TECHNICAL REF

1. Report No. FHWA/LA-92/249	2. Government Accession No.
4. Title and Subtitle Flood Frequency Analysis Using Optimization Techniques	5. Report Date October 1992
	6. Performing Organization Code
7. Author(s) Babak Naghavi and Fang X. Yu	8. Performing Organization Report No. 249
9. Performing Organization Name and Address Louisiana Transportation Research Center 4101 Gourrier Avenue Baton Rouge, LA 70808	10. Work Unit No.
	11. Contract or Grant No. LA HPR 92-1GT(B)
12. Sponsoring Agency Name and Address Louisiana Department of Transportation & Development P.O. Box 94245 Baton Rouge, LA 70804-9245	13. Type of Report and Period Covered Final Report August 1991-October 1992
	14. Sponsoring Agency Code

15. Supplementary Notes

Conducted in cooperation with U.S. Department of Transportation, Federal Highway Administration

16. Abstract

This study consists of three parts. In the first part, a comprehensive investigation was made to find an improved estimation method for the log-Pearson type 3 (LP3) distribution by using optimization techniques. Ninety sets of observed Louisiana flood data and 690 sets of Monte Carlo simulated LP3 data were used for the study. Based on the performance of 20 alternative optimization methods, a superior estimation method (named MALS), was found. As compared with the method of moments (MOM), the MALS method reduced the standard root mean square error (RMSE) by eight percent and the standard bias (BIAS) by 47 percent for the Monte Carlo simulated data. For the observed flood data, the MALS method reduced the relative root average square error (RRASE) by 13. 5 percent and the relative average bias (RAB) by 46 percent as compared with MOM.

In the second part of the study, an indexed regional optimization (IRO) procedure was developed to estimate the parameters of the generalized extreme value (GEV) distribution. The IRO procedure reduced RRASE by 20 percent and RAB by 100 percent as compared with the indexed regional probability weighted moments (IRPWM) procedure for the observed flood data.

In the third part of the study, the IRO procedure was extended to sites where no flood records were available. Limited verification showed that the extended procedure was reasonably accurate for watersheds of smaller than 1000 square miles.

Flood quantiles at the return periods of 2, 10, 25, 50, and 100 years were calculated by both the LP3/MALS and by the GEV/IRO at the 90 stream gauge sites. A procedure for estimating flood quantiles at ungauged sites is outlined with updated regional flood quantiles.

17. Key Words Flood Frequency, Optimization, Monte Carlo Simulation, Frequency Distributions, Parameter Estimation		18. Distribution Statement Unrestricted. This document is available to the public through the National Technical Information Service, Springfield, Virginia 21161	
19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 85	22. Price

Form DOT F1700.7 (1-92)

FLOOD FREQUENCY ANALYSIS USING OPTIMIZATION TECHNIQUES

FINAL REPORT

BY

BABAK NAGHAVI, P.E.
RESEARCH ADMINISTRATOR
LOUISIANA TRANSPORTATION RESEARCH CENTER
4101 GOURRIER AVE., BATON ROUGE, LA 70808

AND

FANG XIN YU
RESEARCH ASSOCIATE
LOUISIANA TRANSPORTATION RESEARCH CENTER
4101 GOURRIER AVE., BATON ROUGE, LA 70808

REPORT NUMBER 249
PROJECT NUMBERS 92-1GT(B)

CONDUCTED BY

LOUISIANA DEPARTMENT OF TRANSPORTATION AND DEVELOPMENT
LOUISIANA TRANSPORTATION RESEARCH CENTER
LOUISIANA STATE UNIVERSITY
IN COOPERATION WITH
U. S. DEPARTMENT OF TRANSPORTATION
FEDERAL HIGHWAY ADMINISTRATION

OCTOBER 1992

ACKNOWLEDGMENTS

This project was funded by the Louisiana Transportation Research Center under LTRC project numbers 92-1GT(B). The Project Review Committee consisted of Mr. Henry J. Barousse, Mr. Jack Manno and Mr. Zahir "Bo" Bolourchi. We would also like to thank Mr. Fred Cifreo for the two bridge analyses performed for the implementation examples. The combined efforts of all of these individuals contributed greatly toward successful completion of this project.

ABSTRACT

This study consists of three parts. In the first part, a comprehensive investigation was made to find an improved estimation method for the log-Pearson type 3 (LP3) distribution by using optimization techniques. Ninety sets of observed Louisiana flood data and 690 sets of Monte Carlo simulated LP3 data were used for the study. Based on the performances of 20 alternative optimization methods, a superior estimation method (named MALS), was found. As compared with the method of moments (MOM), the MALS method reduced the standard root mean square error (RMSE) by eight percent and the standard bias (BIAS) by 47 percent for the Monte Carlo simulated data. For the observed flood data, the MALS method reduced the relative root average square error (RRASE) by 13.5 percent and the relative average bias (RAB) by 46 percent as compared with MOM.

In the second part of the study, an indexed regional optimization (IRO) procedure was developed to estimate the parameters of the generalized extreme value (GEV) distribution. The IRO procedure reduced RRASE by 20 percent and RAB by 100 percent as compared with the indexed regional probability weighted moments (IRPWM) procedure for the observed flood data.

In the third part of the study, the IRO procedure was extended to sites where no flood records were available. Limited verification showed that the extended procedure was reasonably accurate for watersheds of smaller than 1000 square miles.

Flood quantiles at the return periods of 2, 10, 25, 50, and 100 years were calculated by both the LP3/MALS and by the GEV/IRO at the 90 stream gauge sites. A procedure for estimating flood quantiles at ungauged sites is outlined with updated regional flood quantiles.

IMPLEMENTATION STATEMENT

The MALS method developed in this study has been tested by using 90 observed Louisiana flood data and 690 Monte Carlo simulated data sets. All tests showed that the MALS method is superior to the method of moments (MOM) for estimating the parameters of the log-Pearson type 3 distribution. Similarly, the indexed regional optimization (IRO) procedure developed in this study is superior to the indexed regional probability weighted moments (IRPWM) procedure for estimating the parameters of the generalized extreme value distribution. Flood quantiles at 90 stream gauge sites have been predicted for six commonly used return periods by using the MALS method, the MOM method, the IRO procedure, and the IRPWM procedure. A procedure for estimating flood quantiles at ungauged sites is outlined in the third part of this study.

The findings of this study could easily be adopted by the DOTD design personnel. There appears to be no costs associated with the implementation of the recommended MALS method and the indexed regional optimization procedure for flood frequency analysis in Louisiana. It is anticipated that the findings of this study will permit more reliable design of highway drainage structures, resulting in savings in both construction and maintenance.

TABLES OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1
INTRODUCTION

In highway hydraulics design and maintenance work, accurate estimation of stream discharges are needed for a cost-effective design. The U.S. Water Resources Council (USWRC) has recommended the Log-Pearson type 3 (LP3) distribution along with the method of moments (MOM) for parameter estimation since 1967 *[1]* for at-site frequency analysis. Many studies have been carried out to test whether MOM is indeed suitable for estimating the parameters of the LP3 distribution *[2] [3] [4] [5] [6] [7]*. However, no general consensus on the performance of a specific estimation method has been reached to date. An examination of the past studies on parameter estimation indicates that if a method is found to perform well for a specific distribution by using Monte Carlo simulation, it may perform poorly on observed flood data, and vice versa. The reasons for these contradictory results are (1) the underlying population distribution is unknown for the observed data sets, (2) an estimator that performs superior for one distribution may not perform as well for another distribution, and (3) the performance indices for the observed data and for the Monte Carlo simulated data are usually not the same. In practice, a compromise has to be made to select an estimation method that performs relatively well for both types of data so that one may expect the estimation error not to change substantially for different distributions and data sets. Comparatively, MOM has been found to be a relatively simple and reliable estimator for the LP3 distribution.

In recent years, a great deal of effort has been invested in regionalizing statistical parameters *[8] [9]*. Regional frequency analysis consists of fitting a preselected probability distribution by using data from a group of stations with similar response to climatic conditions. Therefore, regionalization techniques have the advantage of reducing the uncertainty inherent in an individual station with short records. Other advantages of regionalization techniques are the ease of the use of regional quantiles for design purposes as well as their applicability to sites where flood records are not available.

Regional frequency techniques have been proposed by Dalrymple *[10]*, Stedinger *[11]* and Kuczera *[12]*. Greis and Wood *[8]* recommended an indexed method similar to that of Dalrymple *[10]*, but with the generalized extreme value (GEV) distribution as the base distribution and the method of probability weighted moments (PWM) as the parameter estimation method. This PWM method, first proposed by Greenwood, et al. *[13]*, has been shown to possess very attractive asymptotic characteristics when used to estimate the parameters of several distributions, especially in cases where the samples exhibit wide variability *[14]*. This characteristic makes the method very useful for regional frequency analysis. In support of this, Potter and Lettenmaier *[15]* tested 10 commonly used frequency methods and found that the indexed regional PWM procedure possessed predictive characteristics superior to the other methods tested. The indexed regional PWM procedure (IRPWM) were also applied to the GEV distribution by Hosking, et al. *[16]* and Schaefer *[17]*, and is the recommended procedure in the United Kingdom.

Flood quantile estimation for an ungauged site is encountered frequently in practice. Many studies have been carried out to find a simple and accurate methodology to estimate flood quantiles at ungauged sites *[18] [19] [20]*. Naghavi, et al., *[21]* applied the indexed regional PWM procedure to the GEV distribution using 85 sets of stream data from Louisiana. In their study, the state was divided into four hydrologically homogeneous regions, which are expected to have similar response to climatic conditions. Four regional regression equations were developed to represent the relationship between the mean annual maximum flood and the drainage area. With these regression equations, a procedure for estimating flood quantiles at any ungauged site was developed.

In spite of all recent statistical developments in the area of hydrological sciences, there still remains many challenging problems that need to be resolved. Some of these problems which directly influence design of hydrologic structures are parameter estimation and regionalization schemes.

CHAPTER 2
OBJECTIVES

The objectives of this study are:

(1) to develop a method (named MALS), which combines the method of moments, the method of least squares, and the conjugate gradient optimization algorithm, to estimate parameters of the log Pearson type 3 (LP3) distribution;

(2) to compare performance of the MALS method with the method of moments (MOM), the maximum likelihood estimate (MLE), and the method of maximum entropy (MME) by using both Monte Carlo simulated data and observed flood data;

(3) to predict flood quantiles for the recurrence intervals of 2, 5, 10, 25, 50, and 100 years at 90 Louisiana stream gage sites by using the MALS method;

(4) to develop an indexed regional optimization (IRO) procedure to estimate the parameters of the generalized extreme value (GEV) distribution;

(5) to compare the performance of the IRO procedure with the indexed regional probability weighted moments (IRPWM) procedure by using the observed flood data;

(6) to calculate flood quantiles for 90 Louisiana stream gauge sites at the recurrence intervals of 2, 10, 25, 50, and 100 years by using the IRO procedure; and,

(7) to extend the indexed regional optimization procedure to estimate flood quantiles at ungauged sites.

CHAPTER 3

SCOPE

The scope of this study encompassed the development and evaluation of at-site and regional parameter estimation procedures in order to improve the prediction accuracy of flood quantiles at both gauged and ungauged sites in Louisiana. Ninety sets of observed annual maximum flood data and 690 sets of Monte Carlo simulated LP3 data were used for the study.

An at-site parameter estimation method, which combines the method of moments and the method of least squares, was selected from 20 alternative methods tested. The selected method, named MALS, was compared with the method of moments (MOM), the method of maximum likelihood estimate (MLE), and the method of maximum entropy (MME). Conjugate gradient search algorithm was used to find the optimal set of parameters of the log Pearson type 3 distribution by minimizing both the relative root average square errors (RRASE) and relative average bias (RAB) between observed and estimated quantiles.

An indexed regional optimization (IRO) procedure was developed to estimate the parameters of the generalized extreme value (GEV) distribution. The parameters estimated by the indexed regional probability weighted moments (IRPWM) procedure serves as the initial estimates for the IRO procedure. The optimal set of regional parameters of the GEV distribution was then obtained by minimizing the relative root average square errors and relative bias between the observed and estimated quantiles in each homogeneous region, using the conjugate gradient search algorithm. The performance of the IRO procedure was compared with that of the IRPWM procedure.

Finally, a flood estimation procedure was developed by combining the IRO procedure with the regional regression equations developed by Naghavi, et al., *[21]*, which related the mean annual maximum flood to the watershed drainage area, to estimate flood quantiles at ungauged sites.

CHAPTER 4
AT-SITE FLOOD FREQUENCY ANALYSIS

Flood Frequency Analysis by the Conventional Procedure

As previously discussed in the introduction section, in conventional frequency analysis one may subjectively select some of the most frequently used distributions and parameter estimation methods and compare the computed results based on some selected performance indices such as the RRASE and RAB using the observed data. Then, the best combination of distribution and parameter estimation method is selected for the prediction of quantiles.

The computed RRASE and RAB for each of the 90 Louisiana stations were obtained by a comprehensive computer program developed by Naghavi, et al., *[7]*. This program can compute the RRASE and RAB for five distributions and three estimation methods. The five distributions considered are:

(1) Two-parameter log-normal (LNO2)

(2) Three-parameter log-normal (LNO3)

(3) Pearson type 3 (PT3)

(4) Log-Pearson type 3 (LP3)

(5) Extreme-value type 1 (EV1 or GUMBEL)

The three parameter estimation methods are:

(1) Method of moments (MOM)

(2) Maximum likelihood estimate (MLE)

(3) Method of maximum entropy (MME)

The average RRASE and average RAB for the 90 stations are listed in Tables 1 and 2 respectively. It is seen from Tables 1 and 2 that the LP3 distribution with MOM gave the smallest average RRASE for the 90 Louisiana stations whereas the EV1 distribution with the MME for parameter estimation gave the smallest RAB. In this situation, the LP3 /MOM would be selected as the best combination of distribution and method for the data because the RRASE is normally considered to be a more significant index than the RAB in practice, provided that the computed RAB is not excessively large as compared with other alternatives. Thus, based on the conventional frequency analysis procedure, the LP3/MOM is the best choice for predicting flood quantiles for the 90 flood gauge stations.

Table 1. Average RRASE for Five Distributions and Three Estimation Methods for 90 Sets of Louisiana Flood Data

		Distribution				
		LNO2	LNO3	PT3	LP3	EV1
	MAX	1.391	1.738	2.635	0.524	4.477
MOM	AVG	0.339	0.343	0.328	**0.171**	0.749
	MIN	0.086	0.067	0.044	0.046	0.082
	MAX	0.717	0.782	4.999	0.694	1.175
MLE	AVG	0.207	0.230	0.990	0.208	0.324
	MIN	0.075	0.085	0.055	0.079	0.080
	MAX	0.717	5.575	5.698	0.721	1.450
MME	AVG	0.206	0.863	1.150	0.208	0.403
	MIN	0.074	0.031	0.050	0.078	0.071

Table 2. Average RAB for Five Distributions and Three Estimation Methods for 90 Sets of Louisiana Flood Data

		Distribution				
		LNO2	LNO3	PT3	LP3	EV1
	MAX	0.577	0.179	1.040	0.153	0.074
MOM	AVG	0.092	-.053	0.027	0.026	-.171
	MIN	-.164	-.688	-.265	-.027	-1.58
	MAX	0.124	0.197	0.064	0.120	0.889
MLE	AVG	0.029	0.036	-.160	0.024	0.043
	MIN	0.007	0.007	-1.51	0.000	-.082
	MAX	0.124	0.063	0.062	0.123	0.110
MME	AVG	0.029	-.140	-.216	0.026	**-.001**
	MIN	0.007	-1.77	-.994	0.006	-.180

The Log-Pearson Type 3 Distribution

Let X and Y be two random variables related as Y=ln(X). If Y is Pearson type 3 distributed, then X is log-Pearson type 3 (LP3) distributed. The probability density function of the LP3 distribution is defined by:

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

where a, b, and c are the scale, shape, and location parameters, respectively. The population mean μ_y, standard deviation σ_y, and the coefficient of skewness γ_y of the variate Y can be expressed in terms of the distribution parameters as:

$$\mu_y = c + a b \tag{2}$$

$$\sigma_y^2 = b\, a^2 \tag{3}$$

$$\gamma_y = \frac{|\,a\,|}{a} \frac{2}{\sqrt{b}} \tag{4}$$

The distribution parameters a, b, and c can be calculated from the above three equations as:

$$b = \frac{4}{\gamma_y^2} \tag{5}$$

$$a = \frac{1}{2} \sigma_y \gamma_y \tag{6}$$

$$c = \mu_y - \frac{2\,\sigma_y}{\gamma_y} \tag{7}$$

The U.S. Water Resources Council *[1]* recommended the use of LP3 distribution along with the method of moments for parameter estimation in 1967. The LP3 has since been extensively studied and applied to hydrological frequency analysis *[22] [23] [24] [2] [25] [26] [4] [6] [7]*. However, the use of method of moments for estimating the parameters of the LP3 distribution is still controversial. Many studies have been carried out to compare different parameter estimation methods using either observed flood data or data generated by Monte Carlo simulation *[3] [4] [5] [6] [7]*. A general consensus on the performance of a specific estimation method has not been reached to date. Based on some of the past studies on parameter estimation, one may find that if a method is found to perform well by Monte Carlo simulation, it may perform poorly on most of the observed flood data, and vice versa. For example, Arora and Singh *[6]* found that the method of mixed moments (MIX) performs the best in their Monte Carlo simulation; Jain *[27]*, however, found MIX performs the worst using 55 observed annual maximum flood data. MOM is normally found to perform better for most observed flood data *[1] [2] [5] [7]*, however, MOM was found to perform poorly in Monte Carlo simulation *[6]*. The reason for these contradictory results is that for the observed data, the underlying population distribution and its parameters are unknown. The underlying distribution(s) could also be a combination of two or even more distributions *[28]*.

For this reason, one normally tries several flexible distributions along-with several parameter estimation methods, and selects the combination that fit the data best *[7]*. This practical procedure, of course, does not always warrant that the final selection is the correct one.

Performance Indices

The commonly used performance indices for the two types of data are generally different. Use of different performance indices may exhibit contradictory evaluation of a specific method. In terms of quantile prediction, the performance indices for Monte Carlo simulation are usually the standard root mean square error (RMSE) and the standard bias (BIAS) *[6]*:

$$RMSE = \sqrt{\frac{1}{m}\sum_{i=1}^{m}\left[\frac{x_c(i) - x}{x}\right]^2} \tag{8}$$

$$BIAS = \frac{\left[\frac{1}{m}\sum_{i=1}^{m} x_c(i)\right] - x}{x} \tag{9}$$

where m is the number of synthesized samples with the sample size n, x is the population quantile generated by using the population parameters for a specific return period T, and $x_c(i)$ is the computed quantile by using the estimated population parameters for the i-th sample with sample size n and the given return period T.

On the other hand, the performance indices for observed data are usually the relative root average square error (RRASE) and the relative average bias (RAB) between the observed and estimated quantiles using some plotting-position formula *[2]*:

$$RRASE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left[\frac{x_c(i) - x_o(i)}{x_o(i)}\right]^2} \tag{10}$$

$$RAB = \frac{1}{n} \sum_{i=1}^{n} \left[\frac{x_c(i) - x_o(i)}{x_o(i)} \right] \tag{11}$$

where $x_o(i)$ is the observed quantile at the i-th plotting position. Most of the performance evaluations of various methods have been based on only one type of data *[2] [5] [6]*. When a parameter estimation method is found to be superior based on one type of data, that method may perform poorly on the other type of data. Therefore, it is desirable to test the performance of an estimation method using both types of data.

Parameter Estimation

Some of the frequently used estimation methods in applied hydrology are: (1) the method of moments (MOM); (2) the maximum likelihood estimate (MLE), (3) the method of maximum entropy (MME); and (4) the least square method (LSM). Each of these methods has its advantages and disadvantages. The method of moments is good if the order of the moments used is no higher than two and the record length is sufficiently long (say at least 20 observations). Estimation involving the third, or higher, moment can be prone to large errors. To alleviate this problem, many combined or mixed parameter estimation methods have been proposed. For example, Houghton *[29]* proposed the method of incomplete means, Greenwood et al. *[13]* proposed the method of probability weighed moments, Rao *[30]* proposed the method of mixed moments, Hosking *[9]* proposed the method of L-moments, among others. The MLE and the MME methods normally perform similarly for both types of data. Numerical difficulties are, however, often experienced when solving the equations resulting from the two methods. The least square method is comparatively less used in frequency analysis partly because the resulting equations are nonlinear and their solutions are not unique. The methods of MOM, MLE, and MME, applied to the LP3 distribution, have been discussed by a number of authors in literature *[2][3][4][5][6][7]*. Therefore, only a summary of these methods is presented here:

MOM: the LP3 distribution parameter a, b, c can be estimated by equations (5) through (7), where the log-transformed population parameters μ_y, σ_y and γ_y are estimated by the sample mean $\overline{y}$, sample standard deviation S_y and sample skewness G_y as

$$\bar{y} = \frac{1}{n} \sum_{i=1}^{n} \ln\left[x_o(i)\right] \tag{12}$$

$$S_y^2 = \frac{1}{n-1} \sum_{i=1}^{n} \left[\ln\left[x_o(i)\right] - \bar{y}\right]^2 \tag{13}$$

$$G_y = \frac{n\,\lambda}{(n-1)(n-2)\,S_y^3} \sum_{i=1}^{n} \left\{\ln[x_o(i)] - \bar{y}\right\}^3 \tag{14}$$

where λ is a bias correction factor for the effect of sample size. One of the frequently used equations for the bias correction factor was given by Bobee and Robitaille *[2]*:

$$\lambda = \frac{(n-1)\left(1 + \frac{8.5}{n}\right)\sqrt{n^2 - n}}{n^2\,(n-2)} \tag{15}$$

MLE: the estimation equations for the distribution parameters are given as

$$a = \frac{S_1}{n\,b} \tag{16}$$

$$b = \frac{S_1\,S_2}{S_1\,S_2 - n^2} \tag{17}$$

$$n\,\Psi(b) - \sum_{i=1}^{n} \ln\left[\frac{\ln\left(x_o(i)\right) - c}{a}\right] = 0 \tag{18}$$

where

$$S_1 = \sum_{i=1}^{n} \left[\ln\left(x_o(i)\right) - c\right] \tag{19}$$

$$S_2 = \sum_{i=1}^{n} \frac{1}{\ln\left(x_o(i)\right) - c} \tag{20}$$

and $\Psi(b)$ is the digamma function and can be approximated by *[32]*:

$$\Psi(b) = \ln(b+2) - \frac{1}{2(b+2)} - \frac{1}{12(b+2)^2} + \frac{1}{120(b+2)^4} - \frac{1}{252(b+2)^6} - \frac{1}{b+1} - \frac{1}{b} \tag{21}$$

A numerical procedure to calculate parameters a, b, and c from equations (16) through (18) has been given by Arora and Singh *[6]*. Once these three parameters are computed, the mean, variance, and coefficient of skewness can be estimated by equations (12) through (14), respectively.

MME: The parameter estimation equations for this method are:

$$S_y = |a| \sqrt{b} \tag{22}$$

$$n\,a\,b = \sum_{i=1}^{n} \left\{\ln\left[x_o(i)\right] - c\right\} \tag{23}$$

$$n\Psi(b) = \sum_{i=1}^{n} \ln\left\{\frac{\ln\left[x_o(i)\right] - c}{a}\right\} \tag{24}$$

A numerical procedure to solve equations (22) through (24) for parameters a, b, and c was also given by Arora and Singh *[6]*. Once the parameters a, b, and c are computed, the log-transformed mean, variance, and the coefficient of skewness can then be calculated by equations (2) through (4), respectively.

Since MOM has been extensively studied and recommended by USWRC for estimating parameters in the LP3 distribution, the question is: does there exist a method that performs better than MOM on both types of data? The objective of this part of the study is to answer this question by conducting the following tasks:

(1) to find a better estimation method for the parameters of the LP3 distribution than MOM for both types of data;

(2) to compare the performance of the proposed method with MOM, MLE, and MME; and

(3) to update all flood quantiles at 90 Louisiana stream gauges.

Data Preparation

Observed Annual Maximum Flood Data:

Annual maximum flood data for all Louisiana stream gauges were obtained from the U.S. Geological Survey. In order to obtain valid statistical analysis, stations having less than 20 years of records or stations at regulated streams ware eliminated from the data. Furthermore, station having drainage areas of less than 10 square miles and record lengths less than 30 years were also eliminated from the data because observations from very small drainage areas with short record lengths are subject to large errors. A total of 90 stations were selected in this study. Figure 1 shows the locations of these stations, in which 84 stations are from Louisiana, one from Mississippi (2492360), two from Arkansas (7364190, 7365800), and three from Texas (8031000, 8030000, 8029500). Table 3 lists the locations, record period, record length, and drainage areas of these 90 stations. Table 4 lists the basic statistics of the raw data as well as the log-transformed data. Average record length for the

90 data sets is 36 years. The coefficient of variation, $\sigma_x/\bar{x}$, of the original data varies from 0.29 to 0.71, and the coefficient of skewness varies from -0.45 to 6.48.

Monte Carlo Simulated Data:

In order to evaluate the performance of a parameter estimation method by Monte Carlo simulation, 690 random samples from LP3 distribution were generated. The population parameters were chosen, based on Louisiana flood data characteristics, as a=0.125, b=16, and c=8; or equivalently, μ_y=10, σ_y=0.5, and γ_y=0.5. The Monte Carlo simulation consisted of two parts, a preliminary test and an extended test. For the preliminary test, 90 LP3 random samples were generated using the selected parameters (nine sample sizes of 15, 20, 25, 30, 40, 60, 80, 100, and 500 with ten samples at each sample size). For the extended test, 600 LP3 random samples were generated (six sample sizes of 20, 30, 40, 60, 100, and 500 with 100 samples at each samples size). The algorithm for generating the LP3 random samples by Monte Carlo simulation is provided in Appendix B.

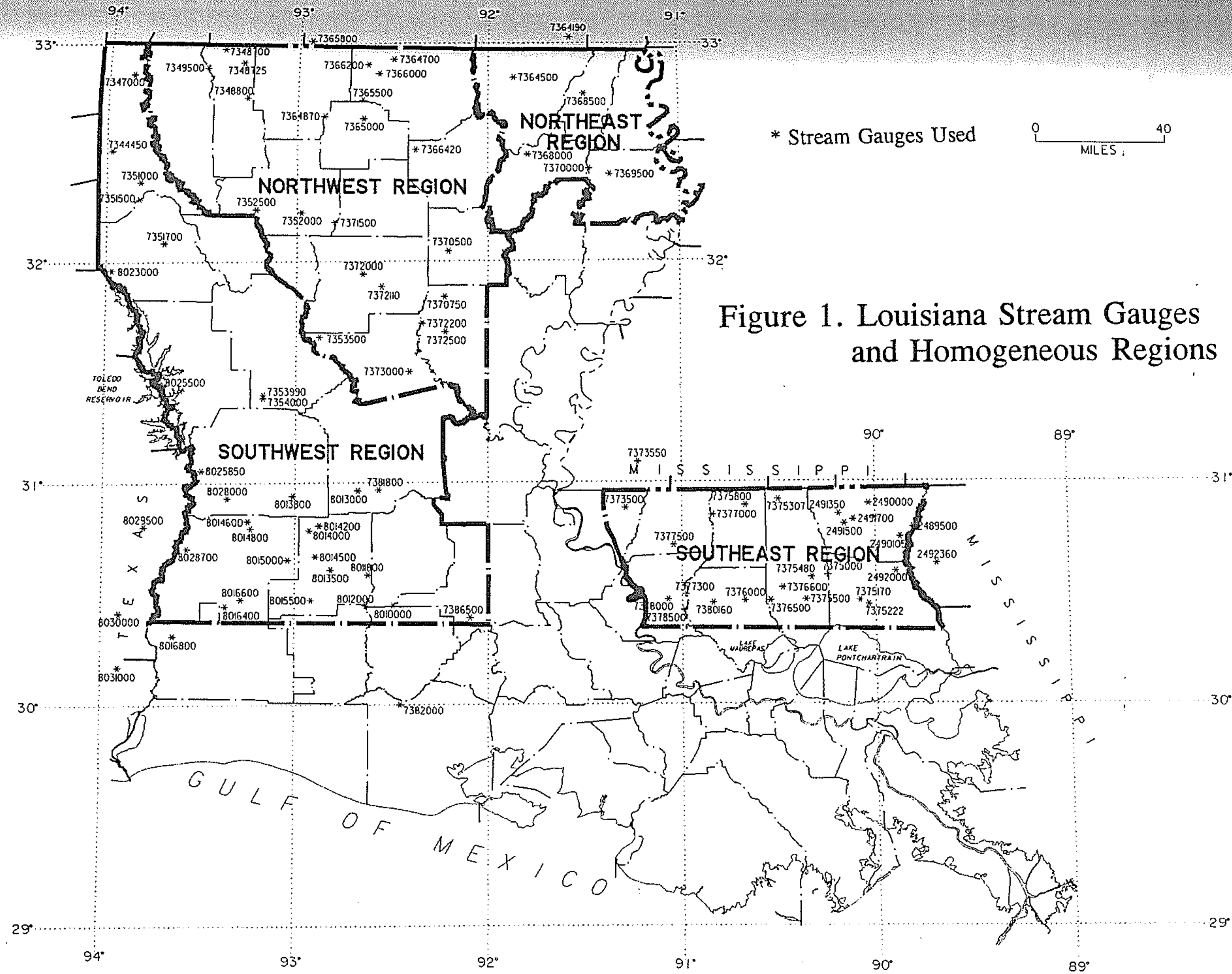

Figure 1. Louisiana Stream Gauges and Homogeneous Regions

Table 3. Louisiana Stream Gauges

STATION NUMBER	RECORD PERIOD	No. OBS.	LATITUDE ##°##'##"	LONGITU. ##°##'##"	AREA (SQ.MI.)	HOMO. REGION
2492000	1938-1990	53	303745	895350	1213	SE
2492360	1966-1986	21	303942	894110	175	SE
2490105	1964-1985	22	304656	895224	72.7	SE
7378500	1939-1990	52	302750	905925	1280	SE
7375222	1966-1990	25	302855	900220	46.1	SE
7380160	1951-1983	33	302945	905030	20.3	SE
7375170	1964-1983	20	302958	900504	88.2	SE
7376000	1941-1990	50	303013	904038	247	SE
7376500	1944-1990	47	303015	903245	79.5	SE
7375500	1939-1990	52	303023	902142	646	SE
7377300	1949-1983	35	303205	905850	884	SE
7376600	1951-1982	32	303340	902855	13.8	SE
7375480	1964-1983	20	303622	901956	91	SE
7375000	1944-1990	47	303657	901455	103	SE
2491500	1922-1990	69	305034	900943	990	SE
2491700	1964-1983	20	305140	900655	44.2	SE
2491350	1966-1986	21	305316	901128	42.2	SE
7377000	1949-1990	42	305320	905040	580	SE
7378000	1944-1990	47	303045	910425	284	SE
7377500	1943-1990	48	304521	910238	145	SE
7373500	1950-1970	21	305520	911735	35.3	SE
7369500	1930-1990	55	322555	912200	309	NE
7370000	1928-1990	63	322725	912830	782	NE
7368500	1950-1977	28	324755	913005	42	NE
7364500	1929-1980	52	325220	915204	1645	NE
7386500	1943-1970	28	302540	920530	19	SW
8012000	1939-1990	52	302850	923755	527	SW
8010000	1939-1990	52	302900	922925	131	SW
8015500	1939-1990	52	303010	925455	1700	SW
8011800	1964-1990	27	303710	923710	43.9	SW
8013500	1939-1990	52	303825	924850	753	SW
8014500	1940-1990	51	304155	925335	510	SW
8014000	1957-1983	27	304852	925534	171	SW
8014200	1950-1986	37	305011	925226	94	SW
8013000	1944-1990	47	305945	924025	499	SW
7382000	1938-1990	53	310000	922246	240	SW
7381800	1954-1986	33	310000	923400	68	SW
7373000	1942-1990	49	313210	922430	51	NW
7353500	1943-1968	26	314115	925240	47	NW
7372500	1940-1970	31	314258	921320	92	NW
7372200	1958-1990	33	314515	922040	1899	NW
7370750	1954-1983	30	315230	921335	47.6	NW
7372110	1965-1990	26	315510	923315	24	NW
7372000	1940-1981	42	315830	913910	654	NW
7370500	1941-1970	30	320455	921225	271	NW

SE=Southeast, SW=Southwest, NE=Northeast, NW=Northwest

Table 3. Louisiana Stream Gauges (cont'd)

STATION NUMBER	RECORD PERIOD	No. OBS.	LATITUDE ##°##'##"	LONGITU. ##°##'##"	AREA (SQ.MI.)	HOMO. REGION
7371500	1939-1990	52	321225	924805	355	NW
7352000	1941-1990	50	321500	925835	154	NW
7373550	1957-1986	30	310520	911430	0.21	SE
7366420	1966-1990	25	323230	922245	113	NW
7365000	1941-1968	28	324050	923910	355	NW
7364870	1966-1990	25	324120	925130	47	NW
7365500	1941-1970	30	324550	923930	178	NW
7366000	1941-1983	43	325315	923425	462	NW
7366200	1956-1990	35	325545	923758	208	NW
7364700	1956-1977	22	325719	922959	141	NW
8031000	1953-1986	34	301110	935430	83	SW
8016800	1954-1984	31	301959	933744	177	SW
8030000	1952-1983	32	322552	935428	69.2	SW
8016400	1946-1984	39	302815	932135	148	SW
8016600	1946-1983	38	303005	931635	82	SW
8015000	1940-1970	31	304055	930215	238	SW
8028700	1956-1981	26	304332	933336	13.1	SW
8029500	1952-1987	36	304908	934707	128	SW
8014600	1964-1983	20	305105	931445	26.3	SW
8028000	1952-1990	39	305710	932110	365	SW
8013800	1963-1983	21	305805	930045	10.4	SW
8025850	1967-1986	20	310432	932922	9.66	SW
8025500	1956-1986	31	311825	933056	148	SW
7354000	1950-1979	30	312430	931015	21.4	SW
7353990	1966-1990	25	312520	931014	37.3	SW
8023000	1956-1983	28	315825	935815	96.5	SW
7351700	1958-1983	26	320605	934145	19.5	SW
7352500	1941-1983	43	321540	931250	423	NW
7351500	1939-1990	52	321800	934940	66	SW
7351000	1939-1981	43	322235	934920	79	SW
7344450	1956-1986	31	323100	935820	80.5	SW
2490000	1949-1968	20	325205	900010	12.1	SE
7348700	1958-1990	33	325940	932347	605	NW
7349500	1939-1990	52	325418	932858	546	NW
7348725	1966-1990	25	325555	931730	33.1	NW
7348800	1954-1977	24	324610	931600	66.9	NW
7347000	1945-1969	25	325125	935220	116	NW
7364190	1926-1970	45	330415	913440	1170	NE
7365800	1956-1984	29	330221	925615	180	NW
7362100	1939-1988	50	332233	924637	385	NW
2489500	1939-1990	52	304735	894915	6573	SE
7375800	1956-1990	35	305550	904024	89.7	SE
7375307	1966-1990	25	305723	903013	52	SE
8014800	1956-1979	24	304909	931351	120	SW
7368000	1928-1990	63	322852	914752	1226	NE

SE=Southeast, SW=Southwest, NE=northeast, NW=Northwest

Table 4. Statistics of 90 Sets of Louisiana Flood Data

STATION NUMBER	No. OF OBS.	$\bar{X}$	S_x	G_x	$\bar{Y}$	S_y	G_y
2492000	53	26378	22745	3.124	9.897	0.765	-0.073
2492360	21	7289	4417	1.445	8.719	0.615	-0.034
2490105	22	3006	2476	1.941	7.702	0.807	0.181
7378500	52	34432	24543	1.795	10.21	0.709	-0.160
7375222	25	2601	1419	0.035	7.673	0.684	-0.958
7380160	33	1194	567	0.528	6.962	0.523	-0.449
7375170	20	4687	2960	1.827	8.287	0.600	0.517
7376000	50	6830	5036	1.627	8.570	0.753	-0.276
7376500	47	3666	2128	1.579	8.054	0.563	-0.038
7375500	52	17840	14915	2.587	9.490	0.800	-0.148
7377300	35	29957	18159	1.644	10.14	0.584	0.217
7376600	32	1372	509	0.092	7.145	0.430	-1.180
7375480	20	8558	7293	2.479	8.695	0.921	-0.352
7375000	47	6154	5775	2.332	8.314	0.959	-0.217
2491500	69	26818	19873	2.330	9.941	0.752	-0.398
2491700	20	4298	4087	2.409	7.827	1.231	-1.062
2491350	21	3073	2603	2.104	7.687	0.869	0.106
7377000	42	27637	22097	1.725	9.893	0.883	-0.444
7378000	47	12561	6940	1.258	9.274	0.614	-0.745
7377500	48	8800	5981	0.776	8.816	0.781	-0.343
7373500	21	7539	4584	1.100	8.731	0.675	-0.492
7369500	55	2774	792	-0.048	7.883	0.313	-0.789
7370000	63	5495	2176	0.434	8.528	0.427	-0.589
7368500	28	1102	372	0.584	6.946	0.359	-0.760
7364500	52	7108	2554	0.067	8.780	0.485	-0.219
7386500	28	1223	532	1.098	7.000	0.521	-1.839
8012000	52	9351	6530	3.048	8.988	0.521	1.071
8010000	52	5201	2490	0.818	8.428	0.551	-1.214
8015500	52	34751	28780	2.875	10.22	0.685	0.049
8011800	27	2482	1930	1.749	7.707	0.765	-0.732
8013500	52	18218	13350	2.443	9.585	0.687	-0.160
8014500	51	17048	21080	5.559	9.352	0.868	0.115
8014000	27	5786	5231	2.932	8.365	0.768	0.404
8014200	37	5226	5544	4.343	8.201	0.857	-0.023
8013000	47	17300	12771	1.849	9.488	0.790	-0.599
7382000	53	2234	3820	7.547	7.366	0.646	2.424
7381800	33	2855	2229	1.657	7.659	0.816	-0.288
7373000	49	4978	5323	2.123	7.949	1.131	-0.053
7353500	26	3064	3412	2.443	7.368	1.274	-0.234
7372500	31	4772	5723	5.566	8.162	0.697	1.544
7372200	33	27215	22839	2.261	9.875	0.884	-0.501
7370750	30	2582	2215	3.085	7.603	0.690	0.713
7372110	26	3888	4635	2.373	7.686	1.068	0.767
7372000	42	9309	6224	1.152	8.851	0.881	-1.373
7370500	30	6399	5123	1.779	8.375	1.044	-1.451

Table 4. Statistics of 90 Sets of Louisiana Flood Data (Cont'd)

STATION NUMBER	No. OF OBS.	$\bar{X}$	S_x	G_x	$\bar{Y}$	S_y	G_y
7371500	52	8997	7946	2.484	8.767	0.874	-0.573
7352000	50	3595	3040	1.817	7.848	0.863	-0.173
7373550	30	231	95	0.829	5.356	0.440	-0.705
7366420	25	5498	6857	3.698	8.100	1.014	0.319
7365000	28	7696	6156	2.313	8.661	0.803	-0.469
7364870	25	2988	2817	3.631	7.622	0.994	-1.407
7365500	30	4001	5035	5.432	7.930	0.776	1.294
7366000	43	8311	9224	4.183	8.675	0.820	0.147
7366200	35	4797	4915	3.837	8.126	0.859	-0.297
7364700	22	4485	6433	3.856	7.874	0.918	1.908
8031000	34	1701	1160	1.715	7.233	0.659	-0.103
8016800	31	4442	3622	3.600	8.173	0.669	0.109
8030000	32	2541	1617	2.066	7.665	0.610	-0.217
8016400	39	4892	3557	2.050	8.278	0.661	0.262
8016600	38	5063	3191	1.314	8.353	0.594	0.459
8015000	31	8611	8363	2.230	8.618	0.988	0.021
8028700	26	935	651	4.047	6.688	0.528	0.956
8029500	36	3843	4434	3.294	7.844	0.844	1.083
8014600	20	2556	2248	2.516	7.532	0.811	0.199
8028000	39	13887	15419	2.349	9.014	1.036	0.400
8013800	21	1320	1018	2.207	6.902	0.817	-0.758
8025850	20	789	771	3.556	6.365	0.747	1.235
8025500	31	6448	7475	2.714	8.303	0.922	0.967
7354000	30	2958	1458	0.531	7.854	0.572	-0.963
7353990	25	4836	4517	2.269	8.072	0.974	-0.312
8023000	28	2487	1847	1.922	7.549	0.782	-0.342
7351700	26	1501	2294	6.482	6.865	0.887	0.505
7352500	43	4894	3411	1.344	8.269	0.682	0.215
7351500	52	5992	4442	1.925	8.392	0.903	-1.440
7351000	43	4237	3041	1.457	8.043	0.907	-1.387
7344450	31	4132	4323	3.193	7.940	0.887	0.067
2490000	20	2412	2448	2.935	7.248	1.208	-0.969
7348700	33	9324	8183	2.478	8.806	0.856	-0.211
7349500	52	5347	3570	1.595	8.359	0.713	-0.492
7348725	25	2092	1626	2.938	7.320	0.977	-2.156
7348800	24	2545	2234	2.925	7.526	0.823	-0.021
7347000	25	1561	774	3.384	7.263	0.422	0.494
7364190	45	4699	1635	-0.447	8.368	0.481	-2.361
7365800	29	7568	12738	5.305	8.322	1.038	0.538
7362100	50	9001	9105	3.484	8.743	0.859	0.006
2489500	52	49525	24690	1.715	10.07	0.468	0.051
7375800	35	5890	6377	3.152	8.231	0.958	0.305
7375307	25	5511	5428	2.031	8.110	1.089	-0.159
8014800	24	5162	3834	2.166	8.283	0.782	-0.435
7368000	63	1900	747	0.610	7.458	0.470	-1.573

Proposed Estimation Method

The purpose of the proposed method in this part of the study is to develop a combination method which combines the method of moments and the method of least squares and performs better than MOM for the LP3 distribution. However, there may exist many combinations, depending upon which parameters are estimated by MOM and which are estimated by the least squares method (LSM), and also which objective function is to be minimized. Practically, there are four types of combined methods. Let $\overline{y}$, S_y and G_y be the estimated values of μ_y, σ_y and γ_y by MOM, respectively, and let $\hat{a}$, $\hat{b}$, and $\hat{c}$ be the estimated values of the LP3 distribution parameters a, b, and c by MOM, respectively. The four combinations of MOM and least squares are:

(1) Estimate μ_y and σ_y by MOM and γ_y by the least square method (LSM), where the coefficient of skewness estimated by MOM serves as an initial value for an optimal solution by LSM.

(2) Estimate μ_y by MOM, and σ_y and γ_y by LSM, where S_y and G_y serve as initial values for LSM.

(3) Estimate μ_y, σ_y and γ_y by LSM where $\overline{y}$, S_y and G_y obtained by MOM serve as initial values of LSM.

(4) Estimate a, b, and c by LSM where $\hat{a}$, $\hat{b}$ and $\hat{c}$, estimated by MOM using Equations (5) through (7), serve as the initial values for LSM.

Since evaluation of an estimator for an observed data set is normally made in terms of the relative root average square error (RRASE) and the relative average bias (RAB) defined in Equations (10) and (11), or in terms of root average squares error (RASE, Equation 25B) and the average bias (ABIAS, Equation 25C), the objective function for LSM should be the RRASE, RAB, RASE, ABIAS, or a combination of them. In this study, five objective functions in terms of RRASE, RAB, RASE, and ABIAS were investigated. The five objective functions were:

$$z = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left[x_c(i) - x_o(i)\right]^2} + \left|\frac{1}{n}\sum_{i=1}^{n}\left[x_c(i) - x_o(i)\right]\right| \qquad (25A)$$

$$z = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left[x_c(i) - x_o(i)\right]^2} \qquad (25B)$$

$$z = \left|\frac{1}{n}\sum_{i=1}^{n}\left[x_c(i) - x_o(i)\right]\right| \qquad (25C)$$

$$z = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left[\frac{x_c(i) - x_o(i)}{x_o(i)}\right]^2} + \left|\frac{1}{n}\sum_{i=1}^{n}\left[\frac{x_c(i) - x_o(i)}{x_o(i)}\right]\right| \qquad (25D)$$

$$z = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left[\frac{x_c(i) - x_o(i)}{x_o(i)}\right]^2} \qquad (25E)$$

A total of 20 alternative methods are possible by combining the above four combinations of MOM and LSM and five objective functions. Table 5 lists these 20 alternatives. The conjugate gradient optimization (CGO) algorithm was employed to find the solution for the least squares method. The CGO algorithm has been described in many optimization textbooks *[33] [34]*. Appendix A gives a detailed description of the CGO algorithm. The MOM estimate(s) is used as a starting point for the CGO search. The estimated quantile for a given cumulative probability, F, computed by a selected plotting-position formula, can be calculated as:

$$X_F = exp(\bar{y} + KS_y) \qquad (26)$$

where K is the frequency factor which is approximated *[32]* by:

$$K = t + (t^2-1)\frac{G_y}{6} + \frac{1}{3}(t^3-6t)\left(\frac{G_y}{6}\right)^2 - (t^2-1)\left(\frac{G_y}{6}\right)^3 + t\left(\frac{G_y}{6}\right)^4 - \frac{1}{3}\left(\frac{G_y}{6}\right)^5 \tag{27}$$

in which t is the standard normal variate and can be calculated *[32]* as:

$$t = \begin{cases} w - \dfrac{C_0 + C_1 w + C_2 w^2}{1 + d_1 w + d_2 w^2 + d_3 w^3}, & F \leq 0.5 \\ -w + \dfrac{C_0 + C_1 w + C_2 w^2}{1 + d_1 w + d_2 w^2 + d_3 w^3}, & F > 0.5 \end{cases} \tag{28}$$

where F is the exceedance probability. Note that the last term on the right-hand side of Equation (27) has a negative sign which is a correction of the original equation given by Kite *[32]*.

$$w = \begin{cases} \sqrt{ln\left(\dfrac{1}{F^2}\right)}, & F \leq 0.5 \\ \sqrt{ln\left[\dfrac{1}{(1-F)^2}\right]} & F > 0.5 \end{cases} \tag{29}$$

The coefficients in Equation (28) are given by Kite *[32]* as:

$C_o = 2.515517$ $\quad$ $d_1 = 1.432788$

$C_1 = 0.802853$ $\quad$ $d_2 = 0.189269$

$C_2 = 0.010328$ $\quad$ $d_3 = 0.001308$

Development of the MALS Method

The values of RMSE and BIAS for seven selected quantiles at the return periods of 2, 5, 10, 25, 50, 100, and 200 years were computed from each of the 20 alternative methods by using the 90 sets of Monte Carlo simulated data. Tables 6 and 7 give examples of the computed results for the sample sizes of 40 and 100. The population quantiles were generated by using the population parameters. The estimated quantiles were generated using the parameters estimated by each of the alternative methods. The values of RMSE and BIAS are the average values for the 10 generated LP3 samples of size of 40 or 100 for each of the seven selected quantiles. As a result, for all of the nine sample sizes, the MALS4, which estimates the μ_y and σ_y by MOM and γ_y by LSM, using the objective function "25D", gave the smallest RMSE and BIAS. Therefore, this method, MALS4, was selected as the best representative method and is hereafter referred to as the MALS method.

The performances of the MALS method for the 90 Monte Carlo simulated samples were further compared with those of MOM, MLE, and MME in terms of standard root mean square error (RMSE) and standard bias (BIAS) using the seven selected quantiles. The computed RMSE and BIAS are the average values for ten samples of each of the nine selected samples sizes. Table 8 shows the RMSE values for the seven selected quantiles and the nine sample sizes of 15, 20, 25, 30, 40, 60, 80, 100, and 500. On the average, the RMSE values for the seven quantiles and the nine sample sizes for MLE, MME, MALS and MOM were 0.1193, 0.1227, 0.1354 and 0.1568, respectively. The MLE performed the best. The RMSE of MALS was 13.6 percent smaller than that of MOM. Table 9 lists the BIAS for the seven selected quantiles and the nine sample sizes. On the average, for the seven quantiles and the nine sample sizes, the BIAS values for MALS, MOM, MLE and MME were 0.0233, 0.0349, 0.0566, and 0.0625, respectively. The MALS method gave the smallest BIAS. It reduced the BIAS by 33 percent as compared with MOM, 59 percent as compared with MLE, and 63 percent as compared with MME. Although MLE and MME are the two best methods in terms of the RMSE test, they tend to under-estimate the quantiles for large return periods (larger than or equal to 50 years). This is shown in Table 9 in which the larger the return period, the larger is the negative RAB for MLE and MME. The MALS method, on the other hand, normally reduces the BIAS by more than 50 percent for large return quantiles as compared with MLE and MME. As compared with MOM, MALS yields larger reductions both in RMSE and in BIAS for predicting flood quantiles of larger return periods.

Testing the MALS Method by the Observed Flood Data

To test the MALS method, first, the 90 sets of observed annual maximum flood data in Louisiana were used. The computed relative root average square error (RRASE) for the 90 data sets are given in Table 10. On the average, the RRASE values for the observed data sets for MOM, MLE, MME and MALS were 0.1699, 0.2080, 0.2083, and 0.1469, respectively. The MALS was found to be the best method and the MLE the worst. The MALS reduced the RRASE by 13.5 percent as compared with MOM, 29.5 percent as compared with MLE, and 29.4 percent as compared with MME. The relative average bias (RAB) for the 90 data sets is shown in Table 11. On the average, the RAB values were 0.0149, 0.0242, 0.0258, and 0.0275, respectively, for MALS, MLE, MME and MOM. The MALS method reduced the RAB by 45.8 percent as compared with MOM, 38.4 percent as compared with MLE, and 42.2 percent as compared with MME.

Extended Test for MALS by Monte Carlo Simulation

To further test the MALS method, 600 additional sets of Monte Carlo simulated LP3 data for sample sizes of 20, 30, 40, 60, 100, and 500 were generated, i.e., one hundred samples were generated for each sample size. Performance indices of RMSE and BIAS for the methods of MOM, MLE, MME, and MALS were computed and compared. Again, seven selected quantiles corresponding to the return periods of 2, 5, 10, 25, 50, 100, and 200 years were used for the comparison. Table 12 shows the computed RMSE values for the seven selected quantiles and six sample sizes. On the average, The RMSE for the seven quantiles and six sample sizes for MOM, MLE, MME, and MALS are 0.194, 0.162, 0.148, and 0.179, respectively. Again, MLE performed the best in terms of RMSE. The RMSE of MALS was eight percent smaller than that of MOM. Table 13 lists the BIAS for seven selected quantiles and six sample sizes. On the average, for the seven quantiles and nine sample sizes, the BIAS values for MOM, MLE, MME, and MALS were 0.015, -0.069,

-0.048, and 0.008, respectively. The MALS performed the best and the MLE the worst. The MALS reduced the BIAS by 47 percent as compared with MOM, 88 percent as compared with MLE, and 83 percent as compared with MME.

To clearly show the relative average performances of the four methods for the six sample sizes, MALS was compared with MOM, MLE, and MME for prediction of the seven quantiles. Table 14 shows the computed results, where the compared values were computed by {RMSE(MALS)-RMSE(XXX)}/RMSE(XXX) for the RMSE and {|BIAS(MALS)|-|BIAS(XXX)|}/|BIAS(XXX)| for the BIAS, in which XXX=MOM, MLE, and MME respectively. From Table 14, one can see that for the return periods 50, 100, and 200 years, MALS reduced RMSE by 7, 11, and 15 percent, and reduced BIAS by 71, 71, and 60 percent respectively as compared with MOM. Therefore, the MALS predicts flood quantiles more accurately for larger return period. Although MLE and MME performed better than MALS in terms of RMSE, MALS reduced the BIAS by at least 90 percent for predicting quantiles at larger return periods as compared with MLE and MME. Table 13 shows that MLE and MME under-estimates the flood quantiles for larger return periods (say larger or equal to 50 years) which is shown by the larger negative BIAS values. Moreover, when sample size gets larger, RMSE and BIAS becomes smaller for both MOM and MALS. This is not true for MLE and MME. The reason for this is that for certain samples, MLE and MME cannot yield a solution *[6]*. Two other advantages of the MALS method are: (1) when different skew correction factors are used, MALS always yields the same RMSE and BIAS, but MOM does not; (2) for large sample sizes (say 500), MALS produces nearly the same RMSE and BIAS regardless of the number of samples used, but the RMSE computed from MOM is largely influenced by the number of samples used.

Predicted Quantiles for the 90 Louisiana Stations

Quantiles for the return periods 2, 5, 10, 25, 50, 100, and 200 years are computed using the MALS method and the MOM, and are listed in Tables 15 and 16 respectively for all of the 90 Louisiana stations. Since the MALS method has been shown to perform significantly better than the MOM for estimating parameters of the LP3 distribution, it is expected that the predicted quantiles by using the MALS method are more accurate.

Three statistical tests were conducted to examine the predicted quantiles by the MALS and the MOM for the return periods of 25, 50, and 100 years for the 90 gauge stations. These three tests are: (1) the t-test, to test whether the population means given by the two methods are significantly different; (2) the F-test, to test whether variances given by the two methods are significantly different; and (3) the Kolmogorov-Smirnov (K-S) test, to test whether the quantiles predicted by the two methods are significantly different. The principles and computational procedures for the above three tests have been described by Press, et al, *(33)*. At the 0.01 significant level, all of the three tests showed that no significant difference in population mean, variance, and predicted LP3 quantiles by the two methods. This is no surprise because the MALS does not completely change the MOM-estimated parameters of the LP3 distribution. In fact, the MALS keeps the same estimated mean and variance as those estimated by MOM but improves the coefficient of skewness using the optimization method.

Summary of the MALS Method

Based on the 90 observed flood data and 690 Monte Carlo simulated data, the following conclusions are drawn:

(1) The MALS method yielded the smallest BIAS or RAB for both types of data as compared with MOM, MLE, and MME.

(2) The MALS method reduced the RMSE by 13.6 percent for the 90 Monte Carlo simulated data and by 8 percent for the 600 samples, and reduced the

RRASE by 13.5 percent for the 90 observed flood data as compared with MOM.

(3) The MALS method reduced BIAS by 33 percent for the 90 Monte Carlo simulated data and by 47 percent for the 600 samples, and reduced RAB by 46 percent for the 90 sets of observed flood data, as compared with MOM.

(4) MLE and MME gave approximately the same accuracy in quantile prediction. They were the two best estimators for 690 Monte Carlo simulated data in terms of RMSE, but the two worst estimators for both the 90 observed flood data and the 690 Monte Carlo simulated samples in terms of RRASE, RAB and BIAS respectively.

(5) For the data used, the MALS method always performed better than MOM regardless of what the performance index was used. For all tests and methods, MALS yielded the best performance for predicting flood quantiles for return periods 50, 100, and 200 years. Thus, the MALS method can be considered as a potential candidate to replace MOM for estimating parameters of the LP3 distribution.

Table 5. Twenty Alternative Combination Methods

Method	Parameter Estimated by		Starting Point	Objective Function
	MOM	LSM		
MALS1	(μ_y, σ_y)	(γ_y)	(G_y)	25A
MALS2	(μ_y, σ_y)	(γ_y)	(G_y)	25B
MALS3	(μ_y, σ_y)	(γ_y)	(G_y)	25C
MALS4	(μ_y, σ_y)	(γ_y)	(G_y)	25D*
MALS5	(μ_y, σ_y)	(γ_y)	(G_y)	25E
MALS6	(μ_y)	(σ_y, γ_y)	(S_y, G_y)	25A
MALS7	(μ_y)	(σ_y, γ_y)	(S_y, G_y)	25B
MALS8	(μ_y)	(σ_y, γ_y)	(S_y, G_y)	25C
MALS9	(μ_y)	(σ_y, γ_y)	(S_y, G_y)	25D
MALS10	(μ_y)	(σ_y, γ_y)	(S_y, G_y)	25E
MALS11		$(\mu_y, \sigma_y, \gamma_y)$	$(\bar{y}, S_y, G_y)$	25A
MALS12		$(\mu_y, \sigma_y, \gamma_y)$	$(\bar{y}, S_y, G_y)$	25B
MALS13		$(\mu_y, \sigma_y, \gamma_y)$	$(\bar{y}, S_y, G_y)$	25C
MALS14		$(\mu_y, \sigma_y, \gamma_y)$	$(\bar{y}, S_y, G_y)$	25D
MALS15		$(\mu_y, \sigma_y, \gamma_y)$	$(\bar{y}, S_y, G_y)$	25E
MALS16		(a, b, c)	$(\hat{a}, \hat{b}, \hat{c})$	25A
MALS17		(a, b, c)	$(\hat{a}, \hat{b}, \hat{c})$	25B
MALS18		(a, b, c)	$(\hat{a}, \hat{b}, \hat{c})$	25C
MALS19		(a, b, c)	$(\hat{a}, \hat{b}, \hat{c})$	25D
MALS20		(a, b, c)	$(\hat{a}, \hat{b}, \hat{c})$	25E

*Finally selected method.

Table 6. Average Standard RMSE and BIAS for 20 Combination Methods Using 10 Samples of Size 40

Method	Return Period (Year)						
	2	5	10	25	50	100	200
RMSE1	0.047	0.072	0.112	0.185	0.251	0.328	0.415
BIAS1	0.013	0.028	0.047	0.081	0.112	0.149	0.192
RMSE2	0.047	0.072	0.112	0.188	0.258	0.340	0.433
BIAS2	0.010	0.026	0.048	0.086	0.122	0.164	0.214
RMSE3	0.099	0.092	0.111	0.173	0.230	0.286	0.340
BIAS3	0.092	0.060	0.009	-0.065	-0.120	-0.174	-0.224
RMSE4	0.056	0.083	0.109	0.149	0.181	0.212	0.242
BIAS4	0.038	0.044	0.041	0.033	0.026	0.020	0.013
RMSE5	0.053	0.080	0.111	0.160	0.199	0.238	0.278
BIAS5	0.034	0.041	0.042	0.042	0.043	0.044	0.047
RMSE6	0.056	0.120	0.170	0.240	0.297	0.357	0.420
BIAS6	0.036	0.089	0.114	0.140	0.157	0.173	0.192
RMSE7	0.062	0.121	0.164	0.226	0.278	0.333	0.392
BIAS7	0.036	0.083	0.105	0.128	0.143	0.159	0.176
RMSE8	0.050	0.117	0.173	0.253	0.316	0.382	0.452
BIAS8	0.030	0.086	0.117	0.153	0.180	0.207	0.235
RMSE9	0.052	0.103	0.144	0.203	0.249	0.298	0.348
BIAS9	0.035	0.072	0.089	0.106	0.117	0.129	0.140
RMSE10	0.050	0.100	0.148	0.218	0.275	0.335	0.398
BIAS10	0.029	0.070	0.092	0.120	0.141	0.163	0.187
RMSE11	0.060	0.110	0.154	0.218	0.270	0.325	0.382
BIAS11	0.043	0.079	0.097	0.118	0.133	0.149	0.166
RMSE12	0.060	0.121	0.164	0.226	0.279	0.336	0.396
BIAS12	0.039	0.084	0.105	0.128	0.144	0.161	0.179
RMSE13	0.068	0.105	0.141	0.195	0.239	0.285	0.332
BIAS13	0.055	0.071	0.077	0.084	0.089	0.095	0.102
RMSE14	0.053	0.104	0.152	0.219	0.272	0.327	0.385
BIAS14	0.036	0.067	0.082	0.101	0.115	0.129	0.145
RMSE15	0.051	0.103	0.153	0.224	0.283	0.344	0.409
BIAS15	0.030	0.071	0.093	0.122	0.144	0.167	0.192
RMSE16	0.070	0.105	0.139	0.192	0.235	0.279	0.325
BIAS16	0.056	0.069	0.073	0.077	0.080	0.085	0.090
RMSE17	0.081	0.121	0.159	0.216	0.264	0.313	0.364
BIAS17	0.069	0.087	0.095	0.103	0.110	0.118	0.127
RMSE18	0.070	0.105	0.139	0.192	0.235	0.279	0.325
BIAS18	0.056	0.069	0.073	0.075	0.081	0.085	0.090
RMSE19	0.054	0.085	0.118	0.168	0.208	0.250	0.292
BIAS19	0.037	0.045	0.046	0.047	0.048	0.050	0.053
RMSE20	0.051	0.100	0.147	0.214	0.267	0.323	0.380
BIAS20	0.033	0.056	0.066	0.079	0.090	0.101	0.114

Table 7. Average Standard RMSE and BIAS for 20 Combination Methods Using 10 Samples of Size 100

Method	Return Period (Year)						
	2	5	10	25	50	100	200
RMSE1	0.086	0.092	0.082	0.110	0.188	0.320	0.516
BIAS1	-0.009	-0.005	0.017	0.062	0.112	0.180	0.272
RMSE2	0.067	0.109	0.105	0.118	0.214	0.435	0.868
BIAS2	-0.010	-0.017	0.008	0.068	0.142	0.256	0.441
RMSE3	0.115	0.197	0.126	0.196	0.566	1.195	2.255
BIAS3	-0.104	-0.187	-0.101	0.172	0.546	1.158	2.172
RMSE4	0.088	0.080	0.065	0.057	0.073	0.108	0.155
BIAS4	0.021	0.021	0.018	0.013	0.010	0.009	0.010
RMSE5	0.089	0.080	0.066	0.058	0.076	0.113	0.162
BIAS5	0.020	0.020	0.018	0.016	0.016	0.017	0.021
RMSE6	0.064	0.086	0.105	0.140	0.175	0.214	0.257
BIAS6	0.041	0.072	0.083	0.093	0.098	0.104	0.110
RMSE7	0.087	0.084	0.074	0.118	0.204	0.332	0.510
BIAS7	-0.001	0.017	0.038	0.080	0.123	0.181	0.258
RMSE8	0.088	0.088	0.086	0.110	0.156	0.224	0.316
BIAS8	0.015	0.043	0.059	0.080	0.098	0.119	0.146
RMSE9	0.089	0.084	0.073	0.072	0.094	0.135	0.191
BIAS9	0.019	0.033	0.038	0.044	0.049	0.056	0.066
RMSE10	0.088	0.084	0.073	0.075	0.102	0.148	0.209
BIAS10	0.017	0.032	0.039	0.048	0.056	0.067	0.080
RMSE11	0.093	0.088	0.089	0.127	0.190	0.280	0.400
BIAS11	0.014	0.042	0.059	0.083	0.106	0.133	0.169
RMSE12	0.085	0.084	0.075	0.119	0.202	0.331	0.509
BIAS12	0.006	0.023	0.040	0.080	0.122	0.180	0.257
RMSE13	0.087	0.083	0.074	0.080	0.110	0.159	0.225
BIAS13	0.030	0.037	0.040	0.044	0.050	0.058	0.070
RMSE14	0.086	0.083	0.075	0.080	0.108	0.157	0.223
BIAS14	0.019	0.034	0.041	0.051	0.060	0.071	0.085
RMSE15	0.088	0.083	0.072	0.076	0.104	0.150	0.213
BIAS15	0.018	0.032	0.039	0.048	0.056	0.067	0.080
RMSE16	0.090	0.082	0.072	0.081	0.115	0.170	0.243
BIAS16	0.031	0.036	0.038	0.042	0.048	0.057	0.070
RMSE17	0.109	0.106	0.091	0.103	0.167	0.277	0.439
BIAS17	0.038	0.044	0.050	0.064	0.082	0.110	0.152
RMSE18	0.089	0.083	0.073	0.078	0.109	0.159	0.226
BIAS18	0.032	0.036	0.037	0.041	0.045	0.053	0.064
RMSE19	0.085	0.079	0.067	0.065	0.089	0.132	0.191
BIAS19	0.020	0.022	0.021	0.022	0.024	0.029	0.037
RMSE20	0.086	0.077	0.065	0.068	0.100	0.151	0.219
BIAS20	0.020	0.023	0.023	0.026	0.030	0.037	0.047

Table 8. Average Standard RMSE for Seven Selected Quantiles Using 10 Samples for Each Sample Size

Method	N	2	5	10	25	50	100	200
		Return Period (Year)						
MOM	15	0.185	0.180	0.205	0.286	0.375	0.487	0.629
MLE	15	0.123	0.159	0.183	0.211	0.230	0.250	0.269
MME	15	0.129	0.169	0.190	0.213	0.229	0.245	0.262
MALS	15	0.150	0.172	0.191	0.225	0.256	0.289	0.324
MOM	20	0.182	0.161	0.142	0.164	0.218	0.292	0.385
MLE	20	0.133	0.135	0.137	0.145	0.158	0.175	0.195
MME	20	0.138	0.141	0.139	0.144	0.156	0.174	0.197
MALS	20	0.156	0.142	0.141	0.156	0.213	0.281	0.347
MOM	25	0.149	0.152	0.149	0.167	0.200	0.247	0.306
MLE	25	0.115	0.132	0.144	0.165	0.183	0.203	0.224
MME	25	0.121	0.135	0.143	0.161	0.181	0.204	0.229
MALS	25	0.130	0.139	0.146	0.167	0.191	0.219	0.251
MOM	30	0.066	0.048	0.067	0.126	0.181	0.241	0.305
MLE	30	0.062	0.051	0.044	0.051	0.071	0.094	0.120
MME	30	0.071	0.062	0.053	0.064	0.089	0.120	0.153
MALS	30	0.066	0.052	0.065	0.110	0.152	0.199	0.248
MOM	40	0.052	0.080	0.112	0.160	0.199	0.238	0.279
MLE	40	0.072	0.085	0.090	0.102	0.117	0.136	0.158
MME	40	0.077	0.094	0.098	0.112	0.130	0.154	0.181
MALS	40	0.056	0.083	0.109	0.149	0.181	0.212	0.244
MOM	60	0.047	0.054	0.064	0.089	0.115	0.145	0.178
MLE	60	0.067	0.060	0.046	0.041	0.056	0.080	0.108
MME	60	0.076	0.069	0.081	0.051	0.074	0.107	0.143
MALS	60	0.049	0.055	0.063	0.086	0.110	0.140	0.173
MOM	80	0.063	0.064	0.067	0.086	0.109	0.135	0.187
MLE	80	0.077	0.068	0.055	0.056	0.074	0.101	0.131
MME	80	0.081	0.074	0.060	0.061	0.082	0.113	0.147
MALS	80	0.066	0.070	0.066	0.079	0.097	0.119	0.143
MOM	100	0.087	0.080	0.067	0.064	0.084	0.123	0.173
MLE	100	0.074	0.070	0.073	0.093	0.117	0.145	0.175
MME	100	0.084	0.074	0.065	0.091	0.102	0.134	0.169
MALS	100	0.088	0.080	0.065	0.057	0.073	0.108	0.155
MOM	500	0.021	0.029	0.037	0.053	0.066	0.081	0.096
MLE	500	0.038	0.033	0.069	0.120	0.158	0.194	0.228
MME	500	0.043	0.034	0.039	0.076	0.112	0.150	0.188
MALS	500	0.022	0.030	0.037	0.048	0.059	0.070	0.083
AVG:								
MOM:		0.095	0.094	0.101	0.133	0.172	0.221	0.282
MLE:		0.084	0.088	0.093	0.109	0.129	0.153	0.179
MME:		0.091	0.095	0.097	0.108	0.128	0.156	0.185
MALS:		0.087	0.091	0.098	0.122	0.149	0.182	0.219

Table 9. Average Standard BIAS for Seven Selected Quantiles Using 10 Samples for Each Sample Size

Method	N	Return Period (Year) 2	5	10	25	50	100	200
MOM	15	0.035	0.023	0.031	0.065	0.110	0.173	0.257
MLE	15	0.055	0.042	0.021	-0.013	-0.040	-0.067	-0.095
MME	15	0.062	0.063	0.044	0.010	-0.020	-0.051	-0.084
MALS	15	0.053	0.054	0.044	0.030	0.020	0.012	0.007
MOM	20	0.074	0.049	0.024	-0.003	-0.014	-0.016	-0.008
MLE	20	0.070	0.050	0.023	-0.017	-0.049	-0.081	-0.114
MME	20	0.078	0.064	0.036	-0.009	-0.045	-0.082	-0.012
MALS	20	0.066	0.053	0.037	0.016	0.004	-0.003	-0.007
MOM	25	0.045	0.027	0.012	-0.002	-0.008	-0.008	-0.002
MLE	25	0.055	0.032	0.004	-0.035	-0.066	-0.098	-0.129
MME	25	0.065	0.044	0.013	-0.035	-0.072	-0.110	-0.148
MALS	25	0.048	0.034	0.018	-0.003	-0.018	-0.032	-0.044
MOM	30	0.030	0.034	0.034	0.035	0.038	0.044	0.053
MLE	30	0.047	0.035	0.014	-0.018	-0.045	-0.072	-0.100
MME	30	0.058	0.048	0.022	-0.021	-0.057	-0.094	-0.131
MALS	30	0.036	0.038	0.033	0.024	0.017	0.012	0.008
MOM	40	0.034	0.042	0.042	0.042	0.041	0.043	0.042
MLE	40	0.053	0.043	0.021	-0.014	-0.044	-0.074	-0.104
MME	40	0.062	0.055	0.029	-0.013	-0.048	-0.085	-0.121
MALS	40	0.038	0.044	0.041	0.033	0.026	0.020	0.013
MOM	60	0.039	0.044	0.042	0.038	0.034	0.030	0.027
MLE	60	0.057	0.048	0.026	-0.009	-0.037	-0.066	-0.095
MME	60	0.069	0.057	0.052	-0.017	-0.055	-0.093	-0.132
MALS	60	0.041	0.045	0.041	0.034	0.027	0.021	0.016
MOM	80	0.040	0.044	0.040	0.032	0.026	0.019	0.039
MLE	80	0.060	0.048	0.024	-0.016	-0.048	-0.081	-0.114
MME	80	0.067	0.056	0.029	-0.016	-0.053	-0.091	-0.129
MALS	80	0.043	0.052	0.039	0.028	0.018	0.007	-0.003
MOM	100	0.018	0.020	0.019	0.019	0.020	0.031	0.027
MLE	100	0.039	0.026	0.005	-0.027	-0.054	-0.080	-0.106
MME	100	0.051	0.036	0.006	-0.030	-0.078	-0.117	-0.155
MALS	100	0.021	0.021	0.018	0.013	0.010	0.009	0.010
MOM	500	0.005	0.003	0.002	0.001	0.001	0.000	0.000
MLE	500	0.027	-0.018	-0.053	-0.097	-0.129	-0.158	-0.187
MME	500	0.040	0.018	-0.014	-0.063	-0.102	-0.141	-0.180
MALS	500	0.006	0.004	0.001	-0.001	-0.003	-0.005	-0.006
AVG:								
MOM:		0.035	0.032	0.027	0.026	0.032	0.041	0.051
MLE:		0.051	0.038	0.021	0.027	0.057	0.086	0.116
MME:		0.061	0.049	0.027	0.024	0.059	0.096	0.121
MALS:		0.039	0.038	0.030	0.020	0.016	0.014	0.013

Table 10. RRASE for 90 Sets of Louisiana Flood Data for Four Estimation Methods

STATION NUMBER	MOM	MLE	MME	MALS
2492000	0.1315	0.1458	0.1413	0.1084
2492360	0.1126	0.1214	0.1132	0.1090
2490105	0.1233	0.1358	0.1224	0.1133
7378500	0.0938	0.1121	0.1081	0.0761
7375222	0.1648	0.2351	0.2391	0.1502
7380160	0.1124	0.1265	0.1259	0.1105
7375170	0.1290	0.1454	0.1400	0.1320
7376000	0.1102	0.1383	0.1423	0.0863
7376500	0.0767	0.0854	0.0809	0.0612
7375500	0.0997	0.1233	0.1169	0.0769
7377300	0.1037	0.1085	0.1055	0.1030
7376600	0.0844	0.1370	0.1386	0.0637
7375480	0.2536	0.2568	0.2497	0.2450
7375000	0.1250	0.1578	0.1650	0.1043
2491500	0.1310	0.1765	0.1797	0.1019
2491700	0.5235	0.5093	0.5156	0.4680
2491350	0.1745	0.1832	0.1748	0.1720
7377000	0.1505	0.1909	0.1961	0.1328
7375800	0.1194	0.1281	0.1245	0.1143
7375307	0.1983	0.2168	0.2035	0.1902
7378000	0.1044	0.1608	0.1637	0.0796
7377500	0.1427	0.1759	0.1793	0.1365
7373500	0.1391	0.1581	0.1515	0.1324
7369500	0.0455	0.0787	0.0778	0.0340
7370000	0.0949	0.1259	0.1259	0.0786
7386500	0.0769	0.0969	0.0961	0.0654
7364500	0.2075	0.3101	0.3148	0.1645
7386500	0.1873	0.2554	0.2608	0.1534
8012000	0.1084	0.1239	0.1251	0.1009
8010000	0.1646	0.2480	0.2529	0.1182
8015500	0.1369	0.1441	0.1386	0.1161
8011800	0.1644	0.2036	0.2090	0.1476
8013500	0.1024	0.1200	0.1189	0.0815
8014500	0.1492	0.1563	0.1475	0.1231
8014000	0.1463	0.1505	0.1447	0.1385
8014200	0.1546	0.1702	0.1619	0.1277
8013000	0.1521	0.2028	0.2071	0.1271
7382000	0.1641	0.1841	0.1954	0.1319
7381800	0.1390	0.1682	0.1721	0.1149
7373000	0.2014	0.2125	0.2158	0.1981
7353500	0.2501	0.2689	0.2766	0.2376
7372500	0.1804	0.1874	0.1902	0.1599
7372200	0.1435	0.1976	0.2028	0.1092
7370750	0.0949	0.1105	0.1049	0.0887
7372110	0.1967	0.2480	0.2419	0.2011
7372000	0.2205	0.3930	0.4027	0.1511
7370500	0.3132	0.5118	0.5281	0.2277

Table 14. Average Relative Improvement of RMSE and BIAS by MALS over MOM, MLE, and MME for Six Sample Sizes

Compared Method	{RMSE(MALS)-RMSE(XXX)}/RMSE(XXX) where XXX=MOM, MLE, and MME Respectively, for the Quantiles at the Following Return Periods:						
	2-YR	5-YR	10-YR	25-YR	50-YR	100-YR	200-YR
MOM	0.000	0.022	0.000	-0.040	-0.066	-0.109	-0.151
MLE	-0.085	-0.078	-0.048	0.031	0.104	0.186	0.279
MME	-0.096	-0.059	0.034	0.160	0.247	0.330	0.427
Compared Method	{|BIAS(MALS)|-|BIAS(XXX)|}/|BIAS(XXX)| where XXX=MOM, MLE, and MME Respectively, for the Quantiles at the Following Return Periods:						
	2-YR	5-YR	10-YR	25-YR	50-YR	100-YR	200-YR
MOM	0.030	0.333	0.250	-0.571	-0.714	-0.708	-0.600
MLE	-0.629	1.667	-0.828	-0.959	-0.963	-0.950	-0.908
MME	-0.667	-0.636	0.000	-0.936	-0.951	-0.939	-0.893

Table 15. Predicted Quantiles for 90 Louisiana Stations Using Log-Pearson Type 3 Distribution With MALS Estimation Method

Station	Return Period (year) 2	5	10	25	50	100	200
2492000	20224	37999	52343	73109	90358	109034	129208
2492360	5984	10190	13639	18797	23253	28263	33895
2490105	2147	4320	6341	9677	12813	16579	21078
7378500	27692	49640	66729	90821	110400	131249	153437
7375222	2367	3853	4736	5708	6331	6876	7356
7380160	1077	1647	2036	2531	2900	3269	3639
7375170	3795	6427	8650	12071	15111	18611	22642
7376000	5449	10008	13506	18343	22190	26210	30404
7376500	3155	5056	6461	8381	9909	11514	13206
7375500	13541	26082	36275	51067	63354	76642	90969
7377300	24594	40994	54543	75009	92881	113176	136229
7376600	1365	1823	2046	2259	2379	2475	2552
7375480	5754	12792	19855	32258	44546	59923	79023
7375000	4222	9225	13631	20387	26246	32778	40009
2491500	22062	39527	51937	67904	79761	91459	103002
2491700	2612	7134	11801	19845	27512	36687	47506
2491350	2050	4422	6857	11265	15777	21598	29060
7377000	20931	42052	58769	82135	100767	120179	140324
7375800	3610	8292	13109	21744	30448	41492	55386
7375307	3294	8293	13518	22855	32158	43787	58157
7378000	11530	18018	21865	26139	28920	31389	33594
7377500	7043	13130	17765	24101	29080	34222	39526
7373500	6307	10971	14505	19382	23275	27363	31657
7369500	2768	3463	3810	4156	4363	4536	4684
7370000	5286	7281	8410	9648	10452	11171	11821
7386500	1075	1413	1602	1807	1940	2060	2168
7364500	7535	9407	9948	10236	10309	10329	10330
7386500	1236	1685	1876	2033	2108	2159	2193
8012000	7529	12054	16012	22341	28188	35164	43496
8010000	5096	7261	8298	9254	9770	10161	10458
8015500	27826	49016	65415	88477	107198	127134	148356
8011800	2367	4279	5641	7395	8698	9983	11251
8013500	14904	26093	34511	46039	55162	64671	74585
8014500	11664	24000	34777	51387	65944	82377	100827
8014000	4093	8044	11784	18099	24178	31637	40755
8014200	3721	7536	10774	15635	19792	24389	29449
8013000	14275	25964	34087	44264	51614	58688	65497
7382000	1370	2482	3725	6226	9101	13246	19234
7381800	2185	4242	5903	8296	10266	12380	14642
7373000	2797	7308	12170	21082	30156	41699	56193
7353500	1575	4621	8140	14920	22093	31473	43541
7372500	3223	6054	8865	13882	18992	25597	34114
7372200	20809	41362	57124	78501	95058	111894	128955
7370750	1909	3514	4977	7378	9635	12356	15630
7372110	1948	5086	8989	17411	27515	42416	64175
7372000	8456	14498	17577	20402	21878	22945	23706
7370500	5422	10318	13009	15589	16983	18014	18766

Table 15. Predicted Quantiles for 90 Louisiana Stations Using Log-Pearson Type 3 Distribution With MALS Estimation Method(Cont'd)

	Return Period (year)						
Station	2	5	10	25	50	100	200
7371500	7074	13550	18101	23792	27869	31755	35452
7352000	2611	5318	7631	11123	14124	17457	21140
7366420	3119	7574	12437	21636	31369	44228	61052
7365000	6094	11464	15501	20938	25143	29426	33782
7364870	2456	4714	6063	7480	8326	9012	9565
7365500	2485	5040	7826	13248	19236	27527	38973
7366000	5921	11705	16615	20425	30407	37519	45411
7366200	3691	7047	9441	12483	14700	16843	18913
7364700	2177	5065	8909	18045	30252	50265	83034
8031000	1398	2416	3198	4294	5182	6127	7132
8016800	3572	6237	8311	11250	13653	16231	18994
8030000	2174	3580	4598	5956	7009	8091	9205
8016400	3855	6812	9285	13036	16313	20025	24227
8016600	4041	6850	9294	13163	16695	20859	25767
8015000	5397	12598	19902	32749	45438	61240	80738
8028700	770	1231	1614	2197	2712	3302	3980
8029500	2214	4802	7877	14382	22152	33676	50709
8014600	1787	3635	5408	8429	11352	14953	19364
8028000	7619	19051	32179	58281	87240	127104	181427
8013800	1075	2001	2657	3493	4104	4697	5274
8025850	535	1048	1568	2509	3478	4740	6380
8025500	3567	8249	13808	25450	39148	59117	88058
7354000	2779	4194	5005	5883	6442	6930	7359
7353990	3318	7331	10883	16346	21092	26389	32257
8023000	1941	3686	5092	7119	8796	10604	12549
7351700	969	2028	2969	4437	5738	7219	8896
7352500	3778	6846	9524	13742	17557	22009	27192
7351500	5482	9222	10968	12432	13119	13567	13849
7351000	3841	6554	7870	9013	9574	9955	10207
7344450	2828	5934	8711	13074	16965	21421	26492
2490000	1534	3947	6174	9615	12571	15812	19321
7348700	6840	13800	19659	28385	35791	43931	52836
7349500	4545	7854	10116	12938	14978	16950	18858
7348725	1966	3283	3827	4221	4374	4455	4492
7348800	1849	3704	5338	7889	10161	12765	15734
7347000	1392	2015	2479	3127	3654	4222	4835
7373550	221	309	361	419	458	494	527
7364190	4923	6292	6755	7055	7161	7211	7231
7365800	4041	9795	15725	26247	36695	49742	65867
7362100	6336	12956	18726	27608	35388	44169	54028
2489500	44283	65807	81129	101567	117528	134090	151356
8014800	4002	7662	10691	15174	18974	23155	27741
7368000	1942	2544	2783	2968	3051	3104	3136

Table 16. Predicted Quantiles for 90 Louisiana Stations Using Log-Pearson Type 3 Distribution With MOM Estimation Method

Station	Return Period (year) 2	5	10	25	50	100	200
2492000	20056	37926	52642	74380	92793	113063	135307
2492360	6141	10280	13431	17836	22406	25211	29269
2490105	2160	4331	6319	9554	12553	16110	20309
7378500	27738	49665	66637	90465	109747	130212	151918
7375222	2394	3854	4778	5540	6066	6507	6880
7380160	1098	1652	2003	2422	2716	2995	3261
7375170	3737	6383	8703	12404	15804	19829	24591
7376000	5455	10013	13494	18293	22098	26059	30183
7376500	3158	5057	6456	8363	9877	11465	13135
7375500	13487	26063	36375	51481	64144	77945	92937
7377300	24842	41177	54283	73563	89986	108244	128555
7376600	1377	1822	2026	2208	2305	2378	2434
7375480	6302	13111	18672	26640	33132	40005	47245
7375000	4226	9229	13619	20335	26142	32603	39741
2491500	21831	39522	52455	69557	82589	95716	108936
2491700	3105	7133	9891	13046	15063	16786	18245
2491350	2148	4510	6706	10303	13649	17620	22305
7377000	21115	42115	58327	80458	97703	115314	133242
7375800	3577	8264	13170	22122	31304	43131	58233
7375307	3424	8380	13165	21060	28341	36859	46715
7378000	11499	18047	21954	26315	29164	31699	33970
7377500	7050	13138	17750	24033	28951	34011	39213
7373500	6540	11030	14071	17855	20589	23230	25789
7369500	2762	3466	3821	4176	4390	4570	4723
7370000	5270	7286	8441	9718	10556	11311	11996
7386500	1087	1414	1584	1759	1866	1958	2037
7364500	7677	9291	9651	9813	9852	9868	9874
7386500	1274	1660	1786	1867	1897	1914	1923
8012000	7309	11825	16092	23388	30553	39546	50828
8010000	5100	7280	8310	9245	9742	10117	10401
8015500	27297	48760	66246	92068	114039	138366	165274
8011800	2439	4283	5476	6877	7829	8699	9496
8013500	14815	26065	34663	46626	56235	66377	77081
8014500	11337	23799	35405	54467	72233	93359	118339
8014000	4078	8036	11805	18212	24413	32061	41447
8014200	3655	7503	10907	16232	20969	26389	32551
8013000	14284	25997	34098	44196	51446	58389	65030
7382000	1259	2231	3569	6799	11196	18552	30892
7381800	2204	4250	5862	8129	9956	11879	13899
7373000	2862	7359	11987	20085	27971	37620	49276
7353500	1665	4685	7828	13261	18437	24616	31882
7372500	2953	5645	8864	15714	23966	36310	54742
7372200	20910	41413	56902	77593	93386	109229	125086
7370750	1848	3455	5031	7808	10598	14159	18687
7372110	1902	5024	9057	18111	29386	46557	72436
7372000	8487	14562	17580	20282	21667	22663	23379
7370500	5530	10299	12703	14821	15877	16614	17129

Table 16. Predicted Quantiles for 90 Louisiana Stations Using Log-Pearson Type 3 Distribution With MOM Estimation Method(Cont'd)

	Return Period (year)						
Station	2	5	10	25	50	100	200
7371500	6973	13571	18387	24625	29247	33776	38199
7352000	2768	5822	7879	10264	11827	13197	14394
7366420	3121	7581	12437	21599	31269	44015	60649
7365000	6147	11480	15386	20516	24390	28254	32106
7364870	2558	4668	5741	6702	7190	7538	7785
7365500	2363	4859	7853	14247	21951	33451	50564
7366000	5738	11594	16940	25605	33601	43045	54147
7366200	3517	7027	9872	13955	17301	20869	24660
7364700	2003	4664	8743	19931	37056	68773	127454
8031000	1399	2417	3195	4285	5163	6097	7089
8016800	3501	6200	8417	11722	14562	17734	21274
8030000	2179	3582	4589	5922	6949	7997	9071
8016400	3823	6791	9325	13244	16732	20748	25365
8016600	4054	6866	9287	13084	16517	20529	25216
8015000	5509	12686	19659	31408	42547	55934	71877
8028700	739	1201	1630	2350	3045	3907	4972
8029500	2197	4796	7915	14549	22488	34247	51573
8014600	1818	3662	5364	8154	10757	13862	17548
8028000	7667	19131	32119	57586	85437	123289	174174
8013800	1101	2002	2594	3293	3768	4203	4600
8025850	500	1001	1580	2777	4173	6201	9138
8025500	3486	8145	13888	26362	41536	64284	98144
7354000	2818	4192	4927	5672	6116	6483	6789
7353990	3368	7352	10754	15807	20051	24653	29611
8023000	1985	3701	5003	6776	8166	9596	11067
7351700	890	1963	3101	5227	7464	10419	14295
7352500	3808	6871	9486	13524	17109	21223	25937
7351500	5437	9331	11204	12826	13628	14186	14576
7351000	3811	6624	8025	9274	9911	10367	10692
7344450	2780	5904	8805	13540	17923	23101	29181
2490000	1703	3936	5528	7432	8706	9836	10831
7348700	6880	13820	19569	28011	35078	42754	51058
7349500	4524	7858	10166	13076	15201	17271	19289
7348725	2074	3265	3481	3649	3697	3720	3730
7348800	1861	3711	5316	7789	9963	12427	15206
7347000	1386	2016	2486	3139	3672	4244	4861
7373550	223	309	356	408	440	469	493
7364190	4932	6253	6678	6949	7049	7105	7135
7365800	3749	9484	16274	30213	46151	68675	100173
7362100	6264	12918	18868	28266	36710	46444	57601
2489500	44266	65802	81139	101625	117649	134292	151661
8014800	4184	7722	10316	13740	16339	18947	21562
7368000	1950	2549	2775	2945	3019	3066	3097

CHAPTER 5
REGIONAL FLOOD FREQUENCY ANALYSIS

Introduction

At-site flood frequency analysis consists of fitting preselected probability distributions to observed data at an individual gauge station or site and then estimating the quantiles for some given exceedance probabilities. These predicted quantiles can be used to design various types of hydraulic structures. However, the use of the observed data at only a single observation station may result in unreliable estimates because the length of records at a single station is relatively short when compared to the recurrence intervals, which are to be estimated from the data. For example, it may be necessary to estimate the 100-year flood from a station with only 20 to 30 years of records. The observed flood data always contains various sources of errors and the underlying distributions for the observed data are rarely known. Over the years, researchers have been striving to search for a robust probability distribution and a superior parameter estimation method for flood frequency analysis. No single superior probability distribution or parameter estimation method for various types of data has been found to date. Even though some investigators may find a superior estimator for a specific distribution by the Monte Carlo simulation, it is highly possible that the estimator found is not superior for the observed data.

On the other hand, regional frequency analysis consists of fitting preselected probability distributions by using data from a group of stations with similar hydrological conditions. Therefore, regionalization techniques have the advantage of reducing the uncertainty inherent in an individual station with short records. Other advantages of regionalization techniques are the ease of the use of regional quantiles for design purposes as well as their applicability to sites where flood records are not available.

Regional frequency techniques have been proposed by a number of researchers *[10] [11] [12]*. Greis and Wood *[8]* recommended an indexing method similar to that of Dalrymple *[10]*, but with the generalized extreme value (GEV) as the base distribution and probability weighted moments (PWM) as the parameter estimation method. This parameter estimation method, first proposed by Greenwood, et al. *[13]*, has been shown to possess very attractive asymptotic characteristics when used to estimate the parameters of several distributions, especially in cases where the samples exhibit wide variability *[14]*. In support of this, Potter and Lettenmaier *[15]* tested ten commonly used frequency methods and found that the GEV index method possessed predictive characteristics superior to the other methods tested.

Although most parameter estimation methods are based on some statistical principles such as maximum likelihood, maximum entropy, principle of moments, and least squares of error, the criteria to evaluate the performance of a parameter estimation method for fitting a selected distribution to observed data are usually similar. The error and bias between the calculated and the observed should be minimized. The relative root average square error (RRASE) and the relative average bias (RAB) are examples of performance indices that are frequently used in flood frequency analysis. These two indices are defined for the regional frequency analysis as:

$$RRASE = \sqrt{\frac{1}{m+n}\sum_{i=1}^{m}\sum_{j=1}^{n}\left[\frac{x_c(i,j) - x_o(i,j)}{x_o(i,j)}\right]^2} \tag{30}$$

$$RAB = \frac{1}{m+n}\sum_{i=1}^{m}\sum_{j=1}^{n}\left[\frac{x_c(i,j) - x_o(i,j)}{x_o(i,j)}\right] \tag{31}$$

where $x_c(i,j)$ and $x_o(i,j)$ are the computed and the observed quantiles at the i-th site and the j-th plotting position, m is the number of sites in the region, and n is the number of observations at the i-th site. The optimal parameter set is obtained by minimizing both the RRASE and the RAB for the data sets used. The objective function that performed the best for the at-site analysis was also used for the regional analysis. This objective function is:

$$MIN\ z = RRASE + |\ RAB\ | \tag{32}$$

The objectives of this part of the study are:

(1) to develop an indexed regional optimization procedure to estimate the parameters of the GEV distribution by minimizing the objective function of Equation (32),

(2) to compute flood quantiles for commonly used return periods at the 90 stream gauge sites by using the indexed regional optimization procedure, and

(3) to compare the performance of the indexed regional optimization procedure with that of the indexed regional probability weighted moments.

Identification of Homogeneous Regions

In a previous study, Naghavi, et al. *[21]* divided the state into four hydrologically homogeneous regions, based on the topographic maps, geological maps, climatic maps, and soil survey maps. These four regions are shown in Figure 1. In this regional study, these four homogeneous regions along with the 90 flood gauge stations selected previous, are used in this part of the study. There are 26 stations in the southeast region, 33 stations in the southwest regions, 25 stations in the northwest region, and six stations in the northeast region. Some pertinent statistics of the 90 station records have been listed in Tables 1 and 2 respectively.

The GEV Distribution and the PWM Method

The GEV distributions is defined in inverse form as:

$$x(F) = \begin{cases} \xi + \alpha\,\{1 - [-\ln(F)]^{k}\}\,/\,k & , \quad k \neq 0 \\ \xi - \alpha\ \ln[-\ln(F)] & , \quad k = 0 \end{cases} \tag{33}$$

where ξ, α and k are the location, scale, and shape parameters, respectively, and F is the cumulative probability. When k = 0, GEV reduces to extreme value type 1 distribution (EV1); when k < 0, GEV becomes EV2 distribution; and when k > 0, GEV becomes EV3 distribution. The mean, variance and coefficient of skewness for the GEV distribution are related to the distribution parameters as *[17]*:

$$\mu_x = \xi + \alpha (1 - \Omega_1)/k \tag{34}$$

$$\sigma_x^2 = \alpha^2 (\Omega_2 - \Omega_1^2)/k^2 \tag{35}$$

$$\gamma_x = - \frac{k}{|k|} \frac{(\Omega_3 - 3\Omega_2\Omega_1 + 2\Omega_1^3)}{(\Omega_2 - \Omega_1^2)^{1.5}} \tag{36}$$

where $\Omega_r = \Gamma(1+rk)$, r=1,2,3. The three parameters for the regional (dimensionless) GEV distribution can be estimated by the PWM method *[16]* as:

$$\hat{\alpha} = \frac{\hat{k}\,[M_{(0)R} - 2M_{(1)R}]}{\Gamma(1+\hat{k})\ (1 - 2^{-\hat{k}})} \tag{37}$$

$$\hat{\xi} = M_{(0)R} + \hat{\alpha}\,[\Gamma(1+\hat{k}) - 1]\,/\,\hat{k} \tag{38}$$

$$\hat{k} = 7.8590\ C + 2.9554\,C^2 \tag{39}$$

where

$$C = \frac{M_{(0)R} - 2M_{(1)R}}{2M_{(0)R} - 6M_{(1)R} + 3M_{(2)R}} - \frac{ln(2)}{ln(3)} \tag{40}$$

$\Gamma(.)$ is the Gamma function, and $M_{(k)R}$ is the standardized and weighted PWM for a region and is estimated by

$$M_{(k)R} = \frac{1}{\sum_{i=1}^{m} n_i} \sum_{j=1}^{m} \left[\frac{M_{(k)}}{M_{(0)}} \right]_j n_j \quad , \qquad k = 0, 1, 2 \ldots \tag{41}$$

where m is the number of gauge stations in the homogeneous region, n_j is the number of observations at station j, and

$$\hat{M}_{(k)} = \hat{M}_{1,0,k} = \frac{1}{n} \sum_{i=1}^{n} \frac{(i-1)(i-2)\ldots(i-k)}{(n-1)(n-2)\ldots(n-k)} \, x_{n+1-i} \quad , \qquad k=0, 1, 2 \ldots \tag{42}$$

is the k-th unbiased PWM from the observed samples.

The Indexed Regional Probability Weighted Moments (IRPWM) Procedure

The indexed procedure has gained more attention in recent years since the introduction of the probability weighted moments (PWM) by Greenwood et al. *[13]*. It has been used by Greis and Wood *[8]*, Landwehr et al.*[14]*, Wallis *[3]*, Stedinger *[11]*. The index procedure has been applied to the GEV distribution by Hosking, et al. *[16]* and Schaefer *[17]*, and is the recommended procedure in the United Kingdom.

Application of the index procedure to the GEV distribution consists of calculating the PWMs from the observed data at each site within a homogeneous region, using Equation (42). Then, the PWMs are standardized at each site by dividing each PWM by the at-site mean. Each of the standardized PWMs are then averaged over the entire homogeneous region, using Equation (41). These regional averaged standardized PWMs are then used to compute the three parameters of the regional GEV distribution by using Equations (37), (38), and (39). Regional quantiles can then be calculated for any exceedance probability from Equation (33).

These regional quantiles are then rescaled for each site of interest by multiplying by the at-site mean. Once the at-site quantiles are computed, comparisons can be made with any other estimation methods using the performance indices RRASE and RAB.

Development of the Indexed Regional Optimization (IRO) Procedure

Once the three regional parameters of the GEV distribution are estimated by the IRPWM procedure, they serve as the inial values of the parameters for the IRO procedure. In the IRO procedure, the conjugate gradient optimization method (CGO) described in the first part of this study is used to find the optimal set of parameters by minimizing the objective function given in Equation (32). The IRO procedure can be described as follows: First, the regional parameters α, ξ, and k are estimated by the IRPWM procedure. The regional quantiles (dimensionless) are then computed by Equation (33) using the unbiased plotting position formula, $F_j = j/(n+1)$, where j is the j-th smallest value at the i-th site, and n is the number of observations at the i-th site. The corresponding at-site quantiles are obtained by multiplying the regional quantiles by the at-site mean. Second, the objective function value of Equation (32) was evaluated using the estimated and observed at-site quantiles. Third, the CGO search algorithm was applied to search for the optimal regional parameters α, ξ, and k by using the estimated regional parameters from IRPWM and the objective function value. Finally, the regional quantiles at some given recurrence intervals are computed by using Equation (33). By using these optimal parameters, the corresponding at-site quantiles are computed by multiplying the regional quantiles by the at-site mean.

Comparison between the IRPWM Procedure and the IRO Procedure

The IRPWM procedure and the IRO procedure have been applied to the four hydrologically homogeneous regions in Louisiana, shown in Figure 1. The computed RRASE and RAB for southeast region are listed in Table 17, for southwest in Table 18, for northwest in Table 19, and for northeast in Table 20. For each of the four regions, the RAB is always reduced to zero (less than 10^{-4}) by the IRO procedure. The RRASE computed by the IRO procedure is always smaller than that by the IRPWM procedure. The larger the RRASE is from the

IRPWM procedure, the larger the reduction is achieved in RRASE by using the IRO procedure. Figure 4 shows this characteristic. The average RRASE and RAB for each region by the two estimation methods are listed in Table 21. Overall, the IRO procedure reduces the RRASE by 20 percent and reduces RAB by 100 percent as compared with the IRPWM procedure. Thus, the IRO procedure is significantly superior to the IRPWM procedure in terms of performance indices RRASE and RAB.

Quantile Prediction

The three parameters estimated by the IRPWM procedure and the IRO procedure for each of the four homogeneous regions are listed in Table 22. Once these regional parameters are estimated, the regional quantiles for a given cumulative probability F can be generated by using Equation (33). At-site quantile can then be obtained by multiplying the regional quantiles by the at-site mean. The predicted quantiles by the IRO procedure for all of the stations in the four regions for 2-, 10-, 25-, 50-, and 100-year return periods are listed in Table 23 and by the IRPWM procedure in Table 24.

Statistical tests were conducted for the 25-, 50-, and 100-year quantiles estimated by IRO, IRPWM, and MOM for the 90 gauge stations. At 0.01 significant level, the t-test, the F-test, and the Kolmogorov-Smirnov test showed that no significant difference in the population mean, variance, and predicted quantiles among the three methods tested.

Summary of the Regional Frequency Analysis

The results of this part of the study show that the generalized extreme value distribution fitted by the indexed regional optimization (IRO) procedure is better than by the indexed regional probability weighted moments (IRPWM) procedure. The IRO procedure reduces the RRASE by 20 percent and reduces the RAB by 100 percent as compared with the IRPWM procedure. The IRO procedure should be quite useful for any other similar regional frequency studies. It should be noted, however, that the predicted regional quantiles may not be applied outside the physical bounds of the region from which it was calculated.

Table 17. RRASE and RAB Computed by Two Indexed Regional Procedures for the Southeast Region

Station Number	RRASE IRPWM	RRASE IRO	RAB IRPWM	RAB IRO
2492000	0.137	0.219	0.037	-0.038
2492360	0.123	0.195	-0.058	-0.106
2490105	0.210	0.195	0.084	0.013
7378500	0.104	0.212	-0.014	-0.079
7375222	0.165	0.214	-0.040	-0.101
7380160	0.199	0.271	-0.099	-0.141
7375170	0.170	0.223	-0.047	-0.090
7376000	0.104	0.200	0.010	-0.063
7376500	0.163	0.261	-0.083	-0.130
7375500	0.147	0.201	0.054	-0.026
7377300	0.163	0.236	-0.063	-0.109
7376600	0.276	0.346	-0.124	-0.160
7375480	0.475	0.369	0.191	0.100
7375000	0.393	0.298	0.214	0.099
2491500	0.117	0.232	0.003	-0.073
2491700	1.546	1.191	0.693	0.492
2491350	0.357	0.295	0.150	0.072
7377000	0.281	0.214	0.113	0.016
7375800	0.466	0.387	0.282	0.180
7375307	0.764	0.578	0.429	0.294
7378000	0.141	0.247	-0.079	-0.134
7377500	0.184	0.218	0.028	-0.047
7373500	0.122	0.170	-0.037	-0.093
2490000	1.307	0.920	0.626	0.427
7373550	0.245	0.318	-0.122	-0.158
2489500	0.236	0.316	-0.107	-0.145
AVERAGE:	0.482	0.400	0.079	0.000

Table 18. RRASE and RAB Computed by Two Indexed Regional Procedures for the Southwest Region

Station Number	RRASE IRPWM	RRASE IRO	RAB IRPWM	RAB IRO
7386500	0.248	0.291	-0.130	-0.152
8012000	0.208	0.276	-0.095	-0.110
8010000	0.227	0.279	-0.121	-0.148
8015500	0.110	0.170	-0.036	-0.071
8011800	0.193	0.140	-0.014	-0.059
8013500	0.084	0.150	-0.049	-0.085
8014500	0.278	0.176	0.145	0.083
8014000	0.166	0.154	0.035	-0.003
8014200	0.264	0.145	0.112	0.052
8013000	0.197	0.124	0.008	-0.045
7382000	0.265	0.329	0.105	0.084
7381800	0.201	0.107	0.038	-0.014
8031000	0.128	0.172	-0.062	-0.093
8016800	0.143	0.187	-0.040	-0.072
8030000	0.135	0.197	-0.091	-0.118
8016400	0.091	0.165	-0.055	-0.085
8016600	0.170	0.229	-0.083	-0.104
8015000	0.494	0.363	0.252	0.177
8028700	0.197	0.259	-0.101	-0.116
8029500	0.327	0.328	0.179	0.138
8014600	0.227	0.168	0.062	0.018
8028000	0.610	0.511	0.376	0.298
8013800	0.304	0.144	0.035	-0.019
8025850	0.196	0.217	0.049	0.019
8025500	0.434	0.401	0.269	0.217
7354000	0.197	0.246	-0.118	-0.143
7353990	0.496	0.307	0.214	0.135
8023000	0.174	0.124	0.007	-0.038
7351700	0.456	0.375	0.252	0.194
7351500	0.642	0.267	0.104	0.001
7351000	0.649	0.248	0.107	0.006
7344450	0.333	0.188	0.151	0.087
8014800	0.208	0.154	0.009	-0.035
AVERAGE:	0.315	0.248	0.046	0.000

Table 19. RRASE and RAB Computed by Two Indexed Regional Procedures for the Northwest Region

Station	RRASE		RAB	
Number	IRPWM	IRO	IRPWM	IRO
7373000	0.633	0.507	0.325	0.219
7353500	1.049	0.819	0.574	0.429
7372500	0.231	0.287	-0.016	-0.052
7372200	0.142	0.119	-0.011	-0.078
7370750	0.156	0.225	-0.087	-0.124
7373110	0.588	0.525	0.347	0.269
7372000	0.249	0.191	-0.046	-0.124
7370500	0.586	0.273	0.115	-0.004
7371500	0.130	0.153	-0.016	-0.089
7352000	0.139	0.186	-0.011	-0.073
7366420	0.380	0.286	0.218	0.137
7365000	0.156	0.174	-0.056	-0.111
7364870	0.541	0.312	0.105	0.004
7365500	0.232	0.261	0.040	-0.002
7366000	0.253	0.197	0.004	-0.052
7366200	0.177	0.172	0.008	-0.057
7364700	0.463	0.465	0.270	0.215
7352500	0.180	0.243	-0.110	-0.147
7348700	0.119	0.140	-0.016	-0.076
7349500	0.170	0.251	-0.119	-0.164
7348725	0.648	0.333	0.068	-0.042
7348800	0.101	0.135	-0.033	-0.085
7347000	0.327	0.373	-0.185	-0.205
7365800	0.541	0.427	0.362	0.269
7362100	0.095	0.162	0.006	-0.057
AVERAGE:	0.404	0.329	0.070	0.000

Table 20. RRASE and RAB Computed by Two Indexed Regional Procedures for the Northeast Region

Station Number	RRASE		RAB	
	IRPWM	IRO	IRPWM	IRO
7369500	0.070	0.138	-0.017	-0.028
7370000	0.088	0.104	0.022	0.002
7368500	0.049	0.091	0.000	-0.012
7364500	0.262	0.157	0.045	0.014
7368000	0.157	0.059	0.035	0.006
7364190	0.266	0.175	0.046	0.017
AVERAGE:	0.173	0.127	0.022	0.000

Table 21. Average RRASE and RAB Computed by Two Indexed Regional Procedures

Region	IRPWM Procedure		IRO Procedure	
	RRASE	RAB	RRASE	RAB
Southeast	0.482	0.079	0.400	0.000
Southwest	0.315	0.046	0.248	0.000
Northwest	0.404	0.070	0.329	0.000
Northeast	0.173	0.022	0.127	0.000
AVERAGE:	0.345	0.054	0.276	0.000

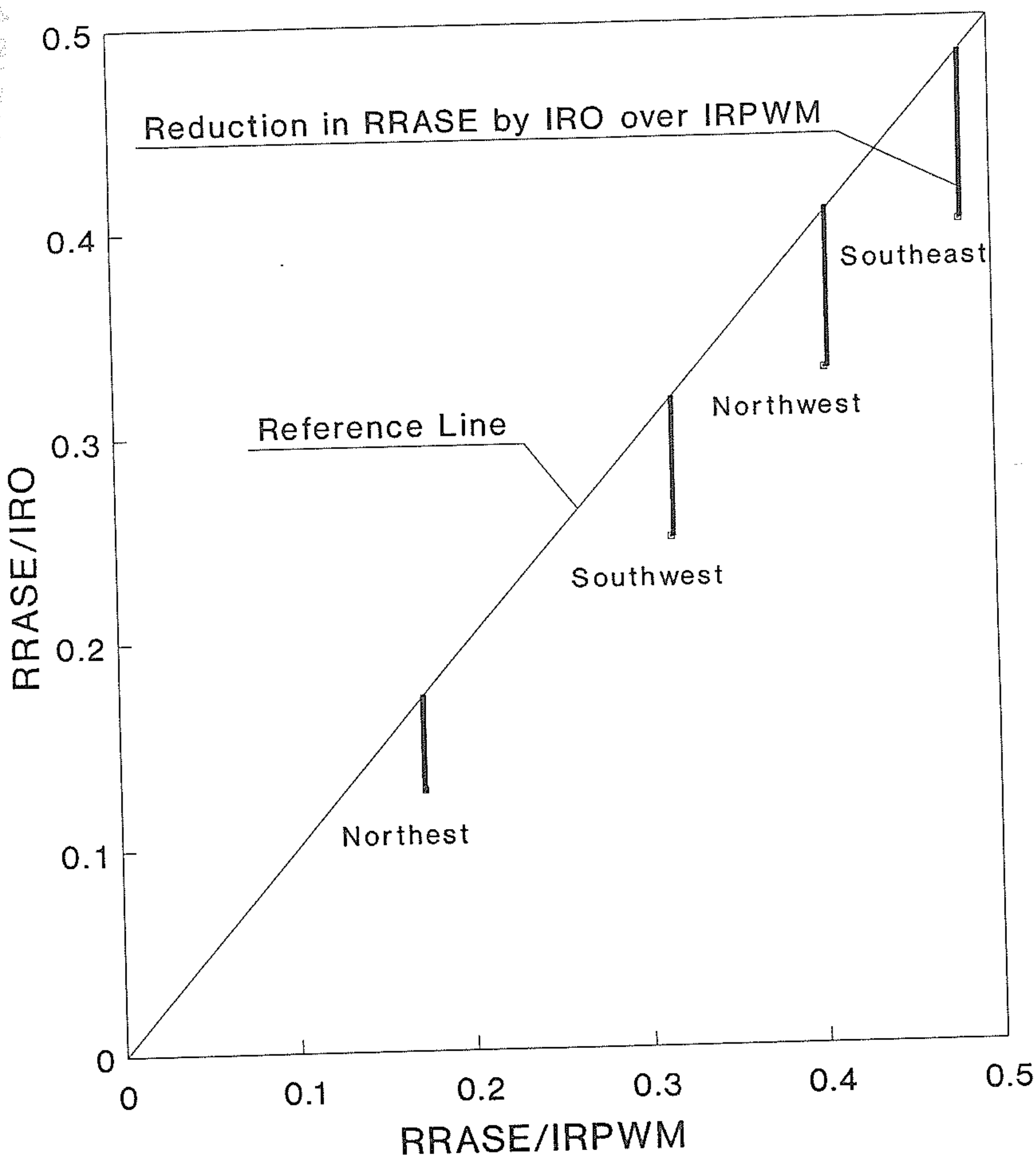

Figure 2. Comparison Between Two Estimation Procedures

Table 22. Regional Parameters Estimated by Two Regional Indexed Procedures

Region	α		ξ		k	
	IRPWM	IRO	IRPWM	IRO	IRPWM	IRO
Southeast	0.457	0.502	0.651	0.623	-0.160	-0.128
Southwest	0.400	0.453	0.601	0.588	-0.302	-0.263
Northwest	0.425	0.459	0.551	0.530	-0.331	-0.307
Northeast	0.342	0.403	0.864	0.859	0.217	0.261

Table 23. Predicted Quantiles for 90 Louisiana Stations by Using the IRO Procedure for the GEV Distribution

	Return Period (year)				
Station	2	10	25	50	100
2492000	21416	50973	68761	83418	99327
2492360	5917	14084	19000	23050	27446
2490105	2440	5808	7835	9505	11318
7378500	27954	66536	89756	108889	129656
7375222	2111	5026	6779	8225	9793
7380160	969	2307	3112	3775	4495
7375170	3804	9056	12217	14821	17647
7376000	5545	13199	17805	21601	25720
7376500	2976	7084	9556	11593	13804
7375500	14484	34474	46504	56417	67177
7377300	24321	57889	78091	94737	112805
7376600	1114	2652	3577	4340	5168
7375480	6948	16537	22309	27064	32226
7375000	4996	11891	16041	19461	23172
2491500	21773	51823	69909	84811	100986
2491700	3490	83061	11205	13593	16186
2491350	2495	5937	8009	9717	11570
7377000	22438	53406	72044	87401	104070
7375800	4782	11383	15355	18628	22181
7375307	4474	10649	14365	17427	20750
7378000	10198	24273	32744	39723	47299
7377500	7144	17004	22938	27828	33135
7373500	6120	14568	19651	23840	28387
7369500	2771	4283	4805	5117	5374
7370000	5492	8486	9520	10137	10648
7368500	1101	1702	1909	2033	2135
7364500	7103	10976	12313	13112	13772
7386500	933	2422	3501	4494	5680
8012000	7134	18514	26761	34352	43410
8010000	3968	10298	14885	19107	24146
8015500	26512	68806	99451	127662	161323
8011800	2168	5627	8133	10440	13192
8013500	13899	36072	52138	66928	84575
8014500	13007	33756	48790	62630	79144
8014000	4414	11456	16559	21256	26860
8014200	3987	10348	14957	19200	24263
8013000	13199	34254	49511	63555	80313
7382000	1705	4424	6394	8208	10372
7381800	2179	5654	8172	10490	13256
7373000	3524	10045	15064	19856	25756
7353500	2169	6183	9273	12223	15855
7372500	3379	9630	14442	19036	24692
7372200	19269	54922	82367	108568	140828
7370750	1828	5211	7815	10301	13361
7372110	2753	7846	11766	15509	20118
7372000	6591	18785	28172	37134	48168
7370500	4531	12914	19367	25528	33113

Table 23. Predicted Quantiles for 90 Louisiana Stations Using the IRO Procedure for the GEV Distribution (Cont'd)

Station	Return Period (year) 2	10	25	50	100
7371500	6370	18156	27228	35890	46554
7352000	2546	7255	10880	14342	18604
7366420	3892	11094	16638	21931	28447
7365000	5449	15531	23292	30701	39824
7364870	2116	6030	9044	11921	15463
7365500	2833	8074	12109	15961	20703
7366000	5885	16773	25154	33156	43008
7366200	3397	9682	14519	19138	24824
7364700	3176	9051	13574	17892	23208
8013000	1298	3368	4869	6250	7898
8016800	3389	8796	12714	16320	20623
8030000	1939	5032	7273	9336	11798
8016400	3732	9687	14001	17972	22711
8016600	3863	10025	14490	18600	23505
8015000	6569	17049	24643	31633	39973
8028700	713	1851	2676	3435	4341
8029500	2932	7608	10997	14117	17839
8014600	1950	5062	7316	9392	11868
8028000	10594	27496	39743	51017	64468
8013800	1007	2614	3778	4850	6128
8025850	602	1563	2259	2899	3664
8025500	4919	12767	18453	23688	29934
7354000	2257	5857	8466	10867	13733
7353990	3689	9575	13839	17765	22449
8023000	1898	4925	7118	9137	11547
7351700	1145	2972	4295	5514	6968
7352500	3465	9876	14811	19523	25324
7351500	4571	11863	17148	22012	27816
7351000	3233	8389	12126	15565	19669
7344450	3153	8182	11826	15181	19184
2490000	1959	4662	6289	7629	9084
7348700	6602	18816	28219	37195	48247
7349500	3786	10791	16183	21331	27669
7348725	1481	4222	6332	8346	10825
7348800	1802	5137	7703	10154	13171
7347000	1106	3151	4725	6228	8079
7373550	188	447	603	732	871
7364190	4696	7256	8140	8668	9104
7365800	5359	15273	22905	30192	39162
7362100	6373	18163	27240	35906	46574
2489500	40208	95702	129099	156619	186489
8014800	3939	10222	14775	18966	23968
7368000	1899	2934	3291	3505	3681

CHAPTER 7

CONCLUSIONS AND RECOMMENDATIONS

The following conclusions are constrained to the study scope of the 90 sets of observed flood data and 690 sets of Monte Carlo simulated data:

(1) The MALS method always performs better than MOM regardless of the type of data or the performance index used. As compared with MOM, the MALS reduced, on the average, the RMSE by 13.6 percent for the 90 sets of Monte Carlo simulated data (preliminary test) and by eight percent for the 600 sets of Monte Carlo simulated data (extended test), and reduced the RRASE by 13.5 percent for the 90 sets of observed data. The MALS reduced BIAS by 33 percent for the 90 sets of Monte Carlo simulated data and by 47 percent for the 600 sets of Monte Carlo simulated data, and reduced the RAB by 47 percent for the 90 sets of observed data.

(2) The MALS yields the smallest BIAS for the 690 sets of Monte Carlo simulated data and the smallest RAB for the 90 sets of observed flood data as compared with MOM, MLE, and MME.

(3) MLE and MME were the best two methods in terms of RMSE but the worst two in terms of BIAS for the 690 sets of Monte Carlo simulated data. They were also found to be the worst two methods in terms of RRASE and RAB for the 90 sets of observed flood data. MLE and MME generally under-estimated the flood quantiles for larger return periods (larger than 50 years).

(4) The MALS predicts flood quantiles more accurately for larger return periods than any other three methods tested.

(5) The MALS yields nearly constant values of RRASE and RAB regardless of what skew-correction factor is used.

(6) When sample size is sufficiently large, for example, 500, the MALS yields nearly constant values of RRASE and RAB, regardless of the number of samples used.

(7) The IRO parameter estimation procedure fitted the generalized extreme value distribution better than the IRPWM procedure. On the average, the IRO procedure reduced the RRASE by 20 percent and the RAB by 100 percent, as compared with the IRPWM procedure for the 90 observed flood data.

(8) The extended IRO procedure was reasonably accurate for predicting flood quantiles at ungauged sites with drainage areas of less than 1000 square miles.

Even though the results in this study indicate that LP3/MALS and GEV/IRO perform reasonably better than other combinations of distributions and estimation methods tested, a more comprehensive Monte Carlo study may be needed for its inclusion in the design procedures.

CHAPTER 8

APPLICATION AND IMPLEMENTATION OF RESULTS

Summary

In the first part of a three-part study, a comprehensive investigation was conducted to find a superior estimation method by using the optimization techniques for the log-Pearson type 3 distribution for at-site flood frequency analysis. In the second part of the study, the selected optimization technique was used to develop a regional estimator for the generalized extreme value distribution (GEV). In the third part of this study, the indexed regional optimization procedure was extended to estimate the flood quantiles at ungauged sites. Ninety sets of observed Louisiana flood data and 690 sets of Monte Carlo simulated data were used for the study. By using conventional flood frequency analysis, five distributions and three estimation methods were used to find the best combination of distribution and method for the Louisiana flood data. The log Pearson type 3 distribution with the method of moments (MOM) was found to provide the best fit to the data.

In order to search for a better estimation method than MOM, 20 combination methods were proposed and tested. The final selection was a combination of the method of moments and the method of optimization (named MALS). By this method, the population mean and variance of the LP3 distribution are estimated by MOM and the population skewness is estimated by the least square method (LSM) with the objective function of minimizing both the relative root average square error (RRASE) and relative average bias (RAB). The MALS performed better than MOM regardless of the type of data or the performance index used. There are several advantages to use MALS: first, MALS predicts flood quantiles more accurately for larger return periods as compared with MOM, MLE, and MME; second, MALS yields a nearly constant RMSE and BIAS when using different bias-correction factors for the coefficient of skewness; third, when the sample size is sufficiently large, the RMSE and BIAS obtained from MALS are nearly the same regardless of the number of samples used; and finally, MALS always yields the smallest BIAS regardless of the type of data used.

In the second part of the study, a combination of the method of probability weighted moments and the method of least squares was developed (named IRO procedure), using the regional index technique. The parameters of the generalized extreme value distribution estimated by the indexed regional probability weighted moments (IRPWM) procedure were used as the initial estimates for the IRO procedure. Computed results show that the IRO procedure yields a smaller RRASE value as compared with the IRPWM procedure for the observed data. Moreover, the IRO procedure reduces the RAB to a nearly zero value (less than 10^{-4}) for the observed data.

In the third part of this study, the IRO procedure was extended to predict flood quantiles at ungauged sites in Louisiana by using the regional regression equations developed by Naghavi, et al. *[21]*. Limited verification showed that the extended estimation procedure was reasonably accurate if the watershed drainage area is between 10 and 1000 square miles.

Significance of Results

Table 27 shows the estimated 100-year quantiles for 11 Louisiana gauge stations at which MOM predicted the 100-year quantiles at least 15 percent larger than those of MALS. On the other hand, Table 28 shows the estimated 100-year quantiles for 10 stations at which MOM predicted the 100-quantiles at least 15 percent smaller than those of MALS. For the stations listed in these tables, the two methods predict significantly different quantiles. For example, the difference in predicted quantiles for the two methods at station 2491700 is as high as 54 percent. Therefore, it is very important to choose an estimation method within a reasonable level of confidence for design work. The following two examples explain how such differences affect the design length of a bridge, consequently affecting the construction cost.

Example 1: The existing bridge at Lawrence Creek (station 2491700) is 458 feet long. The predicted quantiles at 100-year return period by the MALS and MOM are 36687 and

16786 cfs respectively. The corresponding velocities at this site are 7.19 and 3.29 ft/sec. The corresponding design length by MALS is about 500 feet and by MOM, about 400 feet.

Example 2: The existing bridge at Bayou Funny Louis (Station 7372500) is 352 feet long. The predicted quantiles at 100-year return period by the MALS and MOM are 25597 and 36310 cfs respectively. The corresponding velocities at this site are 5.94 and 8.42 ft/sec. The corresponding design length by MALS would probably the same as the existing bridge length, and the design length by MOM would increase to 420 feet.

Final Product Delivery and Training Requirements

The procedures described in this report are available in FORTRAN language on the LaDOTD mainframe system. Users can easily use the computer programs to predict flood quantiles for a set of observed data. The computer programs are well documented and require minimal level of training for use.

REFERENCES CITED

1. U.S. Water Resources Council (USWRC), 1982. Guidelines for determining flood flow frequency. Hydrology Subcommittee, Water Resources Council, Bulletin No. 17B, Reston, Virginia.

2. Bobee, B. and Robitaille, 1977. The use of Pearson type 3 and log-Pearson type 3 distributions revised. Water Resources Research, Vol. 13, No. 2, pp. 427-443.

3. Kuczera, G., 1982. Robust flood frequency models. Water Resources Research, Vol. 18, No. 2, pp. 315-324.

4. Wallis, J. R. and Wood, E. F., 1985. Relative accuracy of log-Pearson III procedures. J. Hydraulic Div., Proc. ASCE, Vol. 111, No. HY7, pp. 1043,1056.

5. Jain, D. and Singh, V. P., 1987. Comparison of some flood frequency distributions using empirical data. Hydrologic Frequency Modeling, V. P. Singh (ed.), D. Reidel Publishing Co., pp. 467-485.

6. Arora, K. and Singh, V. P., 1989. A comparative evaluation of the estimators of the log Pearson type 3 distribution. J. of Hydrology, Vol. 105, pp. 19-37, The Netherlands.

7. Naghavi, B., Singh, V. P. and Yu, F. X., 1991. LADOTD 24-hour rainfall frequency maps and I-D-F curves. Louisiana Transportation Research Center, LTRC Report No. 236, Baton Rouge, LA.

8. Greis, P. and Wood, E., 1981. Regional Flood Frequency Estimation and Network Design, Water Resources Research, Vol. 17, No. 4, pp. 1167-1177.

9. Hosking, J. R. M., 1989. L-Moments: Analysis and Estimation of Distributions Using Linear Combinations of Order Statistics, Journal of Statistical Society, Vol. 5, No. 3.

10. Dalrymple, T., 1960. Flood Frequency Analyses. Manual of Hydrology, Part 3, U.S. Geological Survey Water Supply Paper 1543-A.

11. Stedinger, J., 1983. Estimating a Regional Flood Frequency Distribution, Water Resources Research, Vol. 19, No. 2, 503-510.

12. Kuczera, G., 1982. Combining Site-Specific and Regional Information: An Empirical Bayes Approach, Water Resources Research, Vol. 18, No. 2, pp. 306-314.

13. 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, Vol. 15, No. 5, pp. 1049-1054.

14. Landwehr, J., Matalas, N. and Wallis, J., 1979. Probability Weighted Moments Compared with some Traditional Techniques in Estimating Gumbel Parameters and Quantiles, Water Resources Research, Vol. 15, No. 5, pp. 1055-1064.

15. Potter, K. W., and Lettenmaier, D. P., 1990. A Comparison of Regional Flood Frequency Estimation Methods Using a Resampling Method, Water Resources Research, Vol. 26, No. 3, pp. 415-424.

16. Hosking, J. R. M., Wallis, J. R. and Wood, E. F., 1985. Estimation of the Generalized Extreme Value Distribution by the Method of Probability Weighted Moments, Technometrics, Vol. 27, No. 3, pp. 251-261.

17. Schaefer, M. G., 1990. Regional Analysis of Precipitation Annual Maxima in Washington State, Water Resources Research, Vol. 26, No. 1, pp.119-131.

18. SCS, 1986. Urban hydrology for small watersheds, Technical Release 55, USDA, Washington.

19. Lee, Fred N., 1985. Flood in Louisiana, magnitude and frequency, Water Resources Technical Report No. 36, U.S. Geological Survey, Published by LDOTD, Baton Rouge, LA.

20. Sauer, V. B., W. O. Thomas, Jr., V. A. Striker, and K. V. Wilson, 1983. Flood characteristics of urban watersheds in the United States. U.S. Geological Survey Water Supply Paper 2207.

21. Naghavi, Babak, James F. Cruise, and S. Ekanayake, 1990. Prediction of flood quantiles at ungaged watersheds in Louisiana. LTRC, LDOTD. Report No. 229. 4101 Gourrier Ave., Baton Rouge, LA.

22. Matalas, N. C. and Wallis, J. R., 1973. Eureka! It fits a Pearson type 3 distribution. Water Resources Research, Vol. 9, No. 2, pp. 281-289.

23. Bobee, B., 1975. The log-Pearson type 3 distribution and its application hydrology. Water Resources Research, Vol. 11, No. 5, pp. 681-689.

24. Bobee, B. and Robitaille, 1975. Correction of bias in the estimation of the coefficient of skewness. Water Resources Research, Vol. 11, No. 6, pp. 851-854.

25. Wallis, J. R., Matalas, N. C. and Slack, J. R., 1974. Just a moment!, Water Resources Research, Vol. 10, No. 2, pp. 211-219.

26. Rao, D. V., 1980. Log-Pearson type 3 distribution: a generalized evaluation. J. of Hydraulic Engineering, ASCE, Vol. 106, No. 5, pp. 853-872.

27. Jain, D., 1986. A comparative evaluation of methods of flood frequency analysis and estimation of parameters. Unpublished M.S. thesis, Louisiana State University, Baton Rouge, Louisiana.

28. Rossi, F., M. Fiorentino, and P. Versace, 1984. Two Component Extreme Value Distribution for Flood Frequency Analysis. Water Resources Research, 20(7), 847-856.

29. Houghton, J. C., 1978. The incomplete means estimation procedure applied to flood frequency analysis. Water Resources Research, Vol. 14, No. 6, pp. 1111-1115.

30. Rao, D. V., 1980. Log Pearson type 3 distribution: method of mixed moments. J. of the Hydraulic Division, ASCE, Vol. 106, No. HY6, pp. 999-1019.

31. Wallis, J. R., 1989. Regional Frequency Studies Using L-Moments, IBM Research Report RC14597.

32. Kite, G. W., 1978. Frequency and risk analysis in hydrology. Water Resources Publications, P.O. Box 303, Fort Collins, CO 80522.

33. Press, W. H., Flannery, B. P., Teukolsky, S. A. and Vetterling, W. T., 1986. Numerical Recipes. Cambridge University Press, London, pp. 284-286.

34. Himmelblau, D. M., 1972. Applied Nonlinear Programming. McGraw-Hill Book Co., New York.

35. Naghavi, Babak and Fang Xin Yu, 1991. Generalized Skew Coefficients of Annual Floods for Louisiana Streams. Water Resources Bulletin. Vol. 27, No. 2.

36. International Mathematical and Statistical Library (IMSL) , 1981. Reference Manuals, Vol. 1-4, Ed. 9.

APPENDICES

A: Algorithm of the Conjugate Gradient Optimization (CGO)

Suppose the following generic function is to be minimized:

$$Z = F(x_1, x_2, \ldots, x_n) \tag{A.1}$$

where x_i, $i = 1, 2, \ldots, n$, are n independent variables or parameters. Let ∇F be the gradient vector defined as

$$\nabla F = \left[\frac{\partial F}{\partial x_1} \quad \frac{\partial F}{\partial x_2} \quad \cdots \quad \frac{\partial F}{\partial x_n} \right]^T \tag{A.2}$$

and let $X^i = (x_1^i, x_2^i, \ldots, x_n^i)$ be the minimum point found at the i-th iteration, ϵ_x be the error tolerance for X, and ϵ_f be the error tolerance for F(X). The conjugate gradient optimal search algorithm can be described as follows:

Step 1: Choose a starting point X^0:

Step 2: Compute the conjugate direction:

$$d^i = \nabla F(X^i) + \frac{|\nabla F(X^i)|^2}{|\nabla F(X^{i-1})|^2} d^{i-1} \tag{A.3a}$$

or

$$d^i = \nabla F(X^i) + \frac{|\nabla F(X^i)|^2 + \nabla F(X^i)^T \nabla F(X^{i-1})}{|\nabla F(X^{i-1})|^2} d^{i-1} \tag{A.3b}$$

and

$$d^1 = \nabla F(X^0) \tag{A.3c}$$

where

$$|\nabla F(X^i)|^2 = \sum_{i=1}^{n} \left(\frac{\partial F}{\partial X_i} \right)^2 \tag{A.4}$$

Step 3: Pick up a value t by some one-dimensional search algorithm such that $F(X^i + t\, d^i)$ is minimized.

Step 4: Update the current minimum point found.

$$X^{i+1} = X^i + t_{min}\, d^i \qquad \text{(A.5)}$$

Step 5: Check termination criterion.

If $|\nabla F(X)| < \epsilon_f$

Or $|F(X^{i+1}) - F(X^i)| < \epsilon_f$

Or $|X^{i+1} - X^i| < \epsilon_x$

then stop

Otherwise, repeat beginning with Step 2.

Computation showed that equation (A.3b) is slightly better than equation (A.3a).

B: Algorithm for generating the LP3 random samples

Let NSET be the number of LP3 samples to be generated, N be the sample size, μ_y be the population mean of the log-transformed variable Y, σ_y be the standard deviation of the random variable Y, γ_y be coefficient of skewness, R(i), i=1,2, ..., N, be the i-th random cumulative probability to be generated by the IMSL subroutine RNUN *[36]* for a LP3 random sample, and a, b, and c be the population parameters of the LP3 distribution.

Step 1. Select N, NSET, a, b, and c.

Step 2. Compute μ_y, σ_y, and γ_y by using Equations (2), (3), and (4).

Step 3. Initialize the random-number generator by selecting ISEED=123457 and call the IMSL subroutine RNSET(ISEED).

Step 4. Generate the Nset LP3 samples with size N. This is done by following FORTRAN statements:

```
   DO 10 i=1,NSET
      Call RNUN(N,R)
      DO 5 J=1,N
         Compute the quantile for each random cumulative probability R(i), i=1,2,...,N,
         by using Equations (26) through (29).
 5       CONTINUE
10    CONTINUE
```

This public document is published at a total cost of $ 1,111.08 . One Hundred Seventy-Five (175) copies of this public document were published in this first printing at a cost of $ 653.58 . The total cost of all printings of this document including reprints is $ 1,111.08 . This document was published by Louisiana State University, Graphic Services, 3555 River Road, Baton Rouge, Louisiana 70802, to rep2rt and publish research findings of the Louisiana Transportation Research Center as required in R.S.48:105. This material was printed in accordance with standards for printing by State Agencies established pursuant to R.S.43:31. Printing of this material was purchased in accordance with the provisions of Title 43 of the Louisiana Revised Statutes.