOM	50	-0.100	-0.080	-0.072	-0.062	-0.053	-0.035	-0.008	0.024	0.071	0.101	0.136	0.163	0.247
MLE		-0.098	-0.078	-0.070	-0.060	-0.051	-0.033	-0.007	0.024	0.071	0.100	0.135	0.160	0.243
PWM		-0.065	-0.051	-0.045	-0.039	-0.033	-0.022	-0.005	0.014	0.048	0.067	0.090	0.106	0.158
ENT		-0.094	-0.075	-0.067	-0.058	-0.050	-0.033	-0.007	0.022	0.067	0.095	0.128	0.153	0.231
LEA		-0.007	-0.004	-0.002	-0.003	-0.004	-0.007	-0.008	-0.013	-0.007	-0.007	-0.010	-0.012	-0.023
MIX		-0.987	-0.973	-0.967	-0.957	-0.946	-0.927	-0.903	-0.860	-0.819	-0.792	-0.757	-0.731	-0.648
ICM		-0.115	-0.094	-0.085	-0.074	-0.064	-0.044	-0.014	0.020	0.071	0.103	0.143	0.172	0.264
MOM	75	-0.071	-0.056	-0.049	-0.042	-0.036	-0.024	-0.004	0.017	0.053	0.074	0.100	0.118	0.176
MLE		-0.069	-0.054	-0.047	-0.041	-0.035	-0.023	-0.003	0.017	0.052	0.073	0.098	0.116	0.172
PWM		-0.047	-0.036	-0.031	-0.026	-0.022	-0.015	-0.002	0.011	0.037	0.051	0.068	0.079	0.114
ENT		-0.066	-0.052	-0.045	-0.039	-0.033	-0.022	-0.004	0.016	0.050	0.070	0.093	0.110	0.163
LEA		-0.000	0.002	0.004	0.003	0.001	-0.002	-0.004	-0.010	-0.005	-0.007	-0.010	-0.014	-0.028
MIX		-0.973	-0.963	-0.959	-0.952	-0.944	-0.931	-0.915	-0.882	-0.853	-0.833	-0.809	-0.791	-0.734
ICM		-0.078	-0.062	-0.055	-0.048	-0.042	-0.028	-0.008	0.015	0.052	0.075	0.101	0.121	0.182
MOM	100	-0.057	-0.044	-0.038	-0.033	-0.028	-0.019	-0.003	0.013	0.043	0.060	0.079	0.093	0.137
MLE		-0.055	-0.042	-0.036	-0.031	-0.027	-0.018	-0.003	0.013	0.042	0.059	0.078	0.092	0.134
PWM		-0.038	-0.029	-0.024	-0.020	-0.018	-0.012	-0.002	0.008	0.030	0.042	0.055	0.064	0.090
ENT		-0.053	-0.040	-0.035	-0.030	-0.026	-0.017	-0.003	0.012	0.040	0.056	0.074	0.087	0.127
LEA		0.002	0.004	0.006	0.005	0.003	-0.001	-0.003	-0.009	-0.005	-0.006	-0.010	-0.014	-0.029
MIX		-0.966	-0.959	-0.956	-0.950	-0.944	-0.934	-0.921	-0.894	-0.871	-0.856	-0.837	-0.824	-0.781
ICM		-0.060	-0.047	-0.041	-0.036	-0.031	-0.021	-0.006	0.011	0.042	0.059	0.079	0.094	0.139
MOM	1000	-0.006	-0.006	-0.005	-0.005	-0.004	-0.002	-0.000	0.002	0.005	0.006	0.008	0.009	0.015
MLE		-0.006	-0.006	-0.005	-0.005	-0.004	-0.002	-0.001	0.002	0.005	0.006	0.008	0.009	0.015
PWM		-0.005	-0.004	-0.004	-0.004	-0.003	-0.001	-0.000	0.002	0.004	0.004	0.006	0.007	0.011
ENT		-0.006	-0.006	-0.005	-0.005	-0.004	-0.002	-0.000	0.002	0.004	0.006	0.008	0.009	0.015
LEA		0.004	0.003	0.003	0.002	0.002	0.001	-0.000	-0.001	-0.003	-0.005	-0.017	-0.019	-0.013
MIX		-0.946	-0.945	-0.944	-0.943	-0.943	-0.941	-0.939	-0.937	-0.935	-0.934	-0.932	-0.931	-0.921
ICM		-0.008	-0.007	-0.007	-0.006	-0.005	-0.003	-0.002	0.002	0.004	0.006	0.008	0.009	0.016

TABLE 4.7 - STD. DEVIATION OF QUANTILE ESTIMATES ; SERIALLY CORRELATED SAMPLES

(F = PROB. OF NON EXCEEDENCE, X = QUANTILE)

METHOD	SAMPLE SIZE	F = 0.001 X = -1.933	0.010 -1.527	0.020 -1.364	0.050 -1.097	0.100 -0.834	0.250 -0.327	0.500 0.367	0.750 1.246	0.900 2.250	0.950 2.970	0.980 3.902	0.990 4.600	0.999 6.907
MOM	5	0.833	0.763	0.743	0.715	0.700	0.712	0.807	1.010	1.321	1.566	1.896	2.147	2.958
MLE		0.757	0.712	0.698	0.684	0.682	0.709	0.808	0.997	1.276	1.496	1.791	2.017	2.743
PWM		0.882	0.786	0.758	0.724	0.701	0.703	0.802	1.028	1.377	1.649	2.012	2.289	3.179
ENT		0.763	0.715	0.701	0.686	0.682	0.709	0.808	0.999	1.281	1.504	1.803	2.031	2.766
LEA		0.931	0.815	0.781	0.739	0.709	0.706	0.814	1.065	1.453	1.751	2.148	2.438	3.421
MIX		0.467	0.399	0.385	0.385	0.409	0.512	0.710	1.012	1.385	1.654	2.004	2.257	3.095
ICM		1.410	1.297	1.255	1.192	1.137	1.056	1.011	1.069	1.281	1.495	1.811	2.066	2.957
MOM	10	0.665	0.585	0.566	0.538	0.521	0.526	0.610	0.794	1.077	1.296	1.586	1.806	2.492
MLE		0.580	0.524	0.515	0.500	0.496	0.519	0.610	0.779	1.033	1.228	1.486	1.681	2.282
PWM		0.669	0.579	0.557	0.531	0.513	0.518	0.608	0.800	1.096	1.323	1.622	1.848	2.550
ENT		0.589	0.529	0.518	0.503	0.497	0.519	0.611	0.785	1.045	1.245	1.509	1.709	2.324
LEA		0.736	0.629	0.595	0.559	0.531	0.523	0.613	0.824	1.154	1.403	1.733	1.967	2.760
MIX		0.370	0.311	0.301	0.304	0.330	0.426	0.602	0.874	1.201	1.434	1.736	1.937	2.654
ICM		1.506	1.348	1.288	1.194	1.108	0.962	0.837	0.847	1.098	1.358	1.738	2.039	3.042
MOM	15	0.584	0.501	0.482	0.456	0.437	0.436	0.509	0.673	0.933	1.133	1.394	1.591	2.191
MLE		0.497	0.436	0.427	0.416	0.411	0.429	0.508	0.656	0.885	1.060	1.289	1.461	1.969
PWM		0.571	0.487	0.464	0.443	0.426	0.429	0.508	0.676	0.939	1.139	1.401	1.597	2.192
ENT		0.506	0.441	0.431	0.418	0.412	0.429	0.510	0.664	0.899	1.080	1.315	1.492	2.018
LEA		0.641	0.540	0.506	0.474	0.447	0.435	0.511	0.693	0.987	1.208	1.498	1.701	2.380
MIX		0.321	0.266	0.256	0.259	0.282	0.369	0.526	0.775	1.070	1.280	1.548	1.719	2.354
ICM		1.384	1.225	1.165	1.072	0.986	0.835	0.698	0.700	0.954	1.212	1.581	1.870	2.816
MOM	20	0.530	0.448	0.428	0.404	0.385	0.380	0.444	0.593	0.835	1.020	1.260	1.441	1.976
MLE		0.444	0.383	0.372	0.364	0.358	0.373	0.443	0.576	0.787	0.947	1.155	1.311	1.755
PWM		0.509	0.429	0.405	0.388	0.372	0.374	0.444	0.593	0.834	1.016	1.253	1.428	1.948
ENT		0.453	0.388	0.376	0.366	0.359	0.373	0.446	0.583	0.801	0.967	1.181	1.341	1.803
LEA		0.578	0.483	0.449	0.421	0.394	0.380	0.445	0.608	0.877	1.079	1.341	1.523	2.123
MIX		0.289	0.237	0.227	0.229	0.250	0.329	0.470	0.703	0.975	1.167	1.408	1.559	2.133
ICM		1.257	1.106	1.051	0.964	0.884	0.740	0.609	0.610	0.855	1.098	1.443	1.711	2.575
MOM	30	0.462	0.384	0.361	0.342	0.323	0.314	0.366	0.492	0.708	0.872	1.084	1.241	1.684
MLE		0.379	0.319	0.305	0.301	0.295	0.307	0.365	0.476	0.663	0.803	0.983	1.117	1.470
PWM		0.432	0.358	0.334	0.322	0.307	0.308	0.366	0.490	0.701	0.859	1.062	1.211	1.630

TABLE 4.7 (CONTINUED)

METHOD	SAMPLE SIZE	F =	0.001	0.010	0.020	0.050	0.100	0.250	0.500	0.750	0.900	0.950	0.980	0.990	0.999
		X =	-1.933	-1.527	-1.364	-1.097	-0.834	-0.327	0.367	1.246	2.250	2.970	3.902	4.600	6.907
ENT			0.387	0.324	0.308	0.303	0.296	0.307	0.367	0.482	0.675	0.821	1.007	1.144	1.514
LEA			0.500	0.412	0.381	0.356	0.331	0.315	0.366	0.501	0.738	0.913	1.140	1.293	1.787
MIX			0.248	0.201	0.191	0.191	0.208	0.276	0.396	0.605	0.845	1.014	1.216	1.339	1.827
ICM			1.088	0.952	0.901	0.825	0.754	0.624	0.503	0.499	0.725	0.944	1.249	1.485	2.225
MOM	50		0.384	0.311	0.286	0.275	0.257	0.247	0.286	0.387	0.575	0.716	0.894	1.025	1.359
MLE			0.312	0.252	0.233	0.237	0.231	0.239	0.285	0.372	0.536	0.657	0.809	0.920	1.172
PWM			0.352	0.283	0.258	0.254	0.241	0.240	0.286	0.383	0.565	0.699	0.868	0.988	1.293
ENT			0.317	0.256	0.236	0.238	0.231	0.239	0.287	0.378	0.546	0.671	0.827	0.941	1.206
LEA			0.410	0.331	0.302	0.284	0.263	0.248	0.286	0.391	0.593	0.741	0.928	1.051	1.422
MIX			0.201	0.161	0.152	0.151	0.164	0.219	0.313	0.498	0.703	0.844	1.004	1.091	1.480
ICM			0.871	0.757	0.714	0.655	0.597	0.491	0.392	0.389	0.587	0.771	1.022	1.213	1.787
MOM	75		0.333	0.262	0.236	0.229	0.213	0.202	0.233	0.316	0.488	0.614	0.771	0.884	1.142
MLE			0.271	0.209	0.187	0.195	0.189	0.195	0.233	0.304	0.455	0.564	0.698	0.796	0.976
PWM			0.302	0.235	0.210	0.209	0.198	0.196	0.233	0.313	0.478	0.597	0.745	0.849	1.076
ENT			0.275	0.212	0.190	0.196	0.190	0.195	0.234	0.308	0.463	0.576	0.713	0.811	1.006
LEA			0.352	0.278	0.250	0.237	0.218	0.203	0.233	0.319	0.501	0.632	0.794	0.900	1.183
MIX			0.170	0.134	0.126	0.124	0.134	0.181	0.257	0.427	0.610	0.733	0.864	0.927	1.248
ICM			0.721	0.623	0.584	0.538	0.488	0.400	0.318	0.316	0.497	0.658	0.872	1.032	1.489
MOM	100		0.301	0.232	0.205	0.202	0.186	0.175	0.202	0.274	0.437	0.554	0.697	0.801	1.007
MLE			0.246	0.184	0.161	0.170	0.164	0.169	0.202	0.263	0.407	0.511	0.635	0.723	0.859
PWM			0.273	0.206	0.181	0.183	0.172	0.170	0.202	0.271	0.427	0.538	0.673	0.768	0.944
ENT			0.250	0.186	0.163	0.171	0.165	0.169	0.203	0.267	0.414	0.520	0.647	0.737	0.884
LEA			0.316	0.244	0.216	0.208	0.190	0.176	0.202	0.276	0.446	0.568	0.716	0.809	1.038
MIX			0.151	0.117	0.110	0.108	0.117	0.157	0.222	0.385	0.555	0.667	0.780	0.826	1.102
ICM			0.630	0.540	0.505	0.467	0.423	0.346	0.275	0.273	0.444	0.591	0.782	0.924	1.304
MOM	1000		0.109	0.092	0.085	0.074	0.066	0.056	0.065	0.100	0.147	0.158	0.181	0.210	0.348
MLE			0.091	0.077	0.072	0.064	0.059	0.053	0.065	0.097	0.137	0.140	0.152	0.175	0.300
PWM			0.098	0.083	0.077	0.068	0.061	0.054	0.065	0.099	0.143	0.149	0.168	0.193	0.322
ENT			0.092	0.078	0.072	0.064	0.059	0.053	0.065	0.098	0.139	0.144	0.156	0.180	0.305
LEA			0.110	0.092	0.085	0.075	0.066	0.056	0.065	0.100	0.148	0.159	0.183	0.212	0.352
MIX			0.058	0.045	0.036	0.035	0.038	0.058	0.088	0.132	0.147	0.169	0.209	0.243	0.446
ICM			0.203	0.177	0.167	0.150	0.134	0.107	0.086	0.099	0.149	0.169	0.212	0.254	0.434

TABLE 4.8 - EFFICIENCY OF QUANTILE ESTIMATES : SERIALLY CORRELATED SAMPLES

(F = PROB. OF NON EXCEEDENCE, X = QUANTILE)

METHOD	SAMPLE SIZE	F = 0.001 X = -1.933	0.010 -1.527	0.020 -1.364	0.050 -1.097	0.100 -0.834	0.250 -0.327	0.500 0.367	0.750 1.246	0.900 2.250	0.950 2.970	0.980 3.902	0.990 4.600	0.999 6.907
MOM	5	1.019	1.027	1.024	1.027	1.028	1.023	1.006	0.986	0.974	0.972	0.972	0.971	0.974
MLE		1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
PWM		1.145	1.178	1.176	1.163	1.151	1.102	1.019	0.965	0.945	0.947	0.952	0.956	0.972
ENT		1.008	1.008	1.007	1.006	1.005	1.003	1.000	1.000	1.000	1.000	1.001	1.001	1.003
LEA		1.134	1.185	1.186	1.172	1.159	1.094	0.983	0.909	0.880	0.881	0.886	0.898	0.908
MIX		0.613	0.587	0.576	0.563	0.557	0.572	0.673	0.832	0.953	1.000	1.027	1.044	1.046
ICM		0.265	0.265	0.267	0.274	0.290	0.358	0.572	0.899	0.955	0.889	0.806	0.757	0.651
MOM	10	0.857	0.881	0.896	0.916	0.941	0.986	0.999	0.967	0.933	0.918	0.905	0.898	0.881
MLE		1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
PWM		0.978	1.024	1.040	1.040	1.052	1.050	1.008	0.959	0.928	0.919	0.913	0.910	0.907
ENT		0.993	0.999	1.000	1.002	1.003	1.003	0.996	0.988	0.982	0.981	0.980	0.979	0.979
LEA		0.889	0.946	0.985	0.997	1.024	1.042	0.988	0.912	0.857	0.843	0.831	0.838	0.816
MIX		0.357	0.315	0.307	0.294	0.288	0.300	0.381	0.523	0.678	0.755	0.815	0.864	0.884
ICM		0.187	0.183	0.187	0.197	0.214	0.288	0.519	0.858	0.888	0.815	0.727	0.677	0.569
MOM	15	0.788	0.811	0.832	0.865	0.904	0.973	0.996	0.952	0.907	0.887	0.870	0.861	0.834
MLE		1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
PWM		0.914	0.943	0.979	0.991	1.012	1.032	1.003	0.952	0.916	0.904	0.894	0.891	0.877
ENT		0.984	0.992	0.996	0.999	1.002	1.002	0.992	0.980	0.973	0.970	0.969	0.968	0.965
LEA		0.784	0.822	0.881	0.908	0.954	1.016	0.990	0.908	0.843	0.822	0.804	0.809	0.773
MIX		0.261	0.216	0.210	0.201	0.195	0.205	0.269	0.387	0.539	0.623	0.698	0.760	0.790
ICM		0.158	0.151	0.157	0.170	0.188	0.268	0.526	0.886	0.864	0.768	0.671	0.620	0.507
MOM	20	0.750	0.769	0.792	0.833	0.880	0.965	0.995	0.944	0.894	0.871	0.851	0.840	0.807
MLE		1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
PWM		0.881	0.905	0.945	0.963	0.991	1.022	1.000	0.949	0.912	0.898	0.887	0.883	0.864
ENT		0.980	0.988	0.993	0.997	1.001	1.002	0.991	0.977	0.969	0.966	0.965	0.963	0.958
LEA		0.728	0.757	0.813	0.853	0.910	0.998	0.991	0.906	0.834	0.811	0.790	0.795	0.749
MIX		0.209	0.166	0.158	0.152	0.147	0.155	0.207	0.307	0.452	0.538	0.623	0.691	0.728
ICM		0.147	0.139	0.143	0.157	0.175	0.258	0.527	0.897	0.851	0.749	0.649	0.597	0.480
MOM	30	0.704	0.717	0.737	0.791	0.846	0.952	0.995	0.934	0.879	0.853	0.830	0.819	0.775
MLE		1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
PWM		0.851	0.868	0.906	0.932	0.966	1.010	0.998	0.946	0.909	0.894	0.881	0.879	0.849

TABLE 4.8 (CONTINUED)

METHOD	SAMPLE SIZE	F = 0.001 X = -1.933	0.010 -1.527	0.020 -1.364	0.050 -1.097	0.100 -0.834	0.250 -0.327	0.500 0.367	0.750 1.246	0.900 2.250	0.950 2.970	0.980 3.902	0.990 4.600	0.999 6.907
ENT		0.973	0.982	0.987	0.994	1.000	1.001	0.989	0.974	0.966	0.962	0.959	0.960	0.951
LEA		0.667	0.686	0.727	0.788	0.854	0.972	0.992	0.905	0.828	0.801	0.778	0.784	0.723
MIX		0.154	0.115	0.106	0.104	0.100	0.105	0.143	0.219	0.350	0.434	0.527	0.601	0.644
ICM		0.137	0.126	0.127	0.144	0.162	0.245	0.526	0.911	0.840	0.729	0.627	0.575	0.450
MOM	50	0.678	0.675	0.679	0.752	0.814	0.939	0.995	0.928	0.871	0.846	0.823	0.809	0.751
MLE		1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
PWM		0.835	0.840	0.862	0.906	0.944	0.999	0.996	0.946	0.910	0.897	0.884	0.883	0.845
ENT		0.975	0.979	0.981	0.992	0.998	1.001	0.989	0.973	0.966	0.963	0.960	0.960	0.951
LEA		0.635	0.635	0.650	0.736	0.807	0.949	0.994	0.909	0.831	0.805	0.780	0.789	0.709
MIX		0.105	0.072	0.062	0.064	0.061	0.064	0.089	0.141	0.251	0.330	0.425	0.505	0.549
ICM		0.138	0.120	0.115	0.137	0.155	0.240	0.529	0.920	0.837	0.731	0.631	0.580	0.439
MOM	75	0.674	0.652	0.643	0.728	0.793	0.930	0.995	0.923	0.869	0.847	0.824	0.812	0.736
MLE		1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
PWM		0.836	0.830	0.828	0.893	0.933	0.994	0.995	0.943	0.912	0.903	0.889	0.890	0.840
ENT		0.976	0.979	0.978	0.990	0.997	1.000	0.989	0.971	0.968	0.962	0.962	0.964	0.947
LEA		0.629	0.607	0.599	0.707	0.779	0.933	0.994	0.909	0.834	0.812	0.788	0.797	0.701
MIX		0.080	0.050	0.040	0.043	0.041	0.043	0.060	0.096	0.191	0.263	0.355	0.435	0.469
ICM		0.148	0.120	0.109	0.136	0.154	0.239	0.535	0.924	0.838	0.738	0.645	0.598	0.437
MOM	100	0.676	0.640	0.623	0.718	0.783	0.926	0.994	0.917	0.869	0.852	0.832	0.817	0.732
MLE		1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
PWM		0.836	0.819	0.816	0.888	0.928	0.993	0.994	0.943	0.914	0.907	0.897	0.895	0.841
ENT		0.972	0.980	0.978	0.989	0.996	1.000	0.987	0.971	0.965	0.968	0.965	0.964	0.947
LEA		0.636	0.596	0.578	0.695	0.766	0.927	0.994	0.909	0.840	0.821	0.799	0.812	0.701
MIX		0.066	0.038	0.029	0.033	0.031	0.032	0.045	0.073	0.157	0.225	0.313	0.391	0.414
ICM		0.158	0.121	0.106	0.137	0.154	0.239	0.539	0.926	0.842	0.751	0.662	0.616	0.440
MOM	1000	0.703	0.710	0.731	0.752	0.800	0.915	0.995	0.932	0.866	0.787	0.701	0.695	0.741
MLE		1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
PWM		0.874	0.870	0.897	0.896	0.938	0.986	0.994	0.952	0.923	0.880	0.819	0.826	0.864
ENT		0.975	0.975	1.001	0.992	1.002	1.000	0.987	0.975	0.972	0.951	0.950	0.943	0.963
LEA		0.687	0.702	0.721	0.740	0.796	0.911	0.995	0.925	0.861	0.779	0.688	0.683	0.727
MIX		0.009	0.007	0.006	0.005	0.004	0.003	0.005	0.010	0.021	0.022	0.025	0.033	0.086
ICM		0.202	0.190	0.189	0.184	0.194	0.247	0.567	0.949	0.850	0.688	0.513	0.474	0.478

From sample size 10 onwards, MLE, followed closely by ENT, gave least standard deviation of 'a'. However, up to sample size n = 20, LEAS resulting in comparable standard deviation produced estimates of 'a' with a higher efficiency than MLE. Hence, LEAS can be preferred for such sample sizes.

For 'b', PWM was without doubt the superior method resulting in less bias, least standard deviation and higher efficiency estimates. Although LEAS resulted in lower bias than PWM for n > 15, it showed less efficiency than PWM owing to its higher standard deviation. But it is significant to note that the effect of serial correlation was to markedly lower the performance deviation among the first five methods. In fact the first five methods performed to within 98 percent of the efficiency of MLE method for estimating 'b'.

4.4.2.2 Quantile Estimates

The bias in quantile estimates also increased for serially correlated samples as compared to purely random samples. LEAS provided least biased estimates of the first five methods for sample sizes 5 to 100. PWM provided the next lowest bias. MLE and ENT continued to provide very close-biased estimates, although ENT produced slightly lower bias. MOM provided lower bias than MLE for up to about n = 30, beyond which MOM produced slightly higher bias than MLE. Quite in contrast with other methods, the absolute bias resulting from MIX increased with n for any given F.

MLE resulted in least standard deviation of quantile estimates among the first five methods, even for n = 5, closely followed by ENT. MOM gave lower standard deviation than PWM for n = 5 and 10 only, after which mostly PWM produced lower standard deviation estimates. LEAS provided estimates with rather high standard deviation and this fact is amply demonstrated in Table 4.8 in terms of efficiencies.

Except for n = 5, and some quantiles less than 0.5 for other n, MLE turned out to be the most efficient method for serially correlated samples, followed closely again by ENT. PWM provided the next higher efficiency estimates mostly at all quantile values for all n except at n = 5. MOM came next and LEAS provided the least efficient estimates.

4.5 Bias Correction in Quantile Estimates of MOM

Owing to its simplicity and ease of calculations, MOM has been widely used as an estimator of EV1 distribution parameters. However, MOM results in biased estimates as shown previously. The bias resulting from MOM-quantile estimator can be corrected using simulation results as follows:

From (4.2) and (4.13) we have,

$$x - \hat{x} = b - \hat{b} - \ln(-\ln F) \cdot \left[\frac{1}{a} - \frac{1}{\hat{a}}\right] \tag{4.14}$$

$$E(x - \hat{x}) = E(b - \hat{b}) - \ln(-\ln F) \cdot E\left[\frac{1}{a} - \frac{1}{\hat{a}}\right] \tag{4.15}$$

But from (4.4)

$$E(b - \hat{b}) = -\,0.57721 \cdot E\left[\frac{1}{a} - \frac{1}{\hat{a}}\right] \tag{4.16}$$

Substituting in (4.15)

$$E(x - \hat{x}) = -\frac{1}{a} \cdot E\left[1 - \frac{a}{\hat{a}}\right] \cdot \{(0.57721 + \ln(-\ln F))\} \tag{4.17}$$

$E[1 - (a/\hat{a})]$ is the bias of the scaled random variable $a/\hat{a}$. It is a dimensionless quantity.

To investigate the bias of $a/\hat{a}$ as a function of the sample size and the distribution parameters, three sets of sampling experiments were carried out using N = 25,000 Monte Carlo samples of sizes n (= 5, 7, 10, 15, 20, 30, 40, 50, 75, 100, 150 and 200). The random samples were generated respectively from the following populations:

(1) $a = 1.00$, $b = 0.0$

(2) $a = 0.05$, $b = 100.0$

(3) $a = 0.01$, $b = 200.0$

The bias $E[1 - (a/\hat{a})]$ was computed for various values of n in each parameter set. The results were plotted on a log-log plot and are shown in Fig. 4.1. It is apparent from this figure that the bias of $a/\hat{a}$ is independent of the population parameters from which it was computed and depends only on the sample size n. The regression line is closely fitted by the equation:

$$E(1 - \frac{a}{\hat{a}}) = \frac{0.35}{n^{0.8589}} \cong f(n) \tag{4.18}$$

where f(n) denoting the "true correction" is used in subsequent discussion.

Using (4.18) in (4.17),

$$E(x - \hat{x}) = -\frac{1}{a} \cdot f(n) \cdot [0.57721 + \ln(-\ln F)] \tag{4.19}$$

From (4.18) we can write

$$E[\frac{1}{\hat{a}(1 - f(n))}) = \frac{1}{a} \tag{4.20}$$

which implies that $1/[\hat{a}(1 - f(n))]$ is an unbiased estimator of 1/a. Substituting (4.20) in (4.19) and simplifying, we get

$$E[x - \{\hat{x} - f(n) \cdot [0.57221 + \ln(-\ln F)] \cdot \frac{1}{\hat{a}(1 - f(n))}\}] = 0 \tag{4.21}$$

Hence by definition,

$$\hat{\hat{x}} = \hat{x} - f(n) \cdot [0.57721 + \ln(-\ln F)] \cdot \frac{1}{\hat{a}(1 - f(n))} \tag{4.22}$$

is an unbiased (or corrected) estimator of x. It is henceforth denoted as CMOM estimator of quantile. Simplifying (4.22) further,

TABLE 4.9 - EFFICIENCY OF 'CMOM' QUANTILE ESTIMATES RELATIVE TO 'MOM' ESTIMATES

(F = PROB. OF NON EXCEEDENCE, X = QUANTILE)

METHOD	SAMPLE SIZE	F =	0.001	0.010	0.020	0.050	0.100	0.250	0.500	0.750	0.900	0.950	0.980	0.990	0.999
		X =	-1.933	-1.527	-1.364	-1.097	-0.834	-0.327	0.367	1.246	2.250	2.970	3.902	4.600	6.907
CMOM	5		0.879	0.888	0.908	0.937	0.948	1.004	1.015	0.955	0.914	0.903	0.895	0.902	0.883
MOM			1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
CMOM	10		0.929	0.932	0.944	0.968	0.972	1.004	1.008	0.976	0.952	0.947	0.943	0.955	0.936
MOM			1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
CMOM	15		0.948	0.948	0.955	0.975	0.980	1.003	1.006	0.985	0.966	0.964	0.961	0.974	0.957
MOM			1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
CMOM	20		0.959	0.957	0.958	0.976	0.983	1.003	1.005	0.988	0.973	0.971	0.969	0.983	0.966
MOM			1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
CMOM	30		0.970	0.966	0.964	0.982	0.987	1.002	1.003	0.992	0.978	0.979	0.978	0.992	0.977
MOM			1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
CMOM	50		0.982	0.973	0.970	0.989	0.991	1.003	1.002	0.999	0.986	0.990	0.987	1.001	0.987
MOM			1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
CMOM	75		0.985	0.981	0.977	0.994	0.994	1.003	1.001	0.999	0.989	0.993	0.993	1.008	0.993
MOM			1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
CMOM	100		0.994	0.979	0.975	0.993	0.996	1.004	1.001	1.003	0.991	0.992	0.993	1.004	0.999
MOM			1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
CMOM	200		0.992	0.995	0.992	0.997	1.002	1.000	1.001	1.000	0.995	0.998	0.993	0.994	0.991
MOM			1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

$$\hat{\hat{x}} = \hat{b} - \frac{1}{\hat{a}} \cdot [\ln(-\ln F) + \{0.57721 + \ln(-\ln F\} \cdot \frac{f(n)}{1 - f(n)}] \quad (4.23)$$

and understandably enough the bias corrected quantile estimator is a function of not only 'F' but 'n' also.

The bias and standard deviation of (4.23), with f(n) substituted by the expression in (4.18), are shown in Tables 4.2 and 4.3 respectively, against the method CMOM.

It is easy to see from (4.22) that the variance of $\hat{\hat{x}}$ will be larger than the variance of $\hat{x}$. This is also corroborated by the results in Table 3. The relative efficiency of CMOM quantile estimator (4.23) as compared to the MOM estimator is shown in Table 4.9. Fig. 4.2 shows the bias of original and corrected quantile estimators for 99.9 percent non-exceedance probability. The results are typical of other probabilities too.

4.6 Conclusions

Seven available estimators of EV1 distribution parameters and quantiles were statistically compared using Monte Carlo sampling experiments performed on two cases: a purely random process and a serially correlated process. Additionally, a bias corrected MOM estimator of quantile was also developed for purely random process. The corrected estimator resulted in practically unbiased quantile estimates even for very small sample sizes without causing any appreciable deterioration in the mean square error (MSE).

With regard to the inter-comparison of parameter and quantile estimators, some of the important conclusions are as follows:

(1) The methods of mixed moments and incomplete means resulted in poor estimation of the parameters and the quantiles.

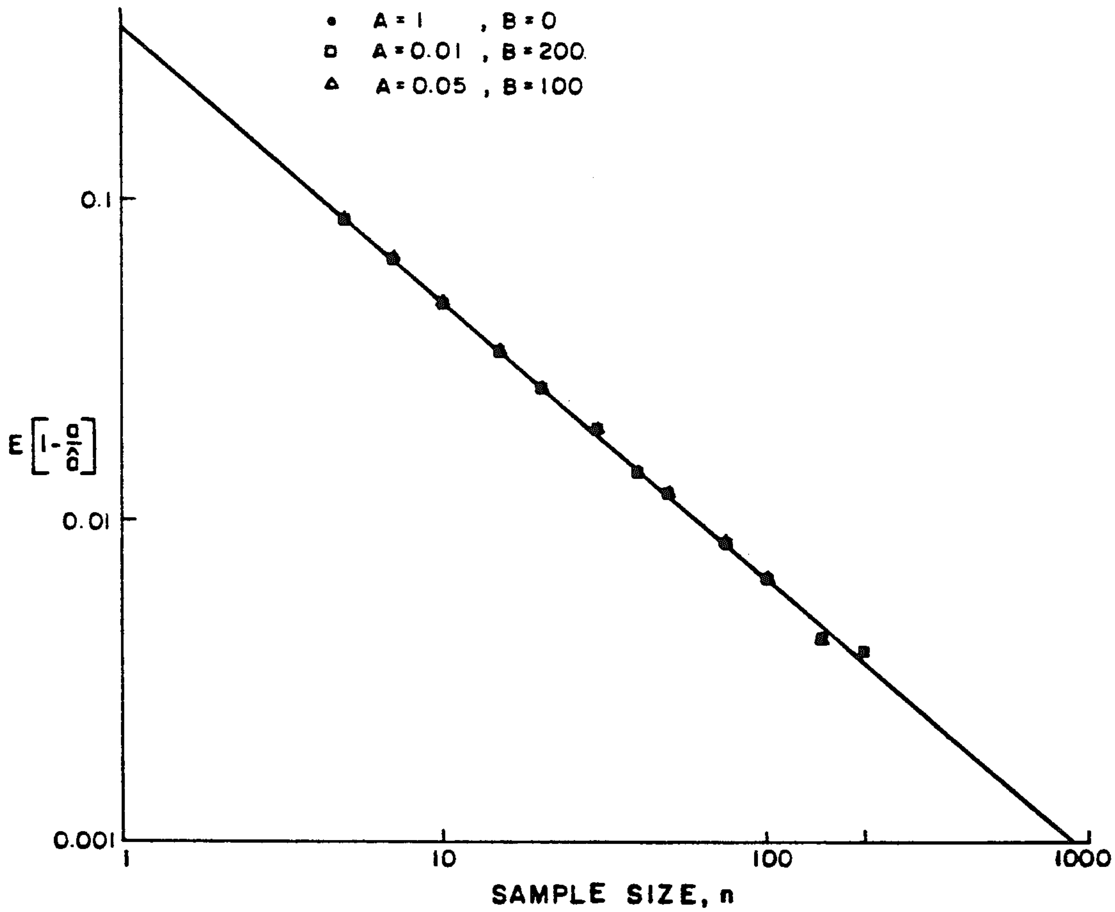

Fig 4.1 : Bias, E(1 - (a/â)) versus sample size n for the EV1 Model

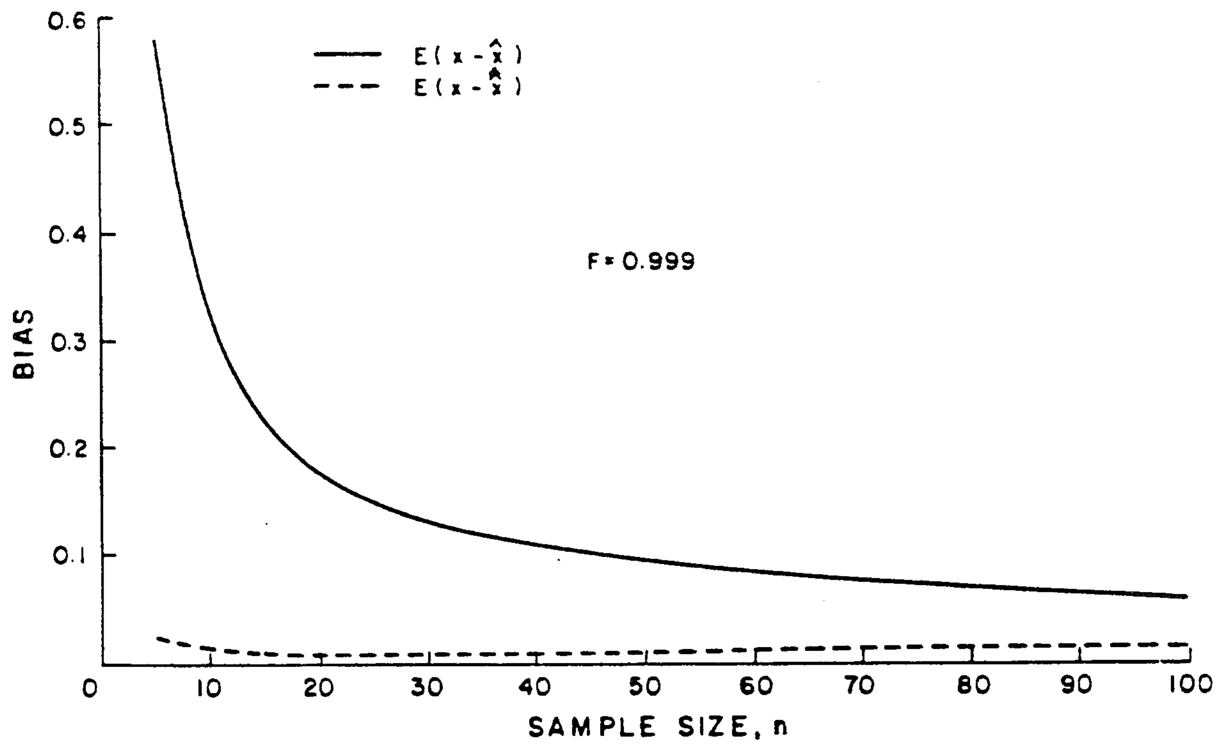

Fig 4.2 : Bias in 'F = 0.999' MOM and CMOM Quantile Estimates of the EV1 Model

(2) The method of least squares provided minimum bias and maximum efficiency estimate of the parameter 'a' for very small samples and also provided competitive estimates of the parameter 'b'.

(3) The maximum likelihood estimation method generally provided most efficient quantile estimates followed closely by the entropy method. In fact, ENT method performed practically in the same manner as MLE.

(4) For small samples, the method of probability weighted moments and the method of moments performed comparably in efficiency of estimating the quantiles. However, the efficiency of PWM improved relative to MLE with increasing sample size. PWM also resulted in nearly unbiased quantile estimates.

(5) The incorporation of serial correlation in samples resulted in deterioration of the performance of all estimators. However, all the methods performed much more similarly in this case.

Chapter 5

LOG PEARSON TYPE 3 (LP3) DISTRIBUTION

5.1 Introduction

The objective of flood frequency analysis is to obtain an estimate of T-year flood quantile at one or more locations in a river system. The length of the available record 'n', if any, is typically less than the recurrence interval T. The quantile estimates are subject to variability on account of both model error and sampling error. Competing estimators can yield estimates markedly differing in bias and mean square error, the commonly used yardsticks to assess statistical performance. Several studies have been reported in the literature comparing the performance of "at-site," "at-site/regional," "regional only" estimators or combinations thereof. In a site-specific framework, a parent distribution representative of real flood experience (Landwehr, et al., 1980) is selected and several competing estimators based on specified model distributions are used to estimate quantiles from the samples drawn from the parent distribution. Landwehr, et al. (1980) used six Wakeby distribution parents as the basis of testing the following estimators: extreme value type 1 (EV1) distribution with method of moments (MOM), maximum likelihood estimation (MLE) and probability weighted moments (PWM); and log normal 3 (LN3) with MOM and Wakeby (WAK) distribution with PWM. Kuczera (1982b) considered the performance of normal distribution with MLE, LN2 with MLE and robust MOM, LP3 with MOM applied to log-transformed data (referred to as indirect method of moment (MMI), EV1 with MLE, PWM and MOM, Log-Gumbel with MOM, WAK with PWM on four Wakeby parent distributions. In a more recent study, Wallis and Wood (1985) compared the performance of general extreme value (GEV) with PWM, WAK with PWM, and log-Pearson type 3 (LP3) with MMI conditioned on GEV, WAK and LP3 populations. While many more studies have been reported, the ones mentioned above have LP3 with MMI figuring as one of the competing estimators. The general conclusion of all these studies seems to be that LP3 distribution with MMI depicts poorer performance. This comes as a little surprise since LP3 distribution being a three parameter model, its MMI quantile estimator depends on the sample skewness estimate, a statistic with a significant downward bias (Wallis, et al., 1974), algebraic bounds (Kirby, 1974), and large sampling variance. Therefore, it is natural to search for better quantile estimators of LP3 in order to make LP3-based estimators more competitive in future robustness studies.

The objective of this study is to evaluate and compare the performance of the quantile estimation methods of LP3 distribution. It is worth noting that LP3 distribution being a 3 parameter model exhibits more versatility than the 2-parameter constant skew models such as EV1 distribution. Theoretically, it is capable of modeling the large skew and kurtosis behavior of annual flood series likely to be encountered in practice. It would, therefore, seem natural to expect that an LP3 estimator, behaving robustly on different "coefficient of variation-skewness" populations of LP3 distribution, would have a better chance of performing satisfactorily on populations based on other distributions such as Wakeby.

Much interest has been generated in the log-Pearson 3 distribution since it was first recommended by the U.S. Water Resources Council (USWRC) (1967), and subsequently updated (1975, 1977, 1981) as the base method of flood frequency analysis in the United States. Bobee (1975) studied the theoretical properties of the LP3 distribution and suggested an estimation method based on the moments of the real data whereas the USWRC has advocated an estimation method based on the moments of the log-transformed data. Condie (1977) studied the MLE estimates of the LP3 distribution and derived analytical results for calculating the asymptotic standard error of the MLE quantile estimator. Nozdryn-Plotnicki, et al. (1979) carried out sampling experiments to compare the performance of the three estimation methods over the LP3 parameter space representative of Canadian flood data. Rao (1980b, 1983) proposed a new method called method of mixed moments (MIX) which obviated the need to use the sample estimates of the skewness coefficients in estimating the parameters and quantiles via sample moment estimates. For the parameter space considered by him, MIX has performed well in comparison to other methods.

5.2 The LP3 Distribution

Let Y = lnX be a Pearson type 3 variate. The density function of Y is given by

$$g(y) = \frac{1}{|a|\Gamma_b} \cdot [\frac{y-c}{a}]^{b-1} \cdot \exp[-(\frac{y-c}{a})] \tag{5.1}$$

X, by definition, is log-Pearson 3 variate, and its density function, easily derived from (5.1), is given by

$$f(x) = \frac{1}{|a|x\Gamma_b} \cdot [\frac{\ln x-c}{a}]^{b-1} \cdot \exp[-\frac{\ln x-c}{a}] \tag{5.2}$$

where a, b and c are the scale, shape, and location parameters respectively. The parameter b is positive and Γ_b denotes the gamma function.

The mean, variance, and skewness coefficient of Y are given by

$$\text{Mean:} \quad \mu_y = c + ab \tag{5.3}$$

Variance: $\sigma_y^2 = ba^2$ (5.4)

Skew: $\gamma_y = \frac{|a|}{a} \cdot \frac{2}{b^{1/2}}$ (5.5)

The moments of X are given by

$$\mu_r' = \frac{\exp(rc)}{(1-ra)^b}, \; 1 - ra > 0, \; r = 1, 2, 3 \quad (5.6)$$

where μ_r' denotes the r-th moment of X about the origin.

From (5.6) the mean, variance, coefficient of variation (CV), skewness coefficient (skew), and kurtosis of X are given by

Mean: $\mu = \frac{\exp(c)}{(1-a)^2}$ (5.7)

Variance: $\sigma_x^2 = \exp(2c) \cdot A$ (5.8)

Coefficient of Variation: $\beta = (1-a)^b \cdot A^{1/2}$ (5.9)

Skewness Coefficient: $\gamma = [\frac{1}{(1-3a)^b} - \frac{3}{(1-a)^b(1-2a)^b} + \frac{2}{(1-a)^{3b}}] \cdot A^{3/2}$ (5.10)

Kurtosis: $\lambda = [\frac{1}{(1-4a)^b} - \frac{4}{(1-a)^b(1-3a)^b} + \frac{6}{(1-a)^{2b}(1-2a)^b} - \frac{3}{(1-a)^{4b}}] \cdot A^{-2}$ (5.11)

where

$$A = [\frac{1}{(1-2a)^b} - \frac{1}{(1-a)^{2b}}]$$

We note that the coefficient of variation, skewness, and kurtosis are independent of the location parameter c.

Consider equation (5.5). If $a > 0$, then $\gamma_y > 0$ implying Y is positively skewed and $c < Y < +\infty$. In this case, X is also positively skewed (Rao, 1980a), and $\exp(c) < X < +\infty$. If $a < 0$, then $\gamma_y < 0$ implying that Y is negatively skewed and $-\infty < Y < c$. In this case, X is either positively or negatively skewed depending upon the values of the parameters a and b, and $-\infty < X < \exp(c)$. For this case, the density $f(x) = 0$, may be arbitrarily defined as zero (Rao, 1980a).

The overall geometric shape of the LP3 distribution is governed by the parameters a and b (Rao, 1980a; Bobee, 1975). The density function is capable of assuming diverse shapes such as reverse-J, U, J, and of course, unimodal (skewed) bell shape. Hoshi and Burges (1981) point out that if $\gamma < \beta^3 + 3\beta$ than $a < 0$, $0 < x < \exp(c)$, $\lambda_{LP3} < \lambda_{LN3}$, and vice versa, where λ_{LP3} is the coefficient of kurtosis for the LP3 distribution (5.11), and λ_{LN3} is the coefficient of kurtosis for the three parameter log normal (LN3) distribution. The LP3 distribution degenerates to the log normal distribution when the parameters a and b become zero and infinity respecitvely (or equivalently, when $\gamma = \beta^3 + 3\beta$ and $\gamma_y = 0$).

5.3 Methods of Parameter Estimation

5.3.1 Method of Moments (Direct) - MMD

This method, proposed by Bobee (1975), uses the sample estimates of moments of the untransformed (real) data. Using (5.6), we can write:

$$\ln \mu_1' = c - b \ln(1-a) \tag{5.12}$$

$$\ln \mu_2' = 2c - b \ln(1-2a) \tag{5.13}$$

$$\ln \mu_3' = 3c - b \ln(1-3a) \tag{5.14}$$

(5.12) - (5.14) can be rearranged to give:

$$\frac{\ln \mu_3' - 3 \ln \mu_1'}{\ln \mu_2' - 2 \ln \mu_1} = \frac{3 \ln(1-a) - \ln(1-3a)}{2 \ln(1-a) - \ln(1-2a)} \;(= B \text{ say}) \tag{5.15}$$

For the sample under consideration, $B = (\ln \mu_3' - 3 \ln \mu_1')/(\ln \mu_2' - 2 \ln \mu_3')$ can be estimated from the sample estimates of the first three moments μ_1', μ_2', μ_3' about the origin ($= n^{-1} \sum_{i=1}^{n} x_i^j$, $j = 1, 2, 3$).

The right-hand side of (5.15), which is a function of the parameter a only (say B(a)), reveals that $a < 1/3$. In the limit, B(a) approaches ∞, 3, and 2 as 'a' approaches 1/3, 0, and $-\infty$ respectively. It should be possible to approximate the B(a) versus 'a' relation by a series of polynomials, as for example in Kite (1977). Then a good approximation of sample estimate of 'a' could directly be found from the sample estimated value of B and should be good enough for most fitting problems. However, for purposes of simulation, a large number of (a-B(a)) points were generated in the region $a < 1/3$ (Bobee, 1975). Table 5.1 lists some of these values. Subsequently, a sample estimate of 'a' was interpolated corresponding to the sample estimated B value from the generated a-B(a) points, and refined using the Newton-Raphson method applied to (5.15). With the interpolated value of 'a' being a very good starting solution, the iterative scheme quickly converged to the true solution to any desired degree of significant digit accuracy. The parameters 'b' and 'c' were then estimated using (5.12) and (5.13).

5.3.2 Method of Moments (Indirect) - MMI1 and MMI2

This is basically the method advocated by the U.S. Water Resources Council. It is the method of moments applied to the log-transformed data. The method utilizes equations (5.3) - (5.5) for estimating the parameters. Details of the method can be found in U.S. Water Resources Council's Bulletin Nos. 15, 17A and 17B, Rao (1980b), and others. Two variations of MMI were tested in simulation studies here. They essentially differed in the sample skewness estimator used on the log-transformed data:

$$g_y = \frac{n}{(n-1)(n-2)} \cdot \sum_{i=1}^{n} \frac{(x_i - \bar{x})^3}{S_x^3} \tag{5.16}$$

TABLE 5.1 : TABLE FOR ESTIMATING SCALE PARAMETER 'A'
DIRECT METHOD OF MOMENTS (MMD)

A	B	*	A	B	*	A	B
-4000.00000	2.03784	*	-0.00050	2.99900	*	0.11111	3.30930
-3000.00000	2.03932	*	-0.00040	2.99920	*	0.12500	3.36624
-2000.00000	2.04162	*	-0.00030	2.99940	*	0.13333	3.40343
-1000.00000	2.04623	*	-0.00020	2.99960	*	0.14006	3.43528
-500.00000	2.05195	*	-0.00010	2.99980	*	0.14045	3.43720
-250.00000	2.05923	*	-0.00005	2.99990	*	0.14085	3.43914
-200.00000	2.06201	*	-0.00001	2.99998	*	0.14124	3.44109
-125.00000	2.06873	*	0.00000	3.00000	*	0.14164	3.44306
-100.00000	2.07242	*	0.00001	3.00002	*	0.14205	3.44505
-50.00000	2.08654	*	0.00005	3.00010	*	0.14225	3.44605
-25.00000	2.10634	*	0.00010	3.00020	*	0.14265	3.44807
-12.50000	2.13498	*	0.00020	3.00040	*	0.14306	3.45010
-10.00000	2.14684	*	0.00030	3.00060	*	0.14327	3.45113
-5.00000	2.19521	*	0.00040	3.00080	*	0.14368	3.45319
-2.50000	2.26716	*	0.00050	3.00100	*	0.14409	3.45527
-2.00000	2.29663	*	0.00060	3.00120	*	0.14443	3.45695
-1.25000	2.36969	*	0.00070	3.00140	*	0.14455	3.45759
-1.00000	2.40942	*	0.00080	3.00160	*	0.14472	3.45843
-0.50000	2.54794	*	0.00090	3.00180	*	0.14684	3.46929
-0.40000	2.59470	*	0.00100	3.00200	*	0.15129	3.49264
-0.33333	2.63252	*	0.00125	3.00251	*	0.15601	3.51846
-0.25000	2.69009	*	0.00167	3.00335	*	0.16103	3.54716
-0.20000	2.73193	*	0.00200	3.00402	*	0.16367	3.56275
-0.12500	2.80904	*	0.00250	3.00503	*	0.16920	3.59680
-0.10000	2.83972	*	0.00333	3.00672	*	0.17513	3.63528
-0.06667	2.88561	*	0.00400	3.00808	*	0.18149	3.67915
-0.05000	2.91106	*	0.00500	3.01013	*	0.18484	3.70348
-0.04000	2.92725	*	0.00667	3.01356	*	0.19194	3.75786
-0.03333	2.93845	*	0.00800	3.01632	*	0.19569	3.78839
-0.02500	2.95293	*	0.01000	3.02051	*	0.20367	3.85765
-0.02000	2.96190	*	0.01111	3.02286	*	0.20790	3.89715
-0.01667	2.96800	*	0.01250	3.02581	*	0.21692	3.98852
-0.01250	2.97576	*	0.01400	3.02878	*	0.22173	4.04177
-0.01111	2.97838	*	0.01667	3.03478	*	0.22676	4.10126
-0.01000	2.98049	*	0.02000	3.04211	*	0.23202	4.16819
-0.00800	2.98431	*	0.02222	3.04706	*	0.23753	4.24411
-0.00500	2.99012	*	0.02500	3.05334	*	0.24331	4.33101
-0.00400	2.99208	*	0.02941	3.06351	*	0.24938	4.43154
-0.00333	2.99339	*	0.03030	3.06559	*	0.25575	4.54934
-0.00250	2.99503	*	0.03333	3.07275	*	0.26247	4.68948
-0.00200	2.99602	*	0.04000	3.08893	*	0.26954	4.85935
-0.00125	2.99751	*	0.05000	3.11437	*	0.27701	5.07016
-0.00100	2.99800	*	0.06667	3.16024	*	0.28490	5.33998
-0.00090	2.99820	*	0.07692	3.19087	*	0.29326	5.70016
-0.00080	2.99840	*	0.08333	3.21106	*	0.30211	6.21160
-0.00070	2.99860	*	0.09091	3.23603	*	0.31153	7.01497
-0.00060	2.99880	*	0.10000	3.26772	*	0.32154	8.56194

NOTE : B = (3 * LN(1-A) - LN(1-3*A)) / (2 * LN(1-A) -LN(1-2*A))

$$g_y' = (1 + \frac{8.5}{n}) \cdot g_y \tag{5.17}$$

where n is the sample size, and $\bar{x}$ and S_x are the sample mean and standard deviation respectively.

5.3.3 Method of Mixed Moments - MIX

Rao (1980b, 1983) proposed this method with the objective of obviating the use of the sample skewness coefficient in parameter estimation. He tried various combinations of mixing the first two moments of the untransformed and log-transformed samples and found one particular combination to be preferable on the basis of sampling properties. The method, referred to here as MIX, is outlined below. A simple procedure, that directly finds the parameters and thus eliminates the need for iterative approach, is also proposed. The practicing hydrologists should find the proposed procedure useful in estimation of parameters by MIX.

MIX conserves the sample mean and variance of the untransformed data, and the sample mean of the log-transformed data. Let $\bar{x}$, s_x^2, $\bar{y}$ be the sample estimates of mean and variance of the untransformed data, and mean of the log-transformed data respectively. Therefore, using (5.3), (5.7), and (5.8) with population statistics replaced by the sample estimates, the MIX parameter estimation equations are:

$$\bar{y} = c + ab \tag{5.18}$$

$$\bar{x} = \frac{\exp(c)}{(1-a)^b} \tag{5.19}$$

$$s_x^2 = \exp(2c) \cdot [\frac{1}{(1-2a)^b} - \frac{1}{(1-a)^{2b}}] \tag{5.20}$$

From (5.18) and (5.19), we can write:

$$\bar{y} - \ln \bar{x} = b[a + \ln(1-a)] \tag{5.21}$$

From (5.19) and (5.20), we can write:

$$\ln\left(\frac{s_x^2 + \bar{x}^2}{\bar{x}^2}\right) = 2b \ln(1-a) - b \ln(1-2a) \tag{5.22}$$

Combining (5.21) and (5.22),

$$\frac{2 \ln(1-a) - \ln(1-2a)}{\ln(1-a) + a} = P \tag{5.23}$$

where $P = \ln[(s_x^2 + \bar{x}^2)/\bar{x}^2]/(\bar{y} - \ln \bar{x})$ can be found from sample estimated $\bar{x}$, $\bar{y}$, s_x^2.

The "left-hand side" of (5.23) depends only on "a" and is defined for $a < 1/2$. It can be easily shown from (5.23) that P approaches $-\infty$, -2, and 0 as "a" approaches 1/2, 0 and $-\infty$ respectively. P is a smooth function over the domain $a < 1/2$.

It should be possible to approximate P versus "a" by a polynomial function in a manner analogous to MMI (Kite, 1977) and estimate "a" from the approximated polynomial. However, for purposes of simulation, an a-P table was generated for large number of values of $a < 1/2$ by making use of (5.23), and "a" was estimated by interpolating from this table. Table 5.2 presents a sample a-P table. For most engineering problems, this estimate should be sufficiently accurate, and if need be, can be further refined by the Newton-Raphson method applied to (5.23). Rao (1983) has reported some problems with convergence, if a "good" starting solution of "a" is not used in the iterative scheme. The technique of choosing the starting solution as the interpolated value as outlined above, always leads to fast convergence to any desired degree of accuracy. After having found the estimate of "a", estimates of "b" and "c" can be found from (5.21) and (5.18) respectively.

TABLE 5.2 : TABLE FOR ESTIMATING SCALE PARAMETER 'A'
METHOD OF MIXED MOMENTS (MIX)

A	P	*	A	P	*	A	P
-4000.00000	-0.00190	*	0.00000	-2.00000	*	0.20367	-2.81039
-3000.00000	-0.00244	*	0.00001	-2.00002	*	0.22173	-2.92392
-2000.00000	-0.00347	*	0.00005	-2.00013	*	0.23753	-3.03270
-1000.00000	-0.00626	*	0.00010	-2.00027	*	0.25575	-3.17098
-500.00000	-0.01119	*	0.00020	-2.00053	*	0.26247	-3.22584
-250.00000	-0.01977	*	0.00030	-2.00080	*	0.27701	-3.35293
-200.00000	-0.02369	*	0.00040	-2.00107	*	0.28490	-3.42712
-125.00000	-0.03451	*	0.00050	-2.00133	*	0.29326	-3.51013
-100.00000	-0.04117	*	0.00060	-2.00160	*	0.30211	-3.60366
-50.00000	-0.07052	*	0.00070	-2.00187	*	0.31153	-3.70990
-25.00000	-0.11887	*	0.00080	-2.00214	*	0.32154	-3.83172
-12.50000	-0.19675	*	0.00090	-2.00240	*	0.33000	-3.94245
-10.00000	-0.23037	*	0.00100	-2.00267	*	0.34000	-4.08398
-5.00000	-0.36956	*	0.00125	-2.00334	*	0.35000	-4.23860
-2.50000	-0.57228	*	0.00167	-2.00445	*	0.36000	-4.40844
-1.25000	-0.84064	*	0.00200	-2.00535	*	0.37000	-4.59608
-1.00000	-0.93752	*	0.00250	-2.00669	*	0.38000	-4.80482
-0.50000	-1.24592	*	0.00333	-2.00893	*	0.39000	-5.03886
-0.40000	-1.34048	*	0.00400	-2.01073	*	0.40000	-5.30371
-0.25000	-1.52001	*	0.00500	-2.01344	*	0.40500	-5.44989
-0.20000	-1.59352	*	0.00800	-2.02161	*	0.41000	-5.60671
-0.10000	-1.76954	*	0.01000	-2.02709	*	0.41500	-5.77555
-0.05000	-1.87641	*	0.01250	-2.03401	*	0.42000	-5.95802
-0.04000	-1.89966	*	0.01389	-2.03787	*	0.42500	-6.15608
-0.02500	-1.93587	*	0.02000	-2.05508	*	0.43000	-6.37210
-0.02000	-1.94830	*	0.02500	-2.06942	*	0.43500	-6.60903
-0.01250	-1.96731	*	0.02941	-2.08227	*	0.44000	-6.87053
-0.01000	-1.97375	*	0.03030	-2.08489	*	0.44500	-7.16128
-0.00800	-1.97893	*	0.03333	-2.09386	*	0.45000	-7.48738
-0.00500	-1.98677	*	0.04000	-2.11391	*	0.45500	-7.85688
-0.00400	-1.98940	*	0.05000	-2.14485	*	0.46000	-8.28086
-0.00250	-1.99336	*	0.06667	-2.19890	*	0.46500	-8.77498
-0.00200	-1.99468	*	0.07692	-2.23381	*	0.47000	-9.36239
-0.00125	-1.99667	*	0.08333	-2.25631	*	0.47500	-10.07940
-0.00100	-1.99734	*	0.10000	-2.31741	*	0.48000	-10.98753
-0.00090	-1.99760	*	0.11111	-2.36039	*	0.48500	-12.20337
-0.00080	-1.99787	*	0.12500	-2.41687	*	0.49000	-13.99187
-0.00070	-1.99813	*	0.13889	-2.47668	*	0.49500	-17.20950
-0.00060	-1.99840	*	0.14104	-2.48628	*	0.49600	-18.27875
-0.00050	-1.99867	*	0.14306	-2.49535	*	0.49700	-19.67587
-0.00040	-1.99893	*	0.14409	-2.50001	*	0.49800	-21.67429
-0.00030	-1.99920	*	0.14451	-2.50190	*	0.49900	-25.14880
-0.00020	-1.99946	*	0.14459	-2.50228	*	0.49910	-25.68150
-0.00010	-1.99973	*	0.14684	-2.51255	*	0.49920	-26.27815
-0.00005	-1.99987	*	0.15601	-2.55545	*	0.49930	-26.95593
-0.00003	-1.99992	*	0.16367	-2.59268	*	0.49940	-27.73999
-0.00001	-1.99997	*	0.18484	-2.70271	*	0.49950	-28.66924

NOTE : P = (2 * LN(1-A) - LN(1-2*A)) / (LN(1-A) + A)

5.3.4 Method of Maximum Likelihood Estimation - MLE

The MLE method in the context of LP3 distribution (or equivalently, P3 with log-transformed data) has been investigated by Matalas (1973), Condie (1977), Condie, et al. (1979), Nozdryn-Plotnicki (1979), and Rao (1986), among others. The methods proposed by all the investigators, except Rao (1986), essentially involve searching for the solution within a limited range of parameter c, for example, ($\ln x_{min} - 0.1$, $\ln x_{min} - 50$) for negative g_y, and ($\ln x_{max} + 0.01$, $\ln x_{max} + 50$) for positive g_y, where x_{min}, x_{max} are the minimum and maximum values of the untransformed sample, and g_y is the estimated skewness coefficient of the log-transformed sample. If no solution is found within this range of c, the MLE estimation effort is considered a failure for the sample under consideration. However, Rao (1986) carried out a thorough investigation of MLE parameter estimation equations and found that, in general, multiple (two or three) solutions of the MLE equations exist with two solutions corresponding to $c \to \pm\infty$ and a third likely solution in the vicinity of "greater than $\ln x_{max}$" or "less than $\ln x_{min}$" regions. In this situation, a criterion is required to select the "best" MLE solution. Rao (1986) proposed a comparison of sample estimated and MLE estimated means and variances as the criteria for selecting the best solution.

In what follows, the method is briefly discussed, together with a proposed algorithm to effect the multiple solutions. An objective criterion, logically resulting from the maximum likelihood estimation philosophy, for selecting the "best" MLE solution is also proposed.

Let (x_1, x_2, ..., x_n) be a sample of size n, drawn from a probability density function f(x;a,b,c). The likelihood function is defined as

$$L = \prod_{i=1}^{n} f(x_i;a,b,c) \qquad (5.24)$$

From (5.24) and (5.2), the log-likelihood function can be written as

$$LL = \sum_{i=1}^{n} \ln f(x_i;a,b,c)$$

$$= - n \ln a - \Sigma \ln x - n \ln \Gamma_b$$

$$+ (b-1) \Sigma \ln [(\ln x-c))/a] - (1/a) \Sigma (\ln x-c) \qquad (5.25)$$

and the summations are over n sample values wherever not specified.

The objective of the method is to maximize the likelihood function (or equivalently, to maximize the log-likelihood function). Thus, the following parameter estimation equations result:

$$\frac{\partial (LL)}{\partial a} = - n\, ab + \Sigma (\ln x-c) = 0 \qquad (5.26)$$

$$\frac{\partial (LL)}{\partial b} = - n\, \psi(b) + \Sigma \ln[\frac{(\ln x-c)}{a}] = 0 \qquad (5.27)$$

$$\frac{\partial (LL)}{\partial c} = \frac{n}{a} - (b-1) \Sigma \frac{1}{(\ln x-c)} = 0 \qquad (5.28)$$

(5.26) and (5.28) can be rearranged (Rao, 1986) to give

$$a = \frac{s_1}{nb} \qquad (5.29)$$

$$b = \frac{s_1\, s_2}{(s_1\, s_2 - n^2)} \qquad (5.30)$$

where $s_1 = \Sigma (\ln x-c)$ and $s_2 = \Sigma 1/(\ln x-c)$.

For a specified value of c, parameters b and a can be explicitly found from (5.30) and (5.29) respectively. Substitution of these a, b and c values in (5.27) yields $\partial(LL)/\partial b = R$. Rao (1986) investigated the variation of R with c and came up with three general patterns of R versus c relationships. These cases are reproduced here in Fig. 5.1 from Rao's paper. For clarity, we note in Fig. 5.1 that before

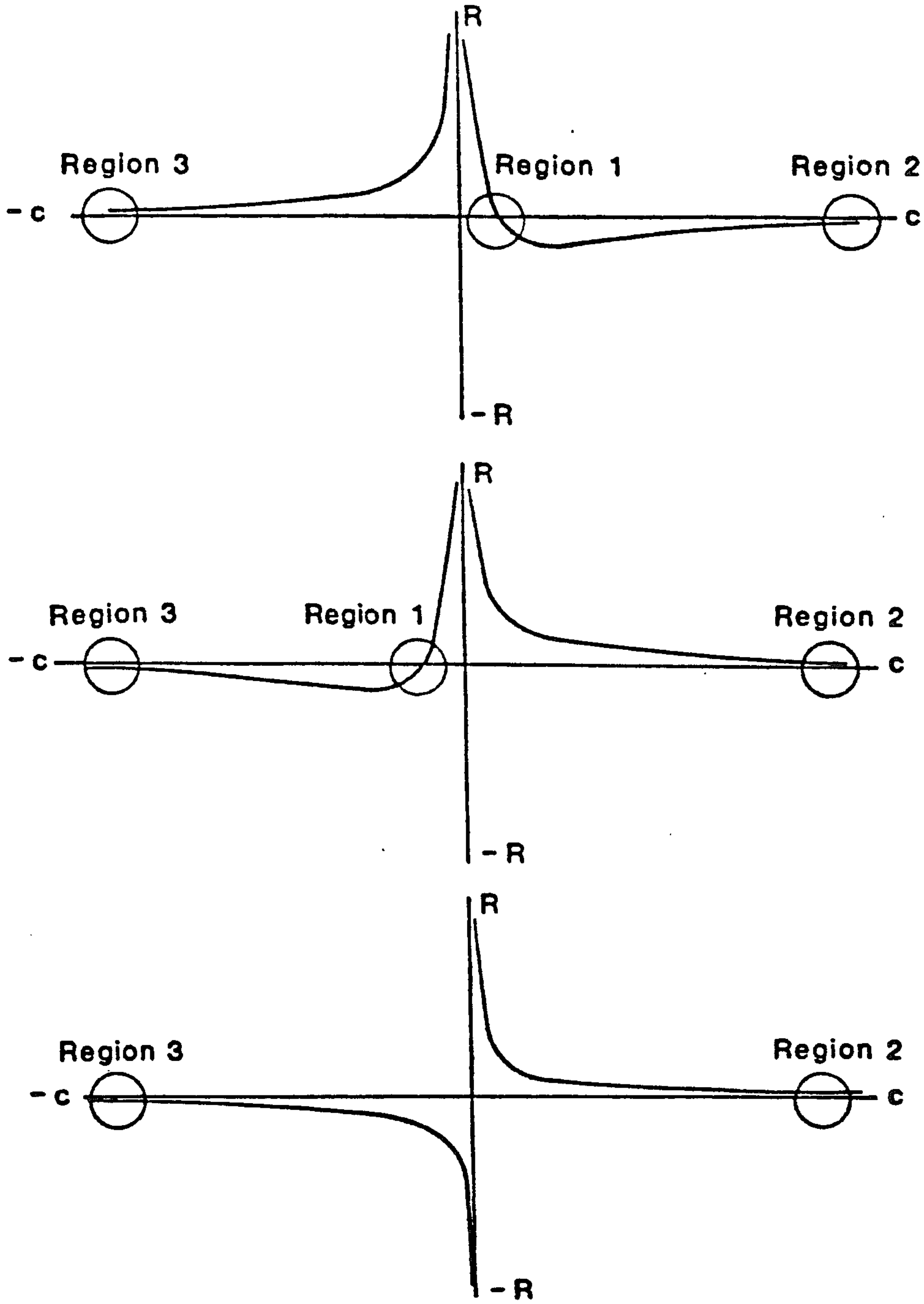

Fig 5.1 : R - c Relationships and Regions of Possible MLE solution of the LP3 Model (from : Rao, 1986)

log-transforming, the samples were standardized by dividing x_i by $\bar{x}$. The range of c investigated was $-\infty < c < \ln(x_{min}/\bar{x}) < 0$, and $0 < \ln(x_{max}/\bar{x}) < c < +\infty$. Similar investigations were made by the author and the following observations are worth mentioning. (1) After steeply falling (or rising), and possibly crossing the c-axis, R tends to be very flat. In fact, if the calculations are not made in double precision accuracy, R erroneously crosses the c-line at several points in a decaying fashion. Apparently, this was the behavior observed by Condie (1977) when he remarked about "several inflexion points" of R. This oversight can obviously lead to erroneous solutions. (2) A "region 1 - positive c" solution, if it exists, excludes the possibility of a "region 1 - negative c" solution and vice versa. In contrast, the region 2 and region 3 solutions corresponding to $c \to \pm\infty$ always exist.

It is proposed that the solution maximizing the log-likelihood function (5.25) is the "best" MLE solution. An investigation was made into the variation of LL (5.25) with parameter c. It was found that (1) a 'region-1' solution, if it existed, did not necessarily correspond to a local maxima of LL. (2) The LL corresponding to $c \to \pm\infty$ always approached the same numerical values. (3) A local maxima of LL corresponding to a 'region-1' solution also proved to be a global maxima. (4) The solution of c maximizing LL, in general, provided better MLE estimates of the mean and variance than the ones satisfying (5.26) - (5.28) but not maximizing LL in the domain of parameter c. Based on these findings, the following procedure was employed to find the MLE parameter estimates in simulation studies. Let the population $c > 0$

Step (1): Search for c in the range $[\ln(x_{max}/\bar{x}) + 0.01,$ $\ln[x_{max}/\bar{x}) + 50]$. If it exists, find it by regula-falsi method, and calculate the corresponding LL. If LL is a local maxima, STOP. c is the MLE solution. If c does not exist, go to step (2).

Step (2): Search for c in the range $[\ln(x_{max}/\bar{x}) - 0.01,$ $\ln[x_{max}/\bar{x}) - 50]$. If it exists and is a local maxima, STOP. MLE solution is found. Else, go to step (3).

Step (3): Find $c_2 \to \pm\infty$, corresponding to $|R| \leq 10^{-9}$. This is the MLE solution.

For population $c < 0$, a complementary procedure was followed.

5.3.5 Method of Entropy - ENT

Singh and Singh (1985) used the concept of entropy to derive a new set of estimation equations for the Pearson type III distribution. The parameter estimation equations, with theoretical expectations replaced by sample moments are

$$\Sigma(\ln x - c) = nab \tag{5.31}$$

$$S_y^2 = a^2 b \tag{5.32}$$

$$\Sigma \ln[\frac{\ln x - c}{a}] = n\psi(b) \tag{5.33}$$

where S_y^2 is the sample variance estimate from the log-transformed data. It is interesting to note that the two equations of this system, namely (5.31) and (5.33), are exactly identical to the two MLE equations, namely (5.26) and 5.27). The third equation (5.32) appears to be simpler than the corresponding (5.28). In fact, (5.32) is the same as (5.4) which is also used in MMI.

The system (5.31) - (5.33) offers similar computational problems as MLE. (5.31) requires that $\Sigma(\ln x - c)$ and a should have the same signs.

Therefore, a is positive, if $c < \ln x_i$, that is, c is a lower bound, and vice versa. (5.31) (or (5.32)) allows for $b > 0$, whereas (5.28) of MLE requires $b > 1$. Therefore, the ENT system of equations is less restrictive than MLE system.

Equations (5.31) - (5.33) have multiple roots. After rearranging, the equations can be written as,

$$a = \pm S_y/\sqrt{b} \tag{5.34}$$

$$c = \bar{y} \pm S_y \sqrt{b} \tag{5.35}$$

Hence, for every value of b, we have two possible values of a and c. The ENT solution was obtained by following similar procedure as in the case of MLE.

5.4 Quantile Estimation

Let

$$W = \frac{\ln X-c}{a} \tag{5.36}$$

W is the standard gamma variate with shape parameter b (and scale parameter 1) having the density function

$$g(w) = \frac{\exp(-w)\ w^{b-1}}{\Gamma(b)},\ w \geq 0 \tag{5.37}$$

Thus, the cumulative distribution function F(x) of the LP3 variate X can be expressed as

$$\begin{aligned} F(x) &= \text{Prob}(X \leq x) \\ &= \text{Prob}\left[\frac{\ln X-c}{a} < \frac{\ln x-c}{a}\right],\ a > 0 \\ &= \text{Prob}\left[\frac{\ln X-c}{a} > \frac{\ln x-c}{a}\right],\ a < 0 \end{aligned} \tag{5.38}$$

That is,

$$F(x) = \begin{cases} G(w_x) & ,\ a > 0 \\ 1 - G(w_x), & a < 0 \end{cases} \tag{5.39}$$

where $G(w_x)$ is the cumulative distribution function of the standard Gamma variate W defined as

$$G(w_x) = \int_0^{w_x} g(w)\, dw, \qquad w_x = \frac{\ln x - c}{a} \tag{5.40}$$

Therefore, the T-year return period quantile, x_T, can be obtained from:

$$\ln x_T = c + a\, w_T \tag{5.41}$$

where w_T is a standard Gamma quantile corresponding to a non-exceedance probability $[G(w_T)]$ of $(1 - 1/T)$ for $a > 0$, or $1/T$ for $a < 0$. w_T can be written as

$$w_T = b^{1/2}(K_T + b^{1/2}) \tag{5.42}$$

where K_T is the frequency factor corresponding to the same probability level as W_T. Using (5.42), we can write (5.41) as

$$\begin{aligned} \ln x_T &= c + ab + ab^{1/2}\, K_T \\ &= \mu_y + \sigma_y\, K_T \end{aligned} \tag{5.43}$$

K_T depends on the probability level and the shape parameter b (or equivalently γ_y). Its values are tabulated in U.S. Water Resources Council Bulletin 17 for $\gamma_y = -9.0$ to 9.0 for a wide range of exceedance probability levels. The values of K_T were interpolated from table and used in (5.43) to obtain the quantile estimate x_T.

5.5 Experimental Design

To assess the performance of various methods of estimation outlined above, Monte Carlo sampling experiments were performed. Annual flood data generally lie in the area of the β-γ diagram delineated by $0.3 < \beta < 0.8$ and γ upward of 1 (Rossi, et al., 1986; Wallis, et al., 1985; Landwehr, et al., 1978). Based on this consideration, five cases of LP3 population, representative of the real flood data were selected for Monte Carlo experiments. These cases are listed in Table 5.3.

Table 5.3. (μ = 1): LP3 population cases considered in sampling experiments.

LP3 Population	CV (β)	Skew (γ)	Parameter a	Parameter b	Parameter c	γ_y
Case 1	0.5	1	-0.11832	19.82269	2.216713	-0.45
Case 2	0.5	3	0.127683	10.30311	-1.407434	0.62
Case 3	0.5	5	0.205678	3.215257	-0.740366	1.12
Case 4	0.3	3	0.150978	2.681889	-0.438946	1.22
Case 5	0.7	3	0.059798	98.38009	-6.066213	0.20

It is noted here by $\lambda_1 < \lambda_{LN3} < \lambda_5 < \lambda_2 < \lambda_4 < \lambda_3$ where the subscripts of λ refer to the LP3 population case.

For each of the population cases listed in Table 5.3, 1000 random samples of size 10, 20, 30, 50 and 75 were generated and parameters and quantiles estimated from methods of estimation outlined earlier.

The 1000 estimated values of estimated parameters and quantiles for each sample size and population case were used to approximate the values of the following performance indices for that case.

$$\text{Standardized Bias, BIAS} = \frac{E(\hat{x}) - x}{x} \tag{5.44}$$

$$\text{Standard Error, SE} = \frac{\sigma(\hat{x})}{x} \tag{5.45}$$

$$\text{Root Mean Square Error, RMSE} = \frac{E[(\hat{x} - x)^2]^{1/2}}{x} \tag{5.46}$$

where $\hat{x}$ is an estimator (parameter or quantile) of x, E(·) denotes statistical expectation, and σ(·) denotes standard deviation of the respective random variable. $E(\hat{x})$ and $\sigma(\hat{x})$ were calculated as

$$E(\hat{x}) = \frac{\Sigma \hat{x}}{N} \qquad (5.47)$$

$$\sigma(\hat{x}) = [\frac{1}{N-1} \Sigma\{\hat{x} - E(\hat{x})\}^2]^{1/2} \qquad (5.48)$$

where the summations are over N estimates $\hat{x}$ of x, N being the number of random samples used in estimation (= 1000 here).

It is easy to write

$$RMSE = [\frac{N-1}{N} SE^2 + BIAS^2]^{1/2} \qquad (5.49)$$

Due to the limited number of random samples used, the results are not expected to reproduce the true values of BIAS, SE and RMSE, but they do provide a means of comparing the performance of various estimation methods.

(5.36) was used to facilitate the generation of LP3 random numbers. To start with, standard Gamma numbers, W_R, were generated through Gamma generator GGAMR (IMSL, 1981). Then the corresponding LP3 numbers, X_R, were obtained as

$$X_R = \exp(a\ W_R + c) \qquad (5.50)$$

5.6 Results and Discussion

In general, unusually high BIAS, SE and RMSE were observed for parameter estimators $\hat{a}$, $\hat{b}$ and $\hat{c}$ of all methods. However, the inter-correlation among the parameter estimates was such that reasonable quantile estimates were obtained. Tables 5.4-5.8, 5.9-5.13, and 5.14-5.18 show the BIAS, SE and RMSE of quantile estimate for five population cases respectively.

TABLE 5.4

BIAS OF SELECTED QUANTILES
(CASE - 1 : C.V. = 0.5 SKEW = 1.0)

		RETURN PERIOD					
METHOD	SAMPLE SIZE	10	25	50	100	200	500
MMD	10	-0.038	-0.082	-0.111	-0.136	-0.158	-0.183
MMI1	10	-0.005	0.015	0.039	0.069	0.107	0.168
MMI2	10	-0.045	-0.037	-0.012	0.030	0.090	0.206
MIX	10	-0.036	-0.030	-0.019	-0.004	0.015	0.045
MLE	10	-0.009	0.030	0.068	0.111	0.161	0.236
ENT	10	-0.047	-0.096	-0.127	-0.155	-0.179	-0.206
MMD	20	-0.018	-0.045	-0.063	-0.078	-0.092	-0.108
MMI1	20	-0.003	0.009	0.023	0.040	0.061	0.094
MMI2	20	-0.021	-0.022	-0.015	-0.003	0.015	0.047
MIX	20	-0.020	-0.018	-0.012	-0.004	0.006	0.024
MLE	20	-0.012	-0.005	0.009	0.030	0.058	0.108
ENT	20	-0.029	-0.061	-0.082	-0.101	-0.118	-0.139
MMD	30	-0.011	-0.030	-0.043	-0.055	-0.066	-0.077
MMI1	30	-0.002	0.009	0.021	0.035	0.051	0.077
MMI2	30	-0.013	-0.012	-0.007	0.001	0.013	0.034
MIX	30	-0.014	-0.012	-0.007	-0.001	0.006	0.019
MLE	30	-0.011	-0.009	-0.001	0.011	0.029	0.061
ENT	30	-0.021	-0.043	-0.058	-0.072	-0.084	-0.099
MMD	50	-0.005	-0.017	-0.026	-0.034	-0.041	-0.050
MMI1	50	0.000	0.008	0.015	0.024	0.035	0.051
MMI2	50	-0.006	-0.006	-0.004	0.000	0.006	0.017
MIX	50	-0.007	-0.006	-0.003	0.000	0.005	0.012
MLE	50	-0.008	-0.011	-0.011	-0.009	-0.005	0.004
ENT	50	-0.011	-0.025	-0.034	-0.043	-0.051	-0.060
MMD	75	-0.002	-0.011	-0.018	-0.024	-0.030	-0.036
MMI1	75	0.001	0.006	0.012	0.018	0.026	0.037
MMI2	75	-0.003	-0.003	-0.002	0.001	0.005	0.012
MIX	75	-0.004	-0.003	-0.002	0.000	0.003	0.008
MLE	75	-0.005	-0.007	-0.008	-0.007	-0.005	-0.001
ENT	75	-0.007	-0.016	-0.023	-0.029	-0.034	-0.040

TABLE 5.5

BIAS OF SELECTED QUANTILES
(CASE - 2 : C.V. = 0.5 SKEW = 3.0)

METHOD	SAMPLE SIZE	RETURN PERIOD 10	25	50	100	200	500
MMD	10	-0.004	-0.078	-0.136	-0.192	-0.247	-0.315
MMI1	10	-0.001	0.000	0.012	0.035	0.072	0.152
MMI2	10	-0.016	0.027	0.096	0.212	0.402	0.863
MIX	10	-0.025	-0.054	-0.074	-0.092	-0.108	-0.126
MLE	10	-0.030	-0.076	-0.108	-0.138	-0.166	-0.197
ENT	10	0.011	0.062	0.115	0.183	0.268	0.413
MMD	20	0.003	-0.050	-0.093	-0.136	-0.178	-0.232
MMI1	20	-0.007	-0.010	-0.008	0.000	0.015	0.047
MMI2	20	-0.009	0.010	0.037	0.077	0.134	0.245
MIX	20	-0.018	-0.045	-0.063	-0.080	-0.095	-0.112
MLE	20	-0.010	-0.010	-0.004	0.010	0.031	0.073
ENT	20	0.001	0.026	0.052	0.085	0.123	0.186
MMD	30	0.010	-0.035	-0.071	-0.109	-0.146	-0.193
MMI1	30	-0.004	-0.007	-0.006	-0.001	0.009	0.029
MMI2	30	-0.003	0.008	0.025	0.049	0.082	0.143
MIX	30	-0.011	-0.034	-0.051	-0.066	-0.080	-0.096
MLE	30	-0.000	0.010	0.024	0.043	0.068	0.111
ENT	30	0.002	0.019	0.035	0.056	0.080	0.119
MMD	50	0.016	-0.018	-0.047	-0.077	-0.107	-0.146
MMI1	50	0.000	-0.001	0.001	0.005	0.012	0.027
MMI2	50	0.001	0.009	0.020	0.035	0.055	0.090
MIX	50	-0.004	-0.022	-0.034	-0.046	-0.056	-0.068
MLE	50	0.004	0.017	0.031	0.047	0.068	0.101
ENT	50	0.004	0.015	0.026	0.039	0.053	0.076
MMD	75	0.017	-0.010	-0.034	-0.058	-0.083	-0.116
MMI1	75	0.000	0.001	0.002	0.006	0.012	0.023
MMI2	75	0.001	0.008	0.016	0.026	0.040	0.063
MIX	75	-0.002	-0.016	-0.025	-0.034	-0.042	-0.050
MLE	75	0.003	0.013	0.022	0.034	0.048	0.069
ENT	75	0.003	0.011	0.019	0.029	0.040	0.058

TABLE 5.6

BIAS OF SELECTED QUANTILES
(CASE - 3 : C.V. = 0.5 SKEW = 5.0)

METHOD	SAMPLE SIZE	RETURN PERIOD 10	25	50	100	200	500
MMD	10	0.009	-0.077	-0.146	-0.214	-0.280	-0.362
MMI1	10	-0.001	-0.008	-0.002	0.017	0.054	0.138
MMI2	10	-0.022	0.025	0.104	0.241	0.474	1.062
MIX	10	-0.019	-0.063	-0.095	-0.127	-0.156	-0.193
MLE	10	-0.031	-0.109	-0.167	-0.222	-0.274	-0.338
ENT	10	0.004	0.021	0.047	0.087	0.144	0.259
MMD	20	0.017	-0.046	-0.100	-0.154	-0.207	-0.275
MMI1	20	-0.007	-0.016	-0.017	-0.011	0.003	0.040
MMI2	20	-0.013	0.008	0.041	0.093	0.169	0.326
MIX	20	-0.012	-0.050	-0.079	-0.107	-0.133	-0.165
MLE	20	-0.011	-0.043	-0.064	-0.081	-0.095	-0.105
ENT	20	-0.004	-0.008	-0.007	-0.003	0.005	0.023
MMD	30	0.026	-0.027	-0.073	-0.120	-0.167	-0.229
MMI1	30	-0.005	-0.011	-0.010	-0.005	0.008	0.037
MMI2	30	-0.008	0.008	0.031	0.066	0.116	0.212
MIX	30	-0.006	-0.038	-0.063	-0.087	-0.109	-0.137
MLE	30	-0.000	-0.008	-0.012	-0.012	-0.009	0.001
ENT	30	-0.002	-0.008	-0.009	-0.009	-0.006	0.002
MMD	50	0.032	-0.011	-0.048	-0.088	-0.129	-0.182
MMI1	50	-0.001	-0.004	-0.004	0.001	0.010	0.031
MMI2	50	-0.002	0.008	0.023	0.044	0.073	0.128
MIX	50	0.001	-0.025	-0.045	-0.065	-0.084	-0.106
MLE	50	0.008	0.015	0.023	0.034	0.048	0.070
ENT	50	0.000	-0.007	-0.012	-0.016	-0.018	-0.019
MMD	75	0.034	-0.000	-0.032	-0.065	-0.100	-0.146
MMI1	75	0.000	-0.001	0.001	0.005	0.013	0.029
MMI2	75	-0.000	0.008	0.019	0.035	0.055	0.091
MIX	75	0.003	-0.017	-0.033	-0.048	-0.063	-0.080
MLE	75	0.008	0.018	0.028	0.039	0.053	0.074
ENT	75	0.001	-0.006	-0.010	-0.013	-0.015	-0.016

TABLE 5.7

BIAS OF SELECTED QUANTILES
(CASE - 4 : C.V. = 0.3 SKEW = 3.0)

METHOD	SAMPLE SIZE	RETURN PERIOD 10	25	50	100	200	500
MMD	10	-0.011	-0.067	-0.111	-0.156	-0.199	-0.255
MMI1	10	-0.006	-0.020	-0.026	-0.028	-0.024	-0.007
MMI2	10	-0.021	0.002	0.039	0.097	0.184	0.367
MIX	10	-0.021	-0.041	-0.056	-0.069	-0.080	-0.094
MLE	10	-0.025	-0.083	-0.128	-0.172	-0.214	-0.268
ENT	10	-0.004	-0.011	-0.013	-0.013	-0.009	0.003
MMD	20	-0.004	-0.046	-0.080	-0.115	-0.149	-0.193
MMI1	20	-0.008	-0.020	-0.027	-0.030	-0.030	-0.024
MMI2	20	-0.013	-0.004	0.012	0.036	0.069	0.133
MIX	20	-0.015	-0.035	-0.051	-0.065	-0.078	-0.094
MLE	20	-0.014	-0.047	-0.072	-0.097	-0.121	-0.149
ENT	20	-0.007	-0.020	-0.030	-0.038	-0.044	-0.051
MMD	30	0.003	-0.031	-0.059	-0.088	-0.117	-0.154
MMI1	30	-0.005	-0.013	-0.017	-0.018	-0.016	-0.007
MMI2	30	-0.008	-0.001	0.011	0.028	0.053	0.098
MIX	30	-0.008	-0.026	-0.038	-0.051	-0.062	-0.075
MLE	30	-0.005	-0.024	-0.039	-0.054	-0.067	-0.082
ENT	30	-0.004	-0.019	-0.029	-0.039	-0.049	-0.059
MMD	50	0.008	-0.018	-0.039	-0.062	-0.084	-0.114
MMI1	50	-0.002	-0.006	-0.008	-0.007	-0.004	0.003
MMI2	50	-0.004	0.002	0.011	0.023	0.038	0.066
MIX	50	-0.004	-0.016	-0.026	-0.035	-0.043	-0.052
MLE	50	-0.000	-0.013	-0.022	-0.031	-0.039	-0.048
ENT	50	-0.002	-0.017	-0.028	-0.039	-0.049	-0.062
MMD	75	0.010	-0.010	-0.028	-0.046	-0.065	-0.090
MMI1	75	-0.000	-0.002	-0.003	-0.002	0.001	0.007
MMI2	75	-0.001	0.004	0.010	0.019	0.030	0.049
MIX	75	-0.001	-0.011	-0.018	-0.025	-0.031	-0.038
MLE	75	0.003	-0.003	-0.007	-0.010	-0.013	-0.016
ENT	75	-0.001	-0.015	-0.026	-0.036	-0.047	-0.060

TABLE 5.8

BIAS OF SELECTED QUANTILES
(CASE - 5 : C.V. = 0.7 SKEW = 3.0)

METHOD	SAMPLE SIZE	RETURN PERIOD 10	25	50	100	200	500
MMD	10	0.004	-0.088	-0.159	-0.227	-0.291	-0.369
MMI1	10	0.006	0.023	0.053	0.101	0.172	0.317
MMI2	10	-0.018	0.034	0.124	0.282	0.554	1.263
MIX	10	-0.031	-0.069	-0.094	-0.115	-0.134	-0.155
MLE	10	-0.028	-0.050	-0.061	-0.069	-0.073	-0.072
ENT	10	0.040	0.164	0.293	0.460	0.678	1.075
MMD	20	0.009	-0.057	-0.110	-0.162	-0.212	-0.274
MMI1	20	-0.005	-0.002	0.007	0.023	0.046	0.090
MMI2	20	-0.009	0.007	0.034	0.075	0.133	0.244
MIX	20	-0.024	-0.057	-0.080	-0.099	-0.117	-0.135
MLE	20	-0.013	-0.003	0.020	0.057	0.111	0.216
ENT	20	0.018	0.084	0.147	0.221	0.308	0.446
MMD	30	0.017	-0.037	-0.081	-0.125	-0.167	-0.221
MMI1	30	-0.003	-0.002	0.005	0.016	0.032	0.062
MMI2	30	-0.004	0.006	0.023	0.047	0.081	0.143
MIX	30	-0.016	-0.044	-0.062	-0.079	-0.094	-0.109
MLE	30	-0.006	0.009	0.032	0.066	0.113	0.200
ENT	30	0.015	0.063	0.108	0.161	0.221	0.313
MMD	50	0.025	-0.015	-0.050	-0.085	-0.120	-0.165
MMI1	50	0.002	0.004	0.009	0.017	0.029	0.051
MMI2	50	0.002	0.009	0.019	0.034	0.055	0.092
MIX	50	-0.006	-0.026	-0.041	-0.053	-0.064	-0.076
MLE	50	-0.001	0.006	0.018	0.034	0.056	0.095
ENT	50	0.015	0.050	0.081	0.116	0.156	0.216
MMD	75	0.024	-0.006	-0.032	-0.059	-0.087	-0.123
MMI1	75	0.002	0.005	0.009	0.015	0.024	0.040
MMI2	75	0.003	0.009	0.016	0.027	0.040	0.064
MIX	75	-0.003	-0.018	-0.028	-0.036	-0.044	-0.051
MLE	75	-0.001	0.005	0.012	0.021	0.034	0.056
ENT	75	0.011	0.038	0.061	0.087	0.115	0.158

TABLE 5.9

STANDARD ERROR OF SELECTED QUANTILES
(CASE - 1 : C.V. = 0.5 SKEW = 1.0)

METHOD	SAMPLE SIZE	RETURN PERIOD 10	25	50	100	200	500
MMD	10	0.181	0.198	0.212	0.227	0.243	0.263
MMI1	10	0.190	0.238	0.291	0.357	0.438	0.572
MMI2	10	0.190	0.266	0.357	0.481	0.653	0.999
MIX	10	0.176	0.201	0.227	0.258	0.294	0.349
MLE	10	0.195	0.254	0.310	0.375	0.449	0.568
ENT	10	0.179	0.200	0.220	0.242	0.267	0.301
MMD	20	0.130	0.145	0.160	0.178	0.197	0.223
MMI1	20	0.133	0.165	0.199	0.239	0.285	0.355
MMI2	20	0.133	0.176	0.222	0.277	0.340	0.439
MIX	20	0.127	0.146	0.167	0.192	0.221	0.264
MLE	20	0.134	0.185	0.246	0.326	0.430	0.614
ENT	20	0.126	0.139	0.152	0.166	0.180	0.199
MMD	30	0.107	0.121	0.136	0.153	0.172	0.198
MMI1	30	0.108	0.135	0.163	0.196	0.233	0.287
MMI2	30	0.109	0.143	0.177	0.218	0.262	0.327
MIX	30	0.105	0.122	0.140	0.162	0.187	0.224
MLE	30	0.107	0.144	0.191	0.254	0.335	0.477
ENT	30	0.104	0.116	0.129	0.142	0.156	0.175
MMD	50	0.084	0.095	0.108	0.123	0.139	0.163
MMI1	50	0.085	0.105	0.127	0.152	0.180	0.221
MMI2	50	0.085	0.109	0.134	0.163	0.195	0.240
MIX	50	0.082	0.096	0.110	0.128	0.148	0.177
MLE	50	0.082	0.103	0.128	0.159	0.194	0.247
ENT	50	0.083	0.095	0.106	0.120	0.135	0.155
MMD	75	0.068	0.079	0.090	0.104	0.120	0.142
MMI1	75	0.068	0.086	0.105	0.127	0.151	0.186
MMI2	75	0.068	0.088	0.109	0.134	0.160	0.197
MIX	75	0.067	0.079	0.092	0.108	0.125	0.151
MLE	75	0.067	0.083	0.103	0.127	0.154	0.195
ENT	75	0.067	0.077	0.087	0.099	0.112	0.130

TABLE 5.10

STANDARD ERROR OF SELECTED QUANTILES
(CASE - 2 : C.V. = 0.5 SKEW = 3.0)

METHOD	SAMPLE SIZE	RETURN PERIOD 10	25	50	100	200	500
MMD	10	0.272	0.320	0.342	0.356	0.361	0.361
MMI1	10	0.238	0.351	0.474	0.642	0.878	1.354
MMI2	10	0.229	0.392	0.625	1.040	1.825	4.220
MIX	10	0.232	0.287	0.331	0.376	0.422	0.486
MLE	10	0.220	0.267	0.304	0.343	0.386	0.451
ENT	10	0.237	0.354	0.477	0.641	0.861	1.278
MMD	20	0.182	0.228	0.257	0.282	0.303	0.325
MMI1	20	0.153	0.223	0.293	0.381	0.493	0.692
MMI2	20	0.152	0.240	0.343	0.489	0.699	1.141
MIX	20	0.155	0.200	0.237	0.277	0.320	0.380
MLE	20	0.157	0.236	0.313	0.407	0.519	0.703
ENT	20	0.153	0.220	0.284	0.362	0.456	0.612
MMD	30	0.155	0.199	0.229	0.256	0.281	0.310
MMI1	30	0.129	0.187	0.244	0.314	0.398	0.540
MMI2	30	0.129	0.198	0.273	0.371	0.500	0.738
MIX	30	0.132	0.174	0.209	0.248	0.289	0.349
MLE	30	0.133	0.200	0.264	0.340	0.427	0.564
ENT	30	0.130	0.183	0.233	0.291	0.358	0.464
MMD	50	0.122	0.165	0.196	0.225	0.253	0.286
MMI1	50	0.100	0.146	0.191	0.248	0.321	0.457
MMI2	50	0.100	0.151	0.205	0.277	0.373	0.568
MIX	50	0.103	0.141	0.173	0.210	0.250	0.307
MLE	50	0.103	0.153	0.200	0.254	0.314	0.406
ENT	50	0.100	0.139	0.173	0.211	0.253	0.315
MMD	75	0.101	0.140	0.169	0.197	0.225	0.260
MMI1	75	0.083	0.123	0.161	0.206	0.258	0.343
MMI2	75	0.083	0.127	0.169	0.221	0.283	0.389
MIX	75	0.086	0.120	0.151	0.185	0.222	0.276
MLE	75	0.084	0.125	0.161	0.201	0.246	0.311
ENT	75	0.083	0.118	0.149	0.184	0.221	0.276

TABLE 5.11

STANDARD ERROR OF SELECTED QUANTILES
(CASE - 3 : C.V. = 0.5 SKEW = 5.0)

		RETURN PERIOD					
METHOD	SAMPLE SIZE	10	25	50	100	200	500
MMD	10	0.323	0.377	0.396	0.402	0.397	0.381
MMI1	10	0.265	0.406	0.555	0.755	1.033	1.588
MMI2	10	0.245	0.443	0.724	1.222	2.160	5.045
MIX	10	0.266	0.332	0.379	0.423	0.465	0.518
MLE	10	0.242	0.280	0.303	0.323	0.342	0.366
ENT	10	0.262	0.406	0.567	0.802	1.154	1.922
MMD	20	0.221	0.276	0.307	0.330	0.346	0.359
MMI1	20	0.173	0.257	0.344	0.454	0.597	0.856
MMI2	20	0.167	0.271	0.398	0.585	0.863	1.471
MIX	20	0.181	0.234	0.275	0.317	0.359	0.417
MLE	20	0.170	0.228	0.283	0.348	0.424	0.543
ENT	20	0.173	0.237	0.296	0.364	0.444	0.571
MMD	30	0.182	0.237	0.271	0.299	0.322	0.344
MMI1	30	0.137	0.206	0.277	0.367	0.480	0.680
MMI2	30	0.134	0.215	0.307	0.436	0.616	0.974
MIX	30	0.146	0.195	0.234	0.276	0.319	0.378
MLE	30	0.139	0.196	0.249	0.310	0.379	0.482
ENT	30	0.137	0.188	0.233	0.285	0.342	0.427
MMD	50	0.148	0.202	0.238	0.269	0.295	0.323
MMI1	50	0.107	0.165	0.225	0.300	0.397	0.575
MMI2	50	0.105	0.169	0.240	0.336	0.466	0.726
MIX	50	0.116	0.161	0.199	0.239	0.282	0.341
MLE	50	0.110	0.161	0.206	0.257	0.312	0.394
ENT	50	0.106	0.146	0.181	0.220	0.263	0.324
MMD	75	0.122	0.170	0.204	0.235	0.263	0.296
MMI1	75	0.088	0.135	0.182	0.239	0.309	0.428
MMI2	75	0.087	0.138	0.191	0.258	0.343	0.494
MIX	75	0.094	0.134	0.168	0.205	0.245	0.302
MLE	75	0.090	0.131	0.167	0.207	0.250	0.313
ENT	75	0.088	0.124	0.156	0.191	0.229	0.284

TABLE 5.12

STANDARD ERROR OF SELECTED QUANTILES
(CASE - 4 : C.V. = 0.3 SKEW = 3.0)

METHOD	SAMPLE SIZE	RETURN PERIOD 10	25	50	100	200	500
MMD	10	0.189	0.229	0.254	0.274	0.290	0.306
MMI1	10	0.175	0.256	0.332	0.424	0.537	0.728
MMI2	10	0.164	0.272	0.407	0.608	0.916	1.614
MIX	10	0.171	0.220	0.257	0.295	0.334	0.388
MLE	10	0.164	0.194	0.214	0.233	0.252	0.277
ENT	10	0.173	0.239	0.297	0.365	0.445	0.575
MMD	20	0.133	0.170	0.196	0.220	0.242	0.268
MMI1	20	0.118	0.174	0.227	0.289	0.363	0.482
MMI2	20	0.114	0.181	0.255	0.353	0.481	0.719
MIX	20	0.119	0.159	0.192	0.227	0.263	0.314
MLE	20	0.115	0.149	0.177	0.208	0.241	0.287
ENT	20	0.117	0.158	0.192	0.231	0.272	0.334
MMD	30	0.111	0.148	0.175	0.202	0.227	0.260
MMI1	30	0.096	0.143	0.190	0.246	0.314	0.427
MMI2	30	0.093	0.147	0.206	0.283	0.383	0.567
MIX	30	0.098	0.135	0.168	0.203	0.240	0.294
MLE	30	0.096	0.132	0.164	0.198	0.234	0.285
ENT	30	0.095	0.127	0.154	0.184	0.216	0.261
MMD	50	0.085	0.118	0.144	0.171	0.196	0.230
MMI1	50	0.072	0.111	0.150	0.196	0.250	0.338
MMI2	50	0.071	0.113	0.158	0.214	0.284	0.402
MIX	50	0.074	0.107	0.136	0.169	0.204	0.255
MLE	50	0.072	0.101	0.128	0.157	0.188	0.230
ENT	50	0.072	0.096	0.119	0.143	0.169	0.206
MMD	75	0.071	0.100	0.124	0.148	0.173	0.206
MMI1	75	0.060	0.092	0.124	0.161	0.204	0.271
MMI2	75	0.060	0.094	0.129	0.172	0.222	0.304
MIX	75	0.062	0.091	0.117	0.146	0.179	0.225
MLE	75	0.060	0.085	0.108	0.133	0.159	0.195
ENT	75	0.060	0.083	0.104	0.127	0.152	0.186

TABLE 5.13

STANDARD ERROR OF SELECTED QUANTILES
(CASE - 5 : C.V. = 0.7 SKEW = 3.0)

METHOD	SAMPLE SIZE	RETURN PERIOD 10	25	50	100	200	500
MMD	10	0.349	0.392	0.404	0.404	0.395	0.375
MMI1	10	0.305	0.435	0.586	0.805	1.127	1.818
MMI2	10	0.301	0.499	0.801	1.371	2.522	6.326
MIX	10	0.291	0.345	0.387	0.431	0.476	0.538
MLE	10	0.282	0.342	0.393	0.450	0.516	0.619
ENT	10	0.315	0.487	0.693	1.006	1.490	2.587
MMD	20	0.237	0.284	0.312	0.334	0.350	0.364
MMI1	20	0.205	0.279	0.352	0.440	0.548	0.730
MMI2	20	0.207	0.304	0.411	0.555	0.748	1.119
MIX	20	0.202	0.248	0.286	0.326	0.369	0.430
MLE	20	0.208	0.315	0.434	0.595	0.818	1.264
ENT	20	0.210	0.293	0.375	0.475	0.600	0.812
MMD	30	0.187	0.232	0.263	0.290	0.313	0.338
MMI1	30	0.159	0.216	0.274	0.342	0.424	0.556
MMI2	30	0.160	0.230	0.306	0.401	0.522	0.734
MIX	30	0.159	0.199	0.235	0.274	0.317	0.378
MLE	30	0.164	0.252	0.347	0.470	0.630	0.919
ENT	30	0.162	0.221	0.278	0.347	0.429	0.561
MMD	50	0.154	0.204	0.240	0.272	0.300	0.332
MMI1	50	0.127	0.175	0.224	0.288	0.373	0.539
MMI2	50	0.128	0.182	0.241	0.321	0.434	0.675
MIX	50	0.130	0.168	0.204	0.243	0.286	0.350
MLE	50	0.128	0.184	0.242	0.312	0.395	0.528
ENT	50	0.127	0.170	0.212	0.262	0.323	0.421
MMD	75	0.124	0.167	0.201	0.234	0.265	0.303
MMI1	75	0.103	0.141	0.179	0.226	0.282	0.377
MMI2	75	0.103	0.145	0.188	0.242	0.309	0.427
MIX	75	0.105	0.139	0.171	0.208	0.248	0.308
MLE	75	0.104	0.145	0.187	0.235	0.289	0.372
ENT	75	0.103	0.138	0.171	0.209	0.253	0.322

TABLE 5.14

ROOT MEAN SQUARE ERROR OF SELECTED QUANTILES
(CASE - 1 : C.V. = 0.5 SKEW = 1.0)

METHOD	SAMPLE SIZE	RETURN PERIOD 10	25	50	100	200	500
MMD	10	0.185	0.214	0.239	0.265	0.290	0.320
MMI1	10	0.190	0.239	0.294	0.364	0.451	0.596
MMI2	10	0.195	0.269	0.357	0.482	0.658	1.019
MIX	10	0.180	0.204	0.228	0.258	0.294	0.351
MLE	10	0.195	0.256	0.317	0.391	0.477	0.615
ENT	10	0.185	0.222	0.254	0.287	0.321	0.365
MMD	20	0.131	0.152	0.172	0.194	0.217	0.248
MMI1	20	0.133	0.165	0.200	0.242	0.292	0.367
MMI2	20	0.134	0.178	0.223	0.277	0.340	0.441
MIX	20	0.129	0.147	0.167	0.192	0.221	0.265
MLE	20	0.135	0.184	0.246	0.328	0.433	0.623
ENT	20	0.129	0.152	0.172	0.194	0.215	0.242
MMD	30	0.107	0.125	0.143	0.163	0.184	0.212
MMI1	30	0.108	0.135	0.164	0.199	0.238	0.297
MMI2	30	0.109	0.143	0.178	0.218	0.262	0.329
MIX	30	0.105	0.122	0.140	0.162	0.187	0.225
MLE	30	0.107	0.144	0.191	0.254	0.336	0.481
ENT	30	0.106	0.124	0.141	0.159	0.177	0.201
MMD	50	0.084	0.097	0.111	0.127	0.145	0.170
MMI1	50	0.085	0.105	0.127	0.154	0.184	0.227
MMI2	50	0.085	0.109	0.134	0.163	0.195	0.240
MIX	50	0.083	0.096	0.110	0.128	0.148	0.177
MLE	50	0.083	0.103	0.128	0.159	0.194	0.247
ENT	50	0.083	0.098	0.112	0.127	0.144	0.166
MMD	75	0.068	0.079	0.092	0.107	0.123	0.146
MMI1	75	0.068	0.086	0.105	0.128	0.153	0.190
MMI2	75	0.068	0.088	0.109	0.134	0.160	0.197
MIX	75	0.067	0.079	0.092	0.107	0.125	0.151
MLE	75	0.067	0.083	0.103	0.127	0.154	0.195
ENT	75	0.067	0.078	0.090	0.103	0.117	0.136

TABLE 5.15

ROOT MEAN SQUARE ERROR OF SELECTED QUANTILES
(CASE - 2 : C.V. = 0.5 SKEW = 3.0)

METHOD	SAMPLE SIZE	RETURN PERIOD 10	25	50	100	200	500
MMD	10	0.272	0.329	0.368	0.404	0.438	0.479
MMI1	10	0.237	0.351	0.474	0.643	0.880	1.361
MMI2	10	0.229	0.392	0.632	1.061	1.868	4.306
MIX	10	0.233	0.292	0.339	0.387	0.436	0.502
MLE	10	0.222	0.277	0.322	0.370	0.420	0.492
ENT	10	0.238	0.359	0.491	0.666	0.901	1.342
MMD	20	0.182	0.233	0.273	0.313	0.351	0.399
MMI1	20	0.153	0.223	0.293	0.381	0.493	0.694
MMI2	20	0.152	0.240	0.345	0.494	0.711	1.166
MIX	20	0.156	0.205	0.245	0.288	0.333	0.396
MLE	20	0.157	0.236	0.313	0.407	0.519	0.706
ENT	20	0.153	0.221	0.289	0.371	0.472	0.639
MMD	30	0.155	0.202	0.240	0.278	0.316	0.365
MMI1	30	0.129	0.187	0.244	0.314	0.398	0.540
MMI2	30	0.129	0.198	0.274	0.374	0.507	0.752
MIX	30	0.132	0.177	0.215	0.256	0.300	0.361
MLE	30	0.133	0.200	0.265	0.342	0.432	0.575
ENT	30	0.130	0.184	0.236	0.296	0.367	0.478
MMD	50	0.123	0.166	0.201	0.238	0.274	0.321
MMI1	50	0.100	0.146	0.191	0.248	0.321	0.457
MMI2	50	0.100	0.151	0.206	0.279	0.377	0.574
MIX	50	0.103	0.142	0.177	0.215	0.256	0.314
MLE	50	0.103	0.154	0.202	0.258	0.321	0.419
ENT	50	0.100	0.140	0.175	0.214	0.259	0.324
MMD	75	0.103	0.140	0.172	0.205	0.240	0.284
MMI1	75	0.083	0.123	0.161	0.206	0.258	0.344
MMI2	75	0.083	0.127	0.170	0.222	0.286	0.394
MIX	75	0.086	0.121	0.153	0.188	0.226	0.280
MLE	75	0.084	0.125	0.162	0.204	0.250	0.319
ENT	75	0.083	0.119	0.150	0.186	0.225	0.282

TABLE 5.16

ROOT MEAN SQUARE ERROR OF SELECTED QUANTILES
(CASE - 3 : C.V. = 0.5 SKEW = 5.0)

METHOD	SAMPLE SIZE	RETURN PERIOD 10	25	50	100	200	500
MMD	10	0.323	0.384	0.422	0.455	0.486	0.526
MMI1	10	0.265	0.406	0.554	0.755	1.034	1.593
MMI2	10	0.246	0.443	0.731	1.245	2.210	5.153
MIX	10	0.267	0.338	0.390	0.441	0.490	0.553
MLE	10	0.244	0.300	0.346	0.392	0.438	0.498
ENT	10	0.262	0.406	0.569	0.807	1.162	1.938
MMD	20	0.222	0.280	0.323	0.364	0.403	0.452
MMI1	20	0.173	0.258	0.344	0.454	0.597	0.857
MMI2	20	0.168	0.271	0.400	0.592	0.879	1.506
MIX	20	0.181	0.239	0.286	0.334	0.383	0.448
MLE	20	0.170	0.232	0.290	0.357	0.434	0.553
ENT	20	0.173	0.237	0.296	0.364	0.444	0.571
MMD	30	0.183	0.238	0.280	0.322	0.363	0.413
MMI1	30	0.137	0.207	0.277	0.366	0.480	0.681
MMI2	30	0.134	0.215	0.309	0.441	0.626	0.997
MIX	30	0.146	0.198	0.242	0.289	0.337	0.402
MLE	30	0.138	0.196	0.250	0.310	0.378	0.482
ENT	30	0.136	0.188	0.233	0.285	0.342	0.427
MMD	50	0.151	0.202	0.243	0.283	0.321	0.370
MMI1	50	0.107	0.165	0.224	0.300	0.397	0.576
MMI2	50	0.105	0.169	0.241	0.338	0.471	0.737
MIX	50	0.116	0.163	0.204	0.248	0.294	0.357
MLE	50	0.110	0.161	0.207	0.259	0.316	0.400
ENT	50	0.106	0.146	0.182	0.221	0.264	0.325
MMD	75	0.127	0.170	0.207	0.244	0.282	0.330
MMI1	75	0.088	0.135	0.182	0.239	0.310	0.429
MMI2	75	0.087	0.138	0.191	0.260	0.347	0.502
MIX	75	0.094	0.135	0.171	0.211	0.253	0.313
MLE	75	0.090	0.132	0.169	0.210	0.255	0.322
ENT	75	0.088	0.124	0.156	0.191	0.230	0.284

TABLE 5.17

ROOT MEAN SQUARE ERROR OF SELECTED QUANTILES
(CASE - 4 : C.V. = 0.3 SKEW = 3.0)

METHOD	SAMPLE SIZE	RETURN PERIOD 10	25	50	100	200	500
MMD	10	0.189	0.239	0.277	0.315	0.351	0.398
MMI1	10	0.175	0.257	0.333	0.425	0.537	0.728
MMI2	10	0.165	0.272	0.408	0.615	0.934	1.655
MIX	10	0.172	0.223	0.263	0.303	0.344	0.399
MLE	10	0.166	0.211	0.249	0.289	0.330	0.385
ENT	10	0.173	0.239	0.297	0.365	0.445	0.574
MMD	20	0.133	0.176	0.212	0.248	0.284	0.330
MMI1	20	0.118	0.175	0.228	0.291	0.364	0.483
MMI2	20	0.115	0.181	0.255	0.354	0.486	0.731
MIX	20	0.120	0.163	0.199	0.236	0.274	0.328
MLE	20	0.116	0.156	0.191	0.230	0.269	0.323
ENT	20	0.118	0.159	0.195	0.233	0.276	0.338
MMD	30	0.111	0.151	0.185	0.220	0.255	0.302
MMI1	30	0.096	0.144	0.190	0.247	0.314	0.427
MMI2	30	0.094	0.147	0.206	0.284	0.386	0.575
MIX	30	0.098	0.138	0.172	0.209	0.248	0.303
MLE	30	0.096	0.134	0.168	0.205	0.244	0.297
ENT	30	0.095	0.128	0.157	0.188	0.221	0.267
MMD	50	0.085	0.120	0.150	0.181	0.213	0.256
MMI1	50	0.072	0.111	0.150	0.196	0.250	0.337
MMI2	50	0.071	0.113	0.158	0.215	0.286	0.407
MIX	50	0.074	0.108	0.139	0.172	0.209	0.260
MLE	50	0.072	0.102	0.130	0.160	0.192	0.235
ENT	50	0.072	0.098	0.122	0.148	0.176	0.215
MMD	75	0.071	0.100	0.127	0.155	0.185	0.225
MMI1	75	0.060	0.092	0.124	0.161	0.204	0.271
MMI2	75	0.060	0.094	0.129	0.172	0.224	0.307
MIX	75	0.062	0.091	0.118	0.148	0.181	0.228
MLE	75	0.060	0.085	0.108	0.133	0.160	0.196
ENT	75	0.060	0.084	0.107	0.132	0.159	0.195

TABLE 5.18

ROOT MEAN SQUARE ERROR OF SELECTED QUANTILES
(CASE - 5 : C.V. = 0.7 SKEW = 3.0)

METHOD	SAMPLE SIZE	RETURN PERIOD 10	25	50	100	200	500
MMD	10	0.349	0.402	0.434	0.463	0.491	0.526
MMI1	10	0.304	0.435	0.588	0.811	1.140	1.844
MMI2	10	0.302	0.500	0.810	1.399	2.581	6.447
MIX	10	0.293	0.351	0.398	0.446	0.494	0.559
MLE	10	0.283	0.346	0.398	0.455	0.520	0.622
ENT	10	0.317	0.513	0.752	1.106	1.636	2.800
MMD	20	0.237	0.290	0.331	0.371	0.409	0.455
MMI1	20	0.205	0.279	0.352	0.441	0.550	0.735
MMI2	20	0.207	0.304	0.413	0.559	0.759	1.145
MIX	20	0.204	0.254	0.296	0.341	0.387	0.451
MLE	20	0.209	0.315	0.434	0.597	0.825	1.282
ENT	20	0.211	0.305	0.402	0.524	0.674	0.926
MMD	30	0.187	0.235	0.275	0.315	0.355	0.404
MMI1	30	0.159	0.216	0.273	0.342	0.425	0.559
MMI2	30	0.160	0.230	0.307	0.404	0.528	0.748
MIX	30	0.159	0.204	0.243	0.285	0.330	0.394
MLE	30	0.164	0.252	0.349	0.475	0.639	0.940
ENT	30	0.163	0.230	0.299	0.382	0.482	0.642
MMD	50	0.156	0.204	0.245	0.285	0.323	0.370
MMI1	50	0.127	0.175	0.224	0.289	0.374	0.541
MMI2	50	0.128	0.182	0.242	0.323	0.437	0.681
MIX	50	0.130	0.170	0.208	0.249	0.293	0.358
MLE	50	0.128	0.184	0.243	0.314	0.399	0.537
ENT	50	0.127	0.177	0.227	0.287	0.358	0.473
MMD	75	0.126	0.167	0.204	0.241	0.279	0.327
MMI1	75	0.103	0.141	0.180	0.226	0.283	0.379
MMI2	75	0.103	0.145	0.189	0.243	0.311	0.431
MIX	75	0.105	0.140	0.173	0.211	0.252	0.312
MLE	75	0.103	0.145	0.187	0.235	0.291	0.376
ENT	75	0.104	0.143	0.181	0.226	0.278	0.358

Our objective here is to identify a robust estimator, based on Tables 5.4-5.18. Kuczera (1982a, 1982b) defined a robust estimator as one that is resistant and efficient over a wide range of population fluctuations. If an estimator performs steadily without undue deterioration in RMSE and BIAS, it can be expected to perform better than other competitive estimators under population conditions different from those which the conclusions are based. Two criteria for identifying a resistant estimator (Matalas and Fiering, 1977; Kuczera, 1982b) are mini-max and minimum average RMSE. According to the mini-max criteria, the preferred estimator is the one whose maximum RMSE for the five population cases is minimal. The minimum average criterion is to select the estimator whose RMSE average over the five cases is minimal. Table 5.19 reports the maximum and average RMSE for each estimator for selected sample size and return period.

The MIX estimator is superior on the basis of the minimum-average RMSE criteria, and comparable to MMD on the basis of mini-max RMSE criteria. Hence, MIX is expected to be the most resistant estimator. Nevertheless, MMD performs comparably. The method proposed by the U.S. Water Resources Council (MMI1) performs poorly, as do MLE and ENT. Taking into consideration the poor performance of MLE and ENT and the tremendous amount of CPU time required by the extensive search routines, there should be no doubt that MLE and ENT are inferior methods for LP3 distribution.

To see the performance of estimators in terms of BIAS, Table 5.20, similar to Table 5.19, was prepared. Interestingly enough, the superior RMSE performance of MIX in comparison to MMD is not deteriorated by BIAS. MIX yields considerably less BIAS then MMD and is clearly superior to MMD in terms of both mini-max BIAS and minimum average BIAS criteria.

5.7 Conclusions

MIX and MDD were found to be clearly superior to other methods in terms of RMSE and BIAS. The method advocated by U.S. Water Resources Council (MMI1) faired poorly. It seems that its continued recommendation for U.S. Agencies is unwarranted. MLE typically required two order of magnitude high CPU time than other methods and faired poorly in performance. Based on the investigations of this study, MIX holds an edge over MMD in performance. However, the results are close for the two methods.

The results of this study indicate that, when considering the fitting of the LP3 model, MIX or MMD should be used as the methods of extimation, as they are clearly superior to the method advocated by the U.S. Water Resources Council. A simple procedure proposed in this work could be used to estimate parameters of LP3 by MIX. This procedure provides a straight-forward solution, and thus obviates the need for iterative procedure.

Table 5.19 : Summary of RMSE performance of 200 - Year Quantile Estimators.

	Maximum RMSE		Average RMSE	
Estimator	n = 10	n = 30	n = 10	n = 30
MMD	0.491	0.363	0.411	0.295
MMI1	1.140	0.480	0.808	0.371
MMI2	2.581	0.626	1.650	0.462
MIX	0.494	0.337	0.412	0.280
MLE	0.520	0.639	0.437	0.406
ENT	1.636	0.482	0.893	0.318

Table 5.20 : Summary of BIAS performance of 200 - Year Quantile Estimators.

	Maximum BIAS		Average ABSOLUTE* BIAS	
Estimator	n = 10	n = 30	n = 10	n = 30
MMD	-0.291	-0.167	-0.235	-0.133
MMI1	0.172	0.032	0.086	0.023
MMI2	0.554	0.116	0.341	0.069
MIX	-0.156	-0.109	-0.099	-0.070
MLE	-0.274	0.113	-0.178	0.057
ENT	0.678	0.221	0.256	0.088

* the sign shows the dominant tendency of the estimator

Chapter 6

TWO COMPONENT EXTREME VALUE (TCEV) DISTRIBUTION

6.1 Introduction

The random variable x is defined to have a two-component extreme value (TCEV) distribution if its probability density function (pdf) is given by

$$f(x) = [\frac{\Lambda_1}{\theta_1} \exp(-x/\theta_1) + \frac{\Lambda_2}{\theta_2} \exp(-x/\theta_2)] \exp[-\Lambda_1 \exp(-x/\theta_1) - \Lambda_2 \exp(-x/\theta_2)]; \quad x > 0 \qquad (6.1)$$

$$= \exp(-\Lambda_1 - \Lambda_2); \quad x = 0 \qquad (6.2)$$

where $\Lambda_1 > 0$, $\Lambda_2 \geq 0$, $\theta_2 \geq \theta_1 > 0$ are parameters. Its cumulative density function (cdf) is

$$F(x) = \exp[-\Lambda_1 \exp(-x/\theta_1) - \Lambda_2 \exp(-x/\theta_2)]; \quad x \geq 0 \qquad (6.3)$$

This cdf has been shown (Versace, et al., 1982; Rossi, et al., 1984) to represent the distribution function of the annual maximum x of a non-negative random variate z whose number of occurrences, k, in a year is a random variate when the following hypotheses hold: (1) z is an independent, identically distributed (iid) variable with probability density function defined by a mixture of two exponential distributions; (2) k is an iid Poisson distributed variate; and (3) z and k are not dependent upon each other.

The two components of the distribution of both z and x are usually referred to as basic component (subscripts of parameters = 1) and outlying component (subscripts of parameters = 2). Theoretical properties of the TCEV distribution have been widely investigated (Rossi, et al., 1984; Beran, et al., 1986; Rossi, et al., 1986). Briefly summarized, the TCEV distribution permits a reasonable interpretation of the physical phenomenon which generates floods and is able to account for most of the characteristics of the real world flood data, important among them being the large variability of the sample skewness coefficient which mostly gives rise to the poor performance of other commonly used flood frequency distributions. The TCEV distribution also offers a practical approach to regional flood frequency estimation.

Thus far, only two parameter estimation methods have been proposed for fitting the TCEV distribution to annual flood series. Canfield (1979) suggested a least squares technique, while Rossi, et al. (1984) presented a procedure based on the maximum likelihood estimation (MLE) method. The latter was further investigated and a regional estimation algorithm based on it was developed (Fiorentino, et al. 1985). Small sample properties of the site-specific and regionalized TCEV-MLE procedure were assessed by Fiorentino and Gabriele (1985), and Arnell and Gabriele (1986). In particular, the latter compared the regionalized TCEV-MLE algorithm with other regional estimators. Although various features of the TCEV-MLE method exhibited a competitive performance, an improvement of the site-specific estimators was suggested by Fiorentino and Gabriele (1985). Furthermore, Fiorentino, et al. (1986) noted that regional estimates of some parameters could be still improved. In light of this discussion, it is desirable to investigate further into the properties and estimation procedures of the TCEV distribution.

The objective of this chapter is to derive this distribution using the principle of maximum entropy (POME). The derivation sheds more light on the nature of the distribution and provides an alternative method for estimation of its parameters. The TCEV-POME estimation procedure is shown to be also suitable for regionalized use. Its performance is assessed by using the Monte Carlo technique and the results are compared with those of the MLE method.

6.2 The Principle of Maximum Entropy (POME)

Entropy is defined as a measure of uncertainty. The entropy function H(f) of a distribution, with pdf f(x), of a random variable X quantifies this uncertainty. When X is a continuous variable and assumes values in a domain c, H(f) is given as

$$H(f) = - \int_c f(x) \ln f(x)\, dx \tag{6.4}$$

This form was first used by Jaynes (1968) and represents an extension for the continuous case of the entropy function applied in communication theory by Shannon and Weaver (1949).

The principle of maximum entropy (POME) was formulated by Jaynes (1957, 1961) and states that "the minimally prejudiced assignment of probabilities is that which maximizes the entropy subject to the given information." Mathematically, "the given information" is quantified by some linearly independent constraints to be imposed while maximizing equation (6.4). It has been argued by Jaynes that the probability distribution resulting from POME is unique in the sense that it does not assume any more information than that quantified by the linearly independent constraints, subject to which the entropy is maximized. In other words, this distribution is most uncertain (or minimally prejudiced) with respect to the missing information. POME has been applied in several hydrologic fields. Sonuga (1972, 1976) was probably the first who applied POME in hydrologic frequency analysis. Jowitt (1979), and Singh and Singh (1985) successfully exploited it to derive the extreme value type I and the Pearson type III distributions respectively. Furthermore, the function H(f) has been given for a number of probability distributions by Verdugo Lazo and Rathie (1978), and Singh, et al. (1985).

In all of these cases, POME has been shown to uniquely specify the amount of information in terms of constraints required to derive a given distribution. The parameter estimation methods based on POME almost always result in an easy, practical solution, and their performance is generally found to be comparable with that of MLE method.

Maximizing H(f) is conceptually easy when constraints are formulated as population means of specified functions of x. In fact, in such a case the required pdf f(x) has been shown (Singh, et al., 1986) to result in

$$f(x) = \exp[- a_0 - \sum_{j=1}^{r} a_j \, y_j(x)] \tag{6.5}$$

where $y_j(x)$ are some functions whose population means are the selected constraints, and a_j ($j = 0, 1, \ldots, r$) are Lagrange multipliers. Since, by definition,

$$\int_c f(x)\, dx = 1 \tag{6.6}$$

for equation (5), the following relationships hold:

$$\begin{aligned} \frac{\partial a_0}{\partial a_j} &= - E[y_j(x)] \\ \frac{\partial^2 a_0}{\partial a_j^2} &= \mathrm{var}[y_j(x)] \qquad j = 1, 2, \ldots, r;\quad i = 1, 2, \ldots, r;\quad i \neq j \\ \frac{\partial^2 a_0}{\partial a_i \partial a_j} &= \mathrm{cov}[y_i(x), y_j(x)] \end{aligned} \tag{6.7}$$

where E[•], var[•] and cov[•] indicate population mean, variance and covariance respectively.

Equations (6.5), (6.6) and (6.7) enable us to derive Lagrange multipliers as functions of apriori information provided by constraints. Finally, for equations (6.4) and (6.5), the entropy function can be written as

$$H(f) = a_0 + \sum_{j=1}^{r} a_j \, E[y_j(x)] \tag{6.8}$$

6.3 Derivation of the TCEV Distribution

6.3.1 Specification of Constraints

The TCEV distribution has four parameters, hence five constraints need to be given for its derivation. Let the constraints have the following form:

$$\int_{-\infty}^{\infty} f(x) \, dx = 1 \tag{6.9}$$

$$\int_{-\infty}^{\infty} x \, f(x) \, dx = E[x] \tag{6.10}$$

$$\int_{-\infty}^{\infty} \exp(- x/\theta_1) \, f(x) \, dx = E[\exp(- x/\theta_1)] \tag{6.11}$$

$$\int_{-\infty}^{\infty} \exp(- x/\theta_2) \, f(x) \, dx = E[\exp(- x/\theta_2)] \tag{6.12}$$

$$\int_{-\infty}^{\infty} \ln(1 + \frac{(\Lambda_2/\theta_2) \exp(- x/\theta_2)}{(\Lambda_1/\theta_1) \exp(- x/\theta_1)}) \, f(x) \, dx$$

$$= E[\ln(1 + \frac{(\Lambda_2/\theta_2) \exp(- x/\theta_2)}{(\Lambda_1/\theta_1) \exp(- x/\theta_1)})] \tag{6.13}$$

The constraints are to be evaluated from data, directly or indirectly, and will suffice to derive the TCEV distribution through maximization of entropy. It may be noted that the first three constraints are the same as those used for deriving EV1 distribution (Jowitt, 1979; Singh, et al., 1985), while the fourth constraint, which is analogous to the third one, provides information on the outlying component. The final constraint combines the information between the basic and outlying components.

6.3.2 Derivation of the Distribution

The pdf in equation (6.5) is determined by POME according to equations (6.9) - (6.13), and takes the form:

$$f(x) = \exp\{- a_0 - a_1 x - a_2 \exp(-x/\theta_1) - a_3 \exp(-x/\theta_2) - a_4 \ln[1 + \frac{\Lambda_2 \theta_1}{\theta_2 \Lambda_1} \exp[- x(\frac{1}{\theta_2} - \frac{1}{\theta_1})]]\} \qquad (6.14)$$

Inserting equation (6.14) in equation (6.9),

$$\exp(a_0) = \int_{-\infty}^{\infty} \exp[- a_1 x - a_2 \exp(- x/\theta_1) - a_3 \exp(-x/\theta_2)] [1 + \frac{\Lambda_2 \theta_1}{\theta_2 \Lambda_1} \exp(- x(\frac{1}{\theta_2} - \frac{1}{\theta_1}))]^{-a_4} dx \qquad (6.15)$$

Let $z = \Lambda_1 \exp(-x/\theta_1)$, $\theta = \theta_2/\theta_1$, $\Lambda = \Lambda_2/(\Lambda_1^{1/\theta})$. After simple manipulation, the zeroth Lagrange multiplier is:

$$a_0 = \ln \theta_1 - a_1 \theta_1 \ln \Lambda_1 + \ln \int_0^{\infty} z^{a_1\theta_1 - 1} \exp(- a_2 z/\Lambda_1) \exp[- a_3 z^{(1/\theta)} \Lambda_1^{(-1/\theta)}] [1 + \frac{\Lambda}{\theta} z^{((1/\theta)-1)}]^{-a_4} dz \qquad (6.16)$$

Inserting equation (6.16) in equation (6.14) we get,

$$f(x) = \frac{1}{\theta_1} \Lambda_1^{a_1\theta_1} \exp[- a_1 x - a_2 \exp(- x/\theta_1) - a_3 \exp(-x/\theta_2)] [1 + \frac{(\Lambda_2/\theta_2) \exp(- x/\theta_2)}{(\Lambda_1/\theta_1) \exp(- x/\theta_1)}]^{-a_4} / I_0 \qquad (6.17)$$

where I_0 is the integral in equation (6.16). When

$$a_1 = 1/\theta_1, \; a_2 = \Lambda_1, \; a_3 = \Lambda_2, \text{ and } a_4 = -1 \qquad (6.18)$$

integral I_0 becomes unity and equation (6.17) becomes

$$f(x) = [\frac{\Lambda_1}{\theta_1} \exp(-x/\theta_1) + \frac{\Lambda_2}{\theta_2} \exp(-x/\theta_2)] \exp[-\Lambda_1 \exp(-x/\theta_1)$$
$$- \Lambda_2 \exp(-x/\theta_2)] \qquad (6.19)$$

which, according to the assumption that equation (6.1) also holds for negative values of x, is the pdf of the TCEV distribution. This distribution is consistent with respect to the information given by equations (6.9) - (6.13).

6.3.3 Relation between Constraints and Parameters

The relationships between the parameters of TCEV distribution and the constraints are specified by exploiting equation (6.7) as follows. Partially differentiating a_0 w.r.t. a_1, a_2, a_3 and a_4 respectively, and taking into account equation (6.18),

$$\frac{\partial a_0}{\partial a_1} = - E[x] = - \theta_1 \ln \Lambda_1 + \theta_1 \int_0^\infty \ln y \exp(- y - \Lambda y^{(1/\theta)})$$
$$(1 + \frac{\Lambda}{\theta} y^{((1/\theta)-1)}) \, dy \qquad (6.20)$$

$$\frac{\partial a_0}{\partial a_2} = - E[\exp(- x/\theta_1)] = - \frac{1}{\Lambda_1} \int_0^\infty y \exp(- y - \Lambda y^{(1/\theta)})$$
$$(1 + \frac{\Lambda}{\theta} y^{((1/\theta)-1)}) \, dy \qquad (6.21)$$

$$\frac{\partial a_0}{\partial a_3} = - E[\exp(- x/\theta_2)] = - \frac{1}{\Lambda_1^{1/\theta}} \int_0^\infty y^{(1/\theta)} \exp(- y - \Lambda y^{(1/\theta)})$$
$$(1 + \frac{\Lambda}{\theta} y^{((1/\theta)-1)}) \, dy \qquad (6.22)$$

$$\frac{\partial a_0}{\partial a_4} = - E[\ln(1 + \frac{(\Lambda_2/\theta_2) \exp(- x/\theta_2)}{(\Lambda_1/\theta_1) \exp(- x/\theta_1)}]$$
$$= - \int_0^\infty \ln(1 + \frac{\Lambda}{\theta} y^{((1/\theta)-1)}) \exp(- y - \Lambda y^{1/\theta})$$
$$(1 + \frac{\Lambda}{\theta} y^{((1/\theta)-1)}) \, dy \qquad (6.23)$$

Solving integrals in equations (6.20) to (6.22) provides,

$$E[x] = \theta_1 \ln \Lambda_1 + \theta_1 \gamma - \theta_1 \sum_{j=1}^{\infty} \frac{(-1)^j \Lambda^j}{j!} \Gamma(j/\theta) \qquad (6.24)$$

$$E[\exp(-x/\theta_1)] = \frac{1}{\Lambda_1} [1 + \frac{1}{\theta} \sum_{j=1}^{\infty} \frac{(-1)^j \Lambda^j}{(j-1)!} \Gamma(j/\theta)] \qquad (6.25)$$

$$E[\exp(-x/\theta_2)] = -\frac{1}{\theta\Lambda_2} \sum_{j=1}^{\infty} \frac{(-1)^j \Lambda^j}{(j-1)!} \Gamma(j/\theta) \qquad (6.26)$$

where $\gamma = 0.5772$ is the Euler constant and $\Gamma(\cdot)$ is the gamma function. θ and Λ are dimensionless parameters, already defined in terms of the four parameters of the TCEV distribution. The integral in equation (6.23) cannot be solved explicitly. But, for $\theta > 1.5$, it is closely approximated by the following function:

$$\int_0^\infty \ln(1 + \frac{\Lambda}{\theta} y^{((1/\theta)-1)}) \exp(-y - \Lambda y^{1/\theta}) (1 + \frac{\Lambda}{\theta} h^{((1/\theta)-1)}) dy$$

$$= 0.1 \exp(-1)(3 + \theta)^{2.059} \Lambda^{\ln 3 - 2(5.5^{-\theta})} \qquad (27)$$

The goodness of this approximation is shown in Figure 6.1. The curves approximating the integral in equation (6.23) have not been plotted for $\theta < 1.5$ to avoid any confusion at the left-bottom where they tend to overlap with each other. Moreover, the goodness of the approximation deteriorates in this range. Thus final constraint can be related to the parameters by

$$E[\ln(1 + \frac{(\Lambda_2/\theta_2) \exp(-x/\theta_2)}{(\Lambda_1/\theta_1) \exp(-x/\theta_1)})] = 0.1 \exp(-1)(3 + \theta)^{2.059} \Lambda^{\ln 3 - 2(5.5)^{-\theta}} \qquad (6.28)$$

Equations (6.24) - (6.26) and (6.28) show that constraints are related to the moments or moment-ratios of the distribution. In fact, besides the obvious case of the constraint $E[y_1(x)]$ representing the population mean of x, it is clear that $E[y_4(x)]$ depends on the dimensionless parameters θ and Λ only, while both $E[y_3(x)]$ and $E[y_2(x)]$

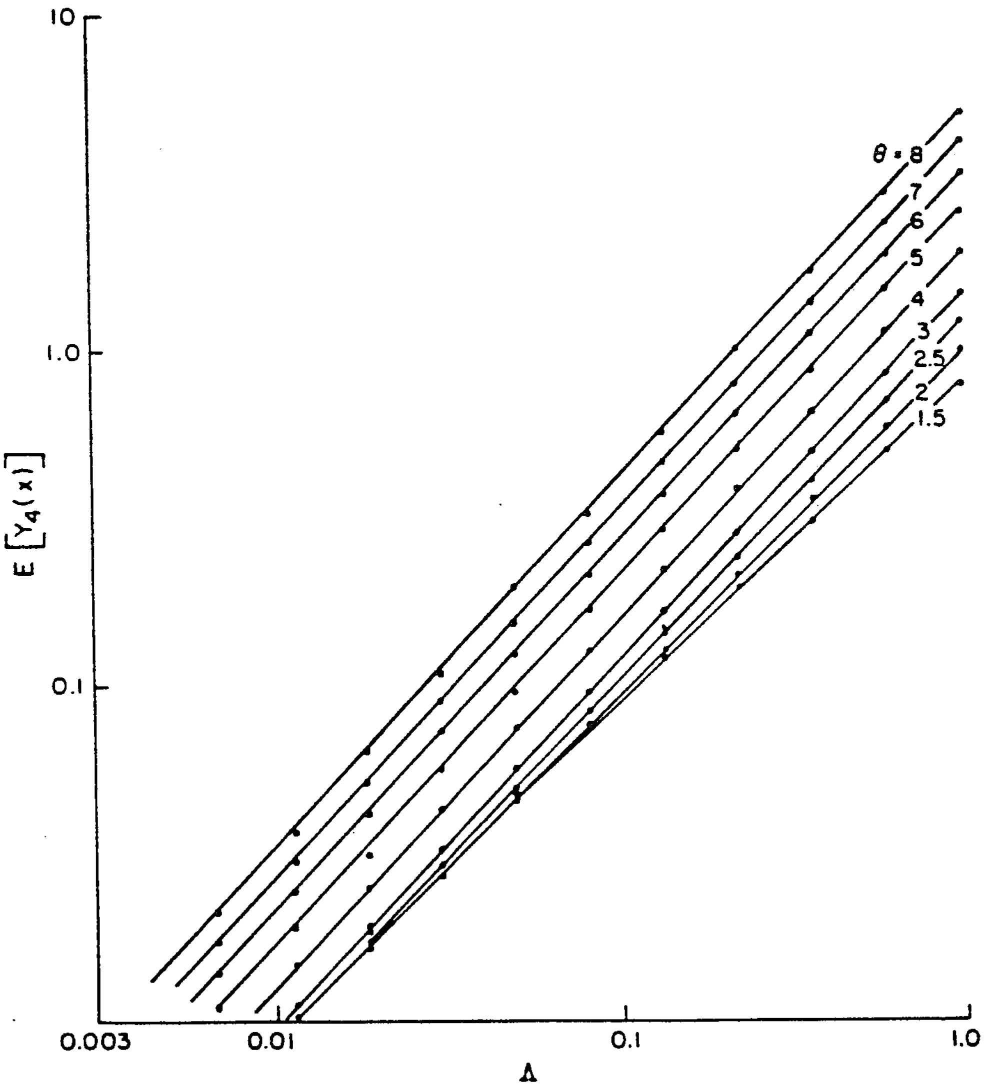

Fig 6.1 : Mean of final constraint $E(y_4(x))$ versus Λ for various values of θ for the TCEV Model

depend on θ, Λ and Λ_1 (note that Λ_2 is a function of Λ_1 via θ and Λ). A similar dependence is exhibited by the theoretical coefficients of skewness (and kurtosis) and variation respectively as can be easily shown using expressions of the moments given by Beran, et al. (1986). This implies that the estimation
of the constraints will likely have a variability increasing with the rank.

Figure 6.2 compares $E[y_4(x)]$ with the moments of the transformed variate

$$y = \frac{x}{\theta_1} - \ln \Lambda_1 \tag{6.29}$$

which is also TCEV distributed and depends on θ and Λ only. Skewness and kurtosis of both y and x variates are the same, while the mean of y is given by dividing the last two terms on the right-hand side of (6.24) by θ_1. One can note that $E[y_4(x)]$ exhibits a shape similar to that of $E[y]$ and that it is more sensitive to changes in either θ or Λ, particularly in the range of low values. This larger sensitivity is much more evident while comparing $E[y_4(x)]$ with skewness and kurtosis; this stipulates that entropy should provide dimensionless parameter estimates much less variable than those based on the method of moments. This POME procedure will be discussed in the next section.

Furthermore, equation (6.26) shows that $E[y_3]$ is also related to the probability, P_2, that the annual maximum value of x comes from the outlying component, P_2 having been derived by Beran, et al. (1986) as

$$P_2 = -\frac{1}{\theta} \sum_{j=1}^{\infty} \frac{(-1)^j \Lambda^j}{(j-1)!} \Gamma(j/\theta) \tag{6.30}$$

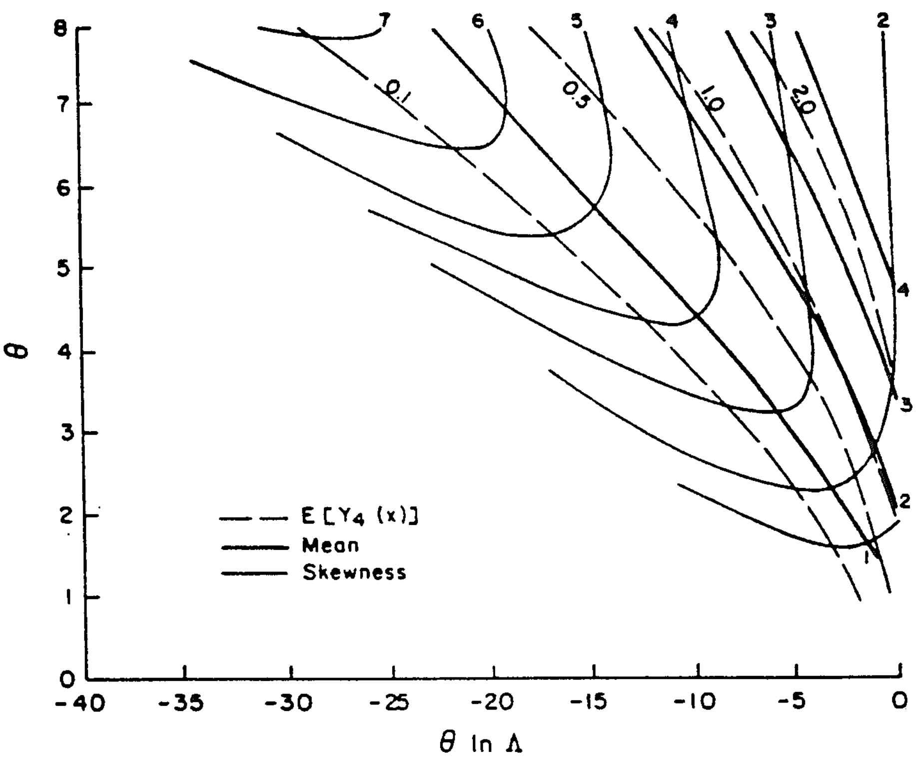

Fig 6.2 : Variation of mean and skewness coefficients of the transformed TCEV variate and mean of the final constraint $E(y_4(x))$ with parameters Λ and θ .

In fact, combining equations (6.26) and (6.30) gives $E[\exp(-x/\theta_2)] = P_2/\Lambda_2$. Analogously, it can be shown that $E[\exp(-x/\theta_1)] = P_1/\Lambda_1$, where P_1 represents the chance that the annual maximum value of x comes from the basic component.

A graph showing P_2 versus θ and Λ values has been provided by Beran, et al. (1986). It shows that for a given value of P_2, θ is a quasi-linear function of $\theta \ln \Lambda$ in almost the entire definition range of Λ. This suggests that an approximate relationship solely between P_2 and Λ can be confidently used for first order calculations. Figure 6.3 shows the goodness of this approximation, which has the following equation:

$$P_2 = 0.65\ \Lambda^{0.85} \tag{6.31}$$

Thus, one can write

$$E[\exp(-x/\theta_2)] = \frac{0.65\ \Lambda^{0.85}}{\Lambda_2} \tag{6.32}$$

6.3.4 Derivation of Constraints

Substituting equation (6.1) in equation (6.4) to express the entropy function H(f) as

$$\begin{aligned} H(f) &= -\int_{-\infty}^{\infty} \ln f(x)\ f(x)\ dx = (\ln \theta_1 - \ln \Lambda_1) \int_{-\infty}^{\infty} f(x)\ dx \\ &+ \frac{1}{\theta_1} \int_{-\infty}^{\infty} x\ f(x)\ dx + \Lambda_1 \int_{-\infty}^{\infty} \exp(-\ x/\theta_1)\ f(x)\ dx \\ &+ \Lambda_2 \int_{-\infty}^{\infty} \exp(-\ x/\theta_2)\ f(x)\ dx \\ &- \int_{-\infty}^{\infty} \ln[1 + \frac{(\Lambda_2/\theta_2)\ \exp(-\ x/\theta_2)}{(\Lambda_1/\theta_1)\ \exp(-\ x/\theta_1)}]\ f(x)\ dx \end{aligned} \tag{6.33}$$

On comparing with equation (6.8), the constraints of equations (6.9) - (6.13) can be obtained.

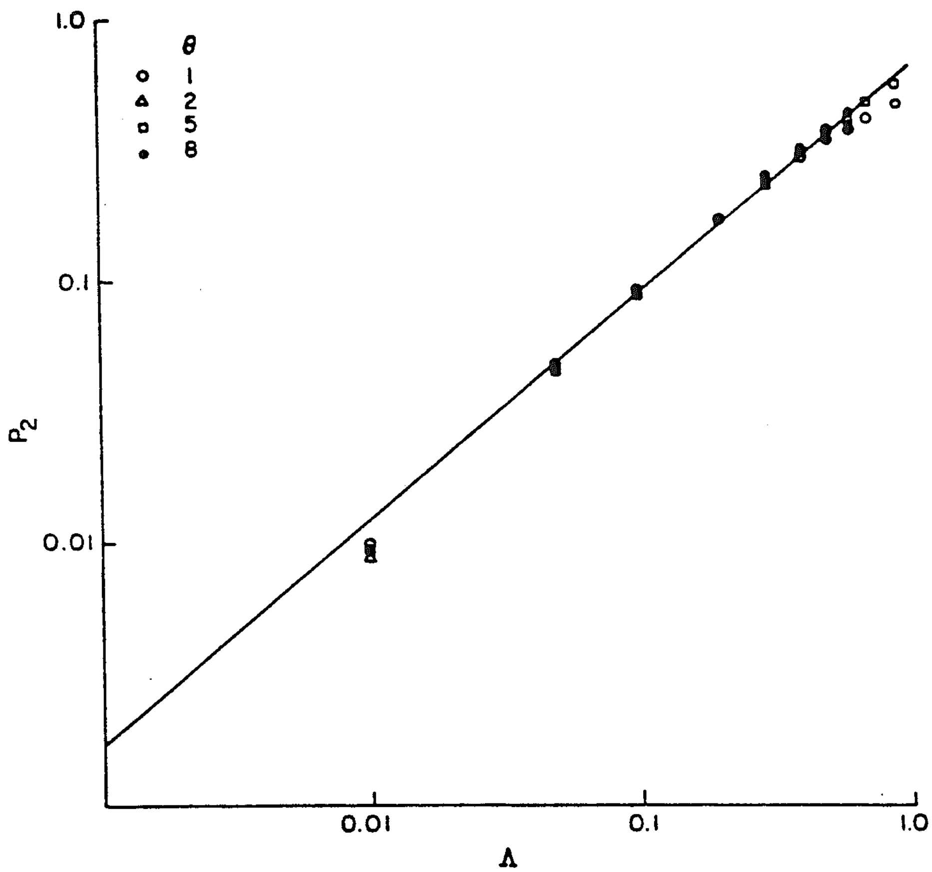

Fig 6.3 : Outlier Probability versus Λ for various values of θ for the TCEV Model

6.4 Estimation of Parameters

6.4.1 Point Estimation

Equations for estimation of parameters can be obtained by substituting sample values for the population means on the left-hand side of equations (6.24) to (6.26) and (6.28). The system of equations to be solved for giving estimates of the four parameters of the TCEV distribution thus takes the form:

$$\bar{x} = \theta_1 \ln \Lambda_1 + \theta_1 \gamma - \theta_1 \sum_{j=1}^{\infty} \frac{(-1)^j \Lambda^j}{j!} \Gamma(j/\theta) \tag{6.34}$$

$$\overline{\exp(-x/\theta_1)} = \frac{1}{\Lambda_1} [1 + \frac{1}{\theta} \sum_{j=1}^{\infty} \frac{(-1)^j \Lambda^j}{(j-1)!} \Gamma(j/\theta)] \tag{6.35}$$

$$\overline{\exp(-x/\theta_2)} = -\frac{1}{\theta\Lambda_2} \sum_{j=1}^{\infty} \frac{(-1)^j \Lambda^j}{(j-1)!} \Gamma(j/\theta) \tag{6.36}$$

$$\overline{\ln[1 + \frac{(\Lambda_2/\theta_2) \exp(-x/\theta_2)}{(\Lambda_1/\theta_1) \exp(-x/\theta_1)}]} = 0.1 \exp(-1)(3+\theta)^{2.059} \Lambda^{\ln 3 - 2(5.5^{-\theta})} \tag{6.37}$$

where the bar indicates that the sample mean of the underlying function is considered. For simplicity, the left-hand sides of equations (6.34) - (6.37) will be hereafter referred to as $\bar{Y}_1, \ldots, \bar{Y}_4$ respectively. Eliminating Λ_2 by way of θ, Λ and Λ_1, and rearranging,

$$\Lambda_1 = \frac{1}{\bar{Y}_2} [1 + \frac{1}{\theta} \sum_{j=1}^{\infty} \frac{(-1)^j \Lambda^j}{(j-1)!} \Gamma(j/\theta)] \tag{6.38}$$

$$\theta_1 = \bar{Y}_1 / \{\ln[1 + \frac{1}{\theta} \sum_{j=1}^{\infty} \frac{(-1)^j \Lambda^j}{(j-1)!} \Gamma(j/\theta)]/\bar{Y}_2 + \gamma - \sum_{j=1}^{\infty} \frac{(-1)^j \Lambda^j}{j!} \Gamma(j/\theta)\} \tag{6.39}$$

$$\Lambda = \frac{1 - \Lambda_1 \bar{Y}_2}{\Lambda_1^{1/\theta} \bar{Y}_3} \tag{6.40}$$

$$0.1 \exp(-1)(3 + \theta)^{2.059} \left(\frac{1 - \Lambda_1 \bar{Y}_2}{\Lambda_1^{1/\theta} \bar{Y}_3}\right)^{\ln 3 - 2(5.5^{-\theta})} = \bar{Y}_4 \tag{6.41}$$

Putting $\bar{Y}_3$ and $\bar{Y}_4$ respectively in the form:

$$\bar{Y}_3 = \overline{\exp[-x/(\theta_1 \theta)]} \tag{6.42}$$

$$\bar{Y}_4 = \overline{\ln\{1 + \frac{1}{\theta\Lambda_1} \frac{1 - \Lambda_1 \bar{Y}_2}{\bar{Y}_3} \exp[-\frac{x}{\theta_1} (\frac{1}{\theta} - 1)]\}} \tag{6.43}$$

One obtains that the only unknown in equation (6.41) is θ. On the other hand: (1) in equation (6.40), Λ does not appear on the right-hand side; (2) in equation (6.39), θ_1 is the only unknown once θ and Λ have been evaluated; and (3) Λ_1 does not appear on the right-hand side in equation (6.38). Therefore, a successive substitution iterative scheme is developed for estimating the four parameters as follows. Assign tentative values to θ and Λ, then successively estimate θ_1 by equation (6.39), Λ_1 by equation (6.38), θ by equation (6.41), and Λ by equation (6.40). Substitute the last values of θ and Λ for those previously obtained and start again from estimation of θ_1. Stop when θ and Λ no longer change. Note that the procedure is fast because equation (6.38) and (6.40) admit solution in closed form and equations (6.39) and (6.41), though not explicit, can be easily solved numerically, for each exhibits one unknown only.

6.4.2 Regional Estimation

A regional flood frequency estimation algorithm can be developed using equations (6.38) - (6.41) (obviously together with equations (6.42) and (6.43)), which can also be used to validate the regionalization model proposed by Fiorentino, et al. (1985) and also described in Fiorentino, et al. (1986). In short, this model assumes that dimensionless parameters θ and Λ do not change over extensive regions, while parameter Λ_1 is constant in smaller areas. In this paper, a regionalization algorithm, based on POME, to estimate θ and Λ is presented.

Suppose there are k gauged sites in a selected region which is assumed to be homogeneous with respect to θ and Λ. Let each site have an annual flood series (AFS) with n years of record. At each site, one must estimate the basic component parameters θ_1 and Λ_1, which vary from site to site, plus the two regional values of θ and Λ. Hence, there are in practice 2k + 2 unknowns. An equal number of independent equations is then needed.

The first 2k equations of the algorithm proposed herein arise from writing equations (6.38) and (6.39) k times, once for each available AFS. The other two equations are derived by taking the average of left-hand sides of equations (6.36) and (6.37) over all k sites,

$$\frac{1}{kn} \sum_{r=1}^{k} \sum_{i=1}^{n} \Lambda_{2r} \exp(-\frac{x_{ir}}{\theta_{2r}}) = -\frac{1}{\theta} \sum_{j=1}^{\infty} \frac{(-1)^j \Lambda^j}{(j-1)!} \Gamma(j/\theta) \tag{6.44}$$

$$\frac{1}{kn} \sum_{r=1}^{k} \sum_{i=1}^{n} \ln[1 + \frac{(\Lambda_2/\theta_2)_r \exp(-x_{ir}/\theta_{2r})}{(\Lambda_1/\theta_1)_r \exp(-x_{ir}/\theta_{1r})}]$$

$$= 0.1 \exp(-1)(3+\theta)^{2.059} \Lambda^{\ln 3 - 2(5.5)^{-\theta}} \tag{6.45}$$

Equations (6.44) and (6.45) can be written in a different manner taking into account the transformation:

$$y = \frac{x}{\theta_1} - \ln \Lambda_1 \tag{6.46}$$

which makes their left-hand sides dependent on θ and Λ only. Since the values of θ and Λ are assumed to be constant at every site, the following forms are thus obtained

$$\frac{1}{kn} \sum_{i=1}^{kn} \Lambda \exp(- y_i/\theta) = - \frac{1}{\theta} \sum_{j=1}^{\infty} \frac{(-1)^j \Lambda^j}{(j-1)!} \Gamma(j/\theta) \qquad (6.47)$$

$$\frac{1}{kn} \sum_{i=1}^{kn} \ln\{1 + \frac{\Lambda}{\theta} \exp[- (\frac{1}{\theta} - 1) y_i]\}$$

$$= 0.1 \exp(-1)(3 + \theta)^{2.059} \Lambda^{\ln 3 - 2(5.5)^{-\theta}} \qquad (6.48)$$

The two procedures are mutually equivalent, at least when the available AFS's have the same length at every site. Whichever is used, estimates of θ_1 and Λ_1 at any site need to be obtained together with regional estimates of θ and Λ. In fact, both sets of equations depend on the basic component parameters, the former in an explicit manner and the latter through the transformation of equation (6.46).

The iterative scheme proposed by Fiorentino and Gabriele (1985) for the regionalized TCEV-MLE procedure also successfully works using the POME-based estimation method. Details of this scheme can be readily found in Fiorentino, et al. (1986).

For a comparison between the proposed procedure and the MLE method, two features of the former look favorable: (1) Estimation of the basic component parameters, once θ and Λ have been evaluated, is relatively simple, for only equation (6.39) needs to be solved numerically. (2) The equations contain a smaller number of exponentials to be solved. However, only a large number of Monte Carlo experiments covering a wide range of situations, can confirm whether the TCEV-POME estimation procedure is competitive. In order to make an assessment, herein shown are the results derived by a limited number of computer simulation experiments.

6.5 Experimental Design and Results

The proposed estimation procedure was assessed, if only approximately, using the Monte Carlo technique, and generating synthetic series from a TCEV distribution with parameters $\theta_1 = 10$, $\Lambda_1 = 10$, $\theta = 3.067$, and $\Lambda = 0.173$, which is what was used to evaluate the TCEV-MLE procedure (Fiorentino and Gabriele, 1985; Arnell and Gabriele, 1986).

Two measures of performance were used, viz, the standardized bias:

$$BIAS = E(\hat{x} - x)/x$$

and the standardized root mean square error:

$$RMSE = E(\hat{x} - x)^2/x$$

where $\hat{x}$ and x denote the estimated and population values respectively of the statistic under examination.

Since the regionalization is the natural field for application of a distribution with a large number of parameters such as four, the attention was principally devoted to the assessment of the regionalized estimators.

One hundred repetitions of forty synthetic series, each with forty years of record, were generated, i.e., 100 homogeneous regions, each with 40 gauged sites, were simulated. Then the regionalization algorithm described above was applied. For each repetition, a regional estimate of θ and Λ together with forty on-site elements of θ_1 and Λ_1 were obtained. BIAS and RMSE of parameter and quantile regional estimators were then evaluated. Of course, due to the very short number of experiments, these results are not expected to reproduce the true values of BIAS and RMSE, but they do provide a first order approximation of the likely results.

In Table 6.1, statistical properties of the regionalized TCEV-POME estimators are compared with those obtained (when available) by Fiorentino and Gabriele (1985), who applied the MLE procedure to the same case. The comparison was also extended to the standardized quantiles $x_F/\bar{x}$, which are important when an index flood regionalization scheme is considered. As usual $\bar{x}$ denotes the sample mean, and x_F the quantile.

Table 6.1. BIAS and RMSE of quantile and regional parametric estimates.

	POME		MLE	
	BIAS	RMSE	BIAS	RMSE
Regional Parameters				
θ	-0.10	0.13	-0.01	0.15
Λ	0.04	0.19	0.32	0.72
Quantiles x_F				
Standardized Quantiles $x_F/\bar{x}$				
F = 0.5	0.01 0.01	0.09 0.02		
0.9	0.00 0.00	0.12 0.05	0.00 0.00	0.10 0.04
0.99	-0.05 -0.06	0.16 0.10	-0.02 -0.02	0.14 0.09
0.999	-0.07 -0.07	0.17 0.12	-0.03 -0.03	0.17 0.13

As regards quantiles (standardized or not), it can be seen that both methods exhibited practically the same RMSE, while the MLE procedure produced less BIAS than the POME method. On the other hand, the

POME procedure produced favorable RMSE with respect to the regional parameter estimates, particularly as regards Λ, whose MLE estimator was highly variable. Regarding the BIAS of θ and Λ estimators, the two competing methods exhibited opposite behavior, POME being outperformed with respect to θ but showing superiority in estimating Λ.

The results shown in Table 6.1 gave rise to the suspicion that relative to the MLE method, POME provided poorer estimators of the TCEV basic component parameters and that only for this reason the superiority substantially exhibited in estimating θ and Λ decayed when quantile estimates were considered. To substantiate it, a large number of experiments, five thousand synthetic series each with forty years of record were generated, and parameters θ_1 and Λ_1 were estimated for each series using equations (6.38) and (6.39) while keeping θ and Λ constant and equal to the population values. BIAS and RMSE of the so obtained estimates are shown in Table 6.2 where they were also compared with the respective values given by the MLE procedure (S. Gabriele, personal communication).

As had been suspected, in this case POME performed worse than MLE with regard to both parameters and quantiles, though no significant difference was shown with respect to BIAS. This is important because it suggests that valuable estimates of TCEV quantiles could be attained by combining a different basic component estimator with the POME one, the latter used to reach good estimates of the regional parameters θ and Λ.

6.6 Conclusions

The use of the principle of maximum entropy (POME) for deriving the two-component extreme value distribution (TCEV) sheds further light on this distribution, which has recently been shown to offer a practical

Table 6.2. BIAS and RMSE of site-specific estimates obtained keeping θ and Λ constant.

	POME		MLE	
	BIAS	RMSE	BIAS	RMSE
Parameters				
θ_1	-0.01	0.18	-0.02	0.14
Λ_1	0.20	0.64	0.18	0.52
Quantiles				
F = 0.5	0.00	0.09		
0.9	0.00	0.12	-0.01	0.10
0.99	-0.01	0.15	-0.01	0.12
0.999	-0.01	0.16	-0.01	0.13

approach to the regional flood frequency analysis. POME specifies the constraints sufficient to derive the TCEV distribution and with respect to which this distribution can be thought to be minimally prejudiced and consistent.

The estimation method based on the POME permits both site-specific and regional estimation. The equations to be solved for giving parameter estimates seem relatively simple when compared with those of the MLE method.

The regionalized TCEV-POME estimation procedure, proposed for application in a homogeneous region with respect to the shape parameters θ and Λ, performs comparably with the analogous regionalization algorithm which employs the MLE method. In particular, regional estimates of θ and Λ obtained by using the proposed procedure are found to exhibit favorable root mean square error (RMSE) in the examined case. Nevertheless, quantile elements obtained using at-site estimates of the basic component parameters together with regional estimates of θ and Λ do not show, relative to the MLE method, any improvement in RMSE. Also, they are, even if only slightly, more biased.

As a consequence, one can argue that valuable regional estimates of TCEV quantiles can be attained by combining the POME estimators of θ and Λ (or at least of Λ when only good performance in BIAS is sought) with another estimation procedure, e.g., the MLE method, which provides good estimates of the basic component parameters.

Chapter 7

CONCLUDING REMARKS

This study was concerned with evaluating and comparing the performance of various estimators of commonly used flood frequency models through Monte Carlo sampling experiments. The motivation for the study stemmed from the fact that for the rather small samples encountered in hydrology, the various competing estimators can yield estimates that differ significantly from each other in terms of the performance indices of MSE, BIAS, and SE. In affect the estimators extract different amounts of useful information from the sample as reflected by their performance indices. The estimate of the T-year return period quantile constitutes an important design variable in many engineering problems. Thus, it is important that the performance of competing estimators is evaluated so that the estimator extracting maximum information from the sample can be recommended for the purpose of design.

The models considered were Gumbel's extreme value type 1 (EV1), log Pearson type 3 (LP3), and the two component extreme value (TCEV) distributions. These models ranged from the simple and oldest two parameter EV1 to the most recent four parameter TCEV. An attempt was made to include all available estimators of a model in the performance evaluation studies. Some of the estimators were computationally not so amenable as others. The MLE estimator applied to the LP3 is one such example. In such cases, investigations were made into the behavior of the estimator, and extensive search procedures designed wherever possible. The underlying objective was to find the estimates for all the samples rather than reject the sample, as is the practice sometimes followed by some investigators. In the random sampling experiments, it does not appear to be reasonable to reject the samples.

The EV1 is perhaps one of the most well known and widely used models in flood frequency analysis. The model is simple to use, and can be particularly useful for moderately skewed flood data. Seven estimators of EV1 were used in the study reported here. Four estimators, namely, MLE, ENT, PWM, and MOM performed well and proved attractive on various grounds. MOM was computationally simplest to use followed by PWM, ENT, and MLE. MLE provided most efficient quantile estimates even for small sample sizes and large recurrence intervals. PWM yielded unbiased estimates of the quantiles, as can also be proved theoretically. ENT performed practically in the same way as MLE, and was somewhat easier to solve than MLE. A bias-corrected MOM quantile estimator was also derived based on the sampling experiments and its validity tested. The incorporation of serial correlation in the samples worsened the performance of all the estimators. However, the estimators performed much more similarly in this case.

LP3 is a three parameter distribution capable of modeling the high skewness and kurtosis likely to be encountered in AFS. The possibility of LP3 minimizing the condition of separation (Landwehr, et al., 1978) has not yet been fully investigated (Hoshi and Burges, 1981). Six estimators of LP3 were included in the simulation study. Many significant results emerged on account of this study. The method advocated by U.S. Water Resources Council faired poorly. An advantage in favor of this method is claimed to be its simplicity. However, simple procedures were devised for the superior estimators. These procedures obviated iterative procedures. Hence, it no longer seems reasonable to continue to use Water Resources Council's estimation method, even on grounds of computational simplicity. MLE estimator required extensive search procedure and large computational effort, and yielded disappointing results. ENT performed similar to MLE, and likewise proved unattractive. MIX and MMD came out as clearly superior to other estimators in terms of overall performance, with MIX seeming to hold an edge over MMD. The simple procedures proposed in the study to solve for MMD and MMI estimates should hopefully make the hydrologists more willing to use these methods of estimation for the LP3 distribution.

The TCEV was suggested by Rossi, et al. (1984), and was shown to account for the condition of separation in Italian floods. The distribution can be considered to be a mixture of two EV1 distributions. The drivation of TCEV via the principle of maximum entropy helped gain more insight into the properties of the distribution. It also yielded the ENT estimator which was shown to be suitable for both site-specific and regional estimation. The ENT estimation system appeared simpler than the MLE system. Performance of the ENT estimator was evaluated in a regional analysis framework, and was found to be comparable to MLE-based regionalization procedure (Rossi, et al., 1984).

REFERENCES

Arnell, N. W. and Gabriele, S., 1986. Regional flood frequency analysis using the two-component extreme value distribution: an assessment using computer simulation experiments. Proceedings of the Seminar, Combined Efficiency of Direct and Indirect Estimates for Point and Regional Flood Prediction, held at Perugia, Italy, December 16-20, 1985 (in press).

Beard, L. R., 1974. Technical Report - Flood flow frequency techniques. University of Texas at Austin.

Benson, M. A., 1968. Uniform flood-frequency estimating methods for Federal Agencies. Water Resources Research, 4, 891-908.

Beran, M. A., Hosking, J. R. M. and Arnell, N. W., 1986. Comment on 'Two-component extreme value distribution for flood frequency analysis,' by Rossi et al., 1984. Water Resources Research, 22(2), 263-266.

Bobee, B., 1975. The log Pearson type 3 distribution and its application in hydrology. Water Resources Research, 11, 681-689.

Bobee, B. and Robitaille, R., 1975. Correction of bias in the estimation of the coefficient of skewness. Water Resources Research, 11, 851-854.

Bobee, B. and Robitaille, R., 1977. The use of the Pearson type 3 and log Pearson type 3 distributions revisited. Water Resources Research, 13, 427-443.

Buckett, J. and Oliver, F. R., 1977. Fitting the Pearson type 3 distribution in practice. Water Resources Research, 13, 851-852.

Burges, S. J., Lettenmaier, D. P. and Bates, C. L., 1975. Properties of the three-parameter log-normal probability distribution. Water Resources Research, 11, 229-235.

Canfield, R. V., 1979. The distribution of the extreme of a mixture random variable with applications in hydrology. Input for Risk Analysis in Water Systems, edited by E. A. McBean, K. W. Hipel, and T. E. Unny, Water Resources Publications, Fort Collins, Colorado, pp. 77-84.

Charbeneau, R. J., 1978. Comparison of the two- and three-parameter log-normal distributions used in streamflow synthesis. Water Resources Research, 14, 149-150.

Condie, R., 1977. The log Pearson type 3 distribution: the T-year event and its asymptotic standard error by maximum likelihood theory. Water Resources Research, 13, 987-991.

Cunnane, C., 1978. Unbiased plotting positions--a review. Journal of Hydrology, 37, 205-222.

Cunnane, C., 1979. A note on the Poisson assumption in partial duration series models. Water Resources Research, 15, 489-494.

Cunnane, C., 1986. Review of statistical models for flood frequency estimation. Keynote paper presented at the International Symposium on Flood Frequency and Risk Analyses, held at Louisiana State University, Baton Rouge, LA.

Dalrymple, T., 1960. Flood frequency methods. USGS, Water Supply Paper 1543A, pp. 11-51, Washington, D.C.

Fiorentino, M. and Gabriele, S., 1984. A correction for the bias of the maximum likelihood estimators of Gumbel parameters. Journal of Hydrology, 7, 39-49.

Fiorentino, M. and Gabriele, S., 1985. Distribuzione TCEV: metodi di stima dei parametri e proprieta statistiche degli stimatori. Consiglio Nazaionale delle Ricerche-lRPl, Geodata n. 25, Cosenza, Italy (in Italian).

Fiorentino, M., Rossi, F. and Versace, P., 1985. Regional flood frequency estimation using the two-component extreme value distribution. Hydrological Sciences Journal, 30, 51-64.

Fiorentino, M., Gabriele, S., Rossi, F. and Versace, P., 1986. Hierarchical approach for regional flood frequency analysis. Paper presented at the International Symposium on Flood Frequency and Risk Analyses held at Louisiana State University, Baton Rouge, Louisiana.

Greenwood, J. A., Landwehr, J. M., Matalas, N. C. and Wallis, J. R., 1979. Probability weighted moments: definition and relation to parameters of several distributions expressible in inverse form. Water Resources Research, 15(5), pp. 1049-1054.

Greis, N. P., 1983. Flood frequency analysis: A review of 1979-1982. Rev. of Geophysics and Space Phys., 21(3), 699-706.

Greis, N. P. and Wood. E. F., 1981. Regional flood frequency estimation and network design. Water Resources Research, 17(4), 1167-1177.

Gumbel, E. J., 1941. The return period of flood flows. Ann. Math. Statist., 12(2), 163-190.

Gumbel, E. J., 1958. Statistics of Extremes. Columbia University Press, 375 pp.

Gumbel, E. J., 1966. Extreme value analysis of hydrologic data. Proc. Hydrol. Symposium No. 5, Statistical Methods in Hydrology, McGill Univ., published by Queen's Printer, Ottawa, pp. 147-169.

Hall, M. J., 1984. Urban Hydrology. Elsevier Applied Science Publishers Ltd., pp. 65-107.

Hoshi, K. and Burges, S. J., 1981a. Sampling properties of parameter estimates for the log Pearson type 3 distribution using moments in real space. Journal of Hydrology, 53, 305-316.

Houghton, J. C., 1978a. Birth of a parent: the Wakeby distribution for modelling flood flows. Water Resources Research, 14, 1105-1109.

Houghton, J. C., 1978b. The incomplete means estimation procedure applied to flood frequency analysis. Water Resources, Research, 14(6), pp. 1111-1115.

International Mathematical and Statistical (IMSL) Library, 1981. Manuals Vol. I-V.

Jain, D., 1986. A comparative evaluation of methods of flood frequency analysis. M.S. Thesis, submitted to the Department of Civil Engineering, Louisiana State University, LA.

Jain, D. and Singh, V. P., 1986. Estimating parameters of EV1 distribution for flood frequency analysis. Water Resources Bulletin, in press.

Jaynes, E. T., 1982. Papers on Probability, Statistics, and Statistical Physics. A reprint collection, Dordrecht, The Netherlands, D. Reidel.

Jaynes, E. T., 1957. Information theory and statistical mechanics, I. Physical Review, Vol. 106, 620-630.

Jaynes, E. T., 1961. Probability Theory in Science and Engineering. McGraw-Hill Book Co., New York.

Jaynes, E. T., 1968. Prior probabilities. IEEE Trans. Syst., Man Cybern., 3(SSC-4), 227-241.

Jowitt, P. W., 1979. The extreme-value type-I distribution and the principle of maximum entropy. Journal of Hydrology, 42, 23-28.

Kirby, W., 1974. Algebraic boundedness of sample statistics. Water Resources Research, 10,220-222.

Kite, G. W., 1977. Frequency and Risk Analysis in Hydrology. Water Resources Publications, Fort Collins, CO.

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

Kuczera, G., 1982b. Robust flood frequency models. Water Resources Research, 18(2), 315-324.

Landwehr, J. M., Matalas, N. C. and Wallis, J. R., 1978. Some comparisons of flood statics in real and log space. Water Resources Research, 14, 902-920.

Landwehr, J. M., Matalas, N. C. and Wallis, J. R., 1979a. Probability weighted moments compared with some traditional techniques in estimating Gumbel parameters and quantiles. Water Resources Research, 15(5), pp. 1055-1064.

Landwehr, J. M., Matalas, N. C. and Wallis, J. R., 1979b. Estimation of parameters and quantiles of Wakeby distributions. 1. Known lower bounds. Water Resources Research, 15, 1361-1372.

Landwehr, J. M., Matalas, N. C. and Wallis, J. R., 1979c. Estimation of parameters and quantiles of Wakeby distributions. 2. Unknown lower bounds. Water Resources Research, 15, 1373-1379.

Landwehr, J. M., Matalas, N. C. and Wallis, J. R., 1980. Quantile estimation with some more or less floodlike distributions. Water Resources Research, 16,547-555.

Lettenmaier, D. P. and Potter, K. W., 1985. Testing flood frequency estimation methods using a regions flood generating model. Water Resources Research, 21(12), 1903-1914.

Lettenmaier, D. P. and Burges, S. J., 1982. Gumbel's extreme value 1 distribution: a new look. Journal of the Hydraulic Division, Proceedings of ASCE, 108(HY4), pp. 502-514.

Linsley, R. K., Kohler, M. A. and Paulhus, J. L. H., 1975. Hydrology for Engineers. 2nd Ed., McGraw Hill Book Co., New York, NY.

Lowery, M. D. and Nash, J. B., 1970. A comparison of methods of fitting the double exponential distribution. Journal of Hydrology, 10, pp. 259-275.

Matalas, N. C. and Wallis, J. R., 1973. Eureka! It fits a Pearson type 3 distribution. Water Resources Research, 9, 281-289.

Matalas, N. C., Slack, J. R. and Wallis, J. R., 1975. Regional skew in search of a parent. Water Resources Research, 11, 815-826.

Montfort, M. A. J. Van, 1970. On testing that the distribution of extremes is of type-I when type-II is the alternative. Journal of Hydrology, 11, 421-427.

Natural Environmental Research Council, 1975. Flood Studies Report. Vol. 1: Hydrological Studies, NERC, London, 550 pp.

Nozdryn-Plotnicki, M. J. and Watt, W. E., 1979. Assessment of fitting techniques for the log Pearson type 3 distribution using Monte Carlo simulation. Water Resources Research, 15, 714-718.

Ochoa, I. D., Bryson, M. C. and Shen, H. W., 1980. On the occurrence and importance of Paretian-tailed distributions in hydrology. Journal of Hydrology, 48, 53-62.

Rao, D. V., 1980a. Log Pearson type 3 distribution: a generalised evaluation. Journal of Hydraulics Division, Proceedings of the ASCE, 106(HY5), 853-872.

Rao, D. V., 1980b. Log Pearson type 3 distribution: method of mixed moments. Journal of Hydraulics Division, Proceedings of the ASCE, 106(HY6), pp. 999-1019.

Rao, D. V., 1983. Estimating log Pearson parameters by mixed moments. Journal of Hydraulics Division, Proceedings of the ASCE, 109(HY8), pp. 1118-1131.

Rao, D. V., 1986. Fitting log Pearson type 3 distribution by maximum likelihood. Paper presented at the International Symposium on Flood Frequency and Risk Analyses, held at Louisiana State University, Baton Rouge, LA.

Rossi, F., Fiorentino, M. and Versace, P., 1984. Two-component extreme value distribution for flood frequency analysis. Water Resources Research, 20(7), 847-856.

Rossi, F., Fiorentino, M. and Versace, P., 1986. Reply to the Comment on 'Two-component extreme value distribution for flood frequency analysis.' Water Resources Research, 22(2), 267-269.

Sangal, B. P. and Biswas, A. K., 1970. The three-parameter lognormal distribution and its applications in hydrology. Water Resources Research, 6, 505-515.

Shannon, C. C. and Weaver, W., 1949. The Mathematical Theory of Communication. University of Illinois Press, Urbana, IL, 117 pp.

Shen, H.-W., Bryson, M. C. and Ochoa, I. D., 1980. Effect of tail behavior assumptions on flood prediction. Water Resources Research, 16, 361-364.

Singh, V. P. and Singh, K., 1985. Derivation of the Pearson type (PT) III distribution by using the principle of maximum entropy (POME). Journal of Hydrology, 80, 197-214.

Singh, V. P., Singh, K. and Rajagopal, K., 1985. Application of the principle of maximum entropy (POME) to hydrologic frequency analysis. Completion Report, Louisiana Water Resources Research Institute, Louisiana State University, Baton Rouge, Louisiana.

Singh, V. P., Rajagopal, A. K. and Singh, K., 1986. Derivation of some frequency distributions using the principle of maximum entropy (POME). Advances in Water Resources, Vol. 9, pp. 91-106.

Slack, J. R., Wallis, J. R. and Matalas, N. C., 1975. On the value of information in flood frequency analysis. Water Resources Research, 11, 629-647.

Sonuga, J. O., 1972. Principal of maximum entropy in hydrologic frequency and analysis. Journal of Hydrology, 17, 117-191.

Sonuga, J. O., 1976. Entropy principle applied to rainfall-runoff process. Journal of Hydrology, 30, 81-94.

Stedinger, J. R., 1980. Fitting log-normal distributions to hydrologic data. Water Resources Research, 16, 481-490.

U.S. Water Resources Council, 1967. Guidelines for determining flood flow frequency. Hydrology Committee, Bulletin 15, Washington, DC (also revised versions, Bulletin 17A, 1977, and Bulletin 17B, 1981).

Verdugo Lazo, A. C. Z. and Rathie, P. N., 1978. On the entropy of continuous probability distributions. IEEE Transactions on Information Theory, Vol. IT-24, 1, 120-122.

Versace, P., Fiorentino, M. and Rossi, F., 1982. Analysis of flood series by stochastic models. in Time Series Methods in Hydrosciences, edited by A. H. El-Shaarawi and S. R. Esterby, 315-324, Elsevier, New York.

Wallis, J. R., Matalas, N. C. and Slack, J. R., 1974. Just a moment!. Water Resources Research, 10, 211-219.

Wallis, J. R., Matalas, N. C. and Slack, J. R., 1977. Apparent regional skew. Water Resources Research, 13, 159-182.

Wallis, J. R. and Wood, E. F., 1985. Relative accuracy of log Pearson III procedures. Journal of Hydraulics Division, Proceedings of the ASCE, 111(HY7), 1043-1056.

Yevjevich, V., 1972. New vistas for flood investigations. Academia Nationale Dei Lincei, Roma, Quaderno N. 169, pp. 515-546.