

(i)

COMPARISON OF FLOOD FREQUENCY ANALYSIS

METHODS FOR RIVER TARABA

BY

OVAJIMOH, ABUBAKAR SALIHU (B.ENG. CIVIL ENGINEERING)

A PROJECT REPORT SUBMITTED TO THE POSTGRADUATE
SCHOOL, AHMADU BELLO UNIVERSITY, IN PARTIAL
FULFILMENT OF THE REQUIREMENTS FOR THE AWARD OF
THE DEGREE OF MASTER OF SCIENCE IN WATER RESOURCES
AND ENVIRONMENTAL ENGINEERING.

DEPARTMENT OF WATER RESOURCES AND

ENVIRONMENTAL ENGINEERING

AHMADU BELLO UNIVERSITY

ZARIA, NIGERIA

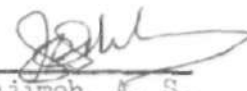
NOVEMBER, 1991

(ii)

DECLARATION

I hereby declare that this project report has been executed
by me and that it is a record of my own research work.

It has not been presented before in any previous application
for a higher degree.


Ova Jimoh, M. S.

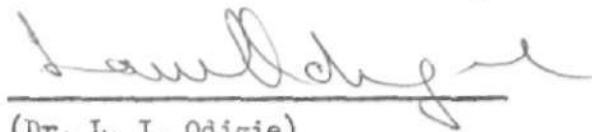
Date: 13/12/91

This declaration is confirmed by


Dr. L. I. Odigie

Date: 2/1/92

This project report entitled "Comparison of Flood Frequency Analysis Methods for River Taraba" by Abubakar Salihu OVAJIMOH meets the regulations governing the award of the degree of Master of Science of Ahmadu Bello University, and is approved for its contribution to knowledge and literary presentation.



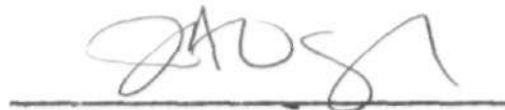
(Dr. L. I. Odigie)
Chairman, Supervisory Committee

Date: 2/1/92



(Prof. J. A. Ogunrombi)
Member, Supervisory Committee

Date: 2/1/92



(Prof. J. A. Ogunrombi)
Head of Department

Date: 2/1/92



Dean, Post-graduate School

Date: 11/3/92

DEDICATION

To God, my wife, Mrs. M. Abubakar
and children (Zuweirat, Rufai and Hashim).

ACKNOWLEDGEMENT

I am grateful to God, the Almighty who has made it possible for me to undertake and complete this thesis.

The enlightening discussions as well as the constructive criticisms of Dr. M. Sowinski, my project supervisor is gratefully acknowledged. The constructive criticisms from Dr. L. I. Odigie, my second supervisor is also gratefully acknowledged.

The encouragement and co-operation from Professor Ogunrombi and Dr. C. O. Okuofu of the Department of Water Resources and Environmental Engineering, Ahmadu Bello University, Zaria are gratefully acknowledged. I am also grateful to Mr. M. A. Ovansa for the assistance given me.

Finally, I am grateful to the Kaduna Polytechnic for sponsoring me for this course.

A B S T R A C T

Water resource designs based on flood are perfect examples of decisions under uncertainty. Three types of uncertainties may exist in a flood design problem - when the classical method is used. These are natural, parameter and model uncertainties.

To reduce the parameter and model uncertainties, the Bayesian procedure has been used. Three common probability distributions (normal, log-normal, and exceedance) were considered. To reduce the parameter uncertainty, the Bayesian procedure was applied to these three distributions to give the corresponding Bayesian probability models. To achieve this objective, flood discharge data collected from Gassol gaging station were taken as the sample information while flood discharge data from eleven other selected gaging stations in the same homogeneous area as Gassol in Upper Benue River Basin were considered to be the prior (regional) information. Natural uncertainty is also reduced by this procedure.

To reduce the model uncertainty, Bayesian procedure was also applied to estimate the marginal likelihoods of the three models. Results showed that only the exceedance model fits the flood discharge data.

Comparison of the classical method with Bayesian method shows that Bayesian method predicts higher flood discharges than the classical method. Also comparison of the regional (prior) method (only in the case of exceedance model) shows that the Bayesian method predicts lower discharges than the regional information.

TABLE OF CONTENTS

	<u>PAGE</u>
TITLE PAGE	i
DECLARATION	ii
CERTIFICATION	iii
DEDICATION	iv
ACKNOWLEDGEMENT	v
ABSTRACT	vi
TABLE OF CONTENT	vii
LIST OF TABLES	xiii
LIST OF FIGURES	xv
LIST OF APPENDICES	xvii
 CHAPTER 1 - INTRODUCTION	 1
1.1 Definition	1
1.2 Flood Frequency Analysis	1
1.3 Literature Review	4
1.4 Objective	6
1.5 Scope of Work	7
 CHAPTER 2 - PARAMETER UNCERTAINTY FORMULATIONS	 9
2.1 Definition of Distributions and Standard Functions	9
2.1.1 Gamma - 1 Distribution	9
2.1.2 Gamma - 2 Distribution	9
2.1.3 Inverted Gamma - 1 Distribution	10
2.1.4 Inverted Gamma - 2 Distribution	10
2.1.5 General Student Function	10
2.1.6 Poisson Distribution	11
2.1.7 General Normal Function	11
2.1.8 Standardized Beta Function	12
2.1.9 Gamma Function	12

	<u>PAGE</u>
2.2 Parameter Uncertainty - General Principle	13
2.3 Parameter Uncertainty For Normal Model	16
2.3.1 Density Function and Likelihood Function	16
2.3.2 Distribution of (μ, h)	19
2.3.3 Estimation of Prior Parameters	25
2.3.4 Estimation of Expectation and Variance of μ and σ^2 ..	28
2.3.5 Posterior Parameters	30
2.3.6 Classical Proceedure for normal model	31
2.4 Parameter uncertainty for log-normal model	33
2.4.1 Density Function, Likelihood function and distribution of (μ, h)	33
2.4.2 Classical procedure for log-normal model	34
2.5 Parameter Uncertainty for Exceedance model	35
2.5.1 Density function and likelihood function	35
2.5.2 Distribution (α, v)	41
2.5.3 Estimation of Prior Parameters of Exceedance Model ...	42
(a) Prior Parameter for α	43
(b) Prior Parameter for v	44
2.5.4 Posterior parameters of exceedance model	45
2.5.5 Classical procedure for exceedance model	45
CHAPTER 3 - REGIONAL FREQUENCY ANALYSIS	47
3.1.1 Taraba River - Description of the basin	47
3.1.2 - Gassol Station	52
3.2 - Data	54
3.2.1 - Data Verification	54
3.2.2 - Data Analysis	59
3.3 - Fitting of frequency curves	59
3.4 Homogeneity Test	74
3.5 Regression Analysis	78

	<u>PAGE</u>
3.5.1 - Review of previous work on Regression Analysis.....	78
3.5.2 - General Procedure.....	78
3.6 Applications of Regression Analysis	80
3.6.1 - General Procedure	80
3.6.2 Normal Model	80
(a) Regression Model for Mean Annual Maximum Flood Discharges	80
(b) Regression Model for Variances of the Annual Maximum Flood Discharges	83
3.6.3 Log-Normal Model	85
(a) Regression Model for the Mean of The Natural Logarithm of the Annual Maximum Flood Discharges (LnQ)	85
(b) Regression Model for the Variance of the Natural Logarithm of the Annual Maximum Flood Discharges Ln Var(Q).....	87
3.6.4 Exceedance Model	88
(a) Regression Model for Mean Exceedance Flood Discharges	88
(b) Regression Model for Variance of Exceedance Flood Discharges	90
3.7 Summary of the Regression Models	92
 CHAPTER 4 - APPLICATION OF FORMULATED MODELS TO RIVER TARABA AT GASSOL GAGING STATION IN UPPER BENUE RIVER BASIN, NIGERIA	 93
4.1 Sample Parameter Estimation	93
4.1.1 General Information	93
4.1.2 Normal Distribution	93
4.1.3 Log-Normal Distribution	94
4.1.4 Exceedance Distribution	94

4.2	Prior Parameter Estimation	95
4.2.1	General Principle	95
4.2.2	Normal Model	97
	(a) Calculation of Expectation of Mean Discharge	97
	(b) Calculation of Variance of the Mean Discharge ...	98
	(c) Calculation of the Expectation of Variance of the Annual Discharge	98
	(d) Calculation of the Variance of the Annual Flood Discharges	99
4.2.3	Log-Normal Model - Prior Parameters Estimation	100
4.2.4	Exceedance Model - Prior Parameters Estimation	102
	(a) For Event Magnitude Distribution, ...	102
	(b) Prior Parameters Estimation for the Average Arrival Rate, v	102
4.3	Posterior Parameter Estimation	104
4.3.1	Estimation of Posterior Parameters for Normal Model..	104
4.3.2	Estimation of Posterior Parameters for Log-Normal Model	105
4.3.3	Estimation of Posterior Parameters for Exceedance Model	105
4.4	Bayesian Methods of Analysis	107
4.4.1	Bayesian Transformed - Normal Model	107
4.4.2	Bayesian Transformed - Log-Normal Model	107
4.4.3	Bayesian Transformed - Exceedance Model	108
4.5	Classical Method of Analysis	108
4.5.1	Normal Model	108
4.5.2	Log-Normal Model	109
4.5.3	Exceedance Model	109
4.6	Graphical Comparison of Classical, Regional and Bayesian Methods.....	109

	<u>PAGE</u>
CHAPTER 5 - MODEL UNCERTAINTY FORMULATIONS AND APPLICATIONS ..	113
5.1 Model Uncertainty - General Principle	113
5.2 Analytical Derivation of the Marginal Density Function ..	117
5.2.1 Comment	117
5.2.2 Marginal Density Function for Normal Model	117
5.2.3 Marginal Density Function for Log-Normal Model	119
5.2.4 Marginal Density Function for Exceedance Model	120
5.3 Application	124
5.3.1 - Calculation of Marginal Likelihood Function, K_N , for the Normal Model	125
5.3.2 Calculation of Marginal Likelihood Function, K_{LN} , for the Log-Normal Model	125
5.3.3 Calculation of Marginal Likelihood Function, K_E , for the Exceedance Model	125
5.3.4 Calculation of Normalizing Constant, K^*	126
5.3.5 Calculation of Posterior Model Probabilities, $P^n(\theta_i)$	127
(a) For Normal	127
(b) For Log-Normal	127
(c) For Exceedance	127
5.3.6 Calculation of the Composite Bayesian Model	128
CHAPTER 6 - DISCUSSIONS	129
6.1 Comparison of flood discharges predicted by Classical, Bayesian and Regional Methods.....	129
6.2 Selection of model that fits the Flood Data at Gassol.....	132
6.3 Sample, Prior and Posterior Mean Flood Discharges for Normal Model.....	134
6.4 Regional frequency curves for Normal and Log-Normal Models.....	134

	<u>PAGE</u>
CHAPTER 7 - CONCLUSIONS/RECOMMENDATIONS	135
7.1 Conclusions	135
7.2 Recommendations	137
References	138
Appendices	141

	<u>LIST OF TABLES</u>	<u>PAGE</u>
Table 2.1	Relationship Between Return Period By the Partial - Duration Series and the Annual Series	16
Table 3.0	Double Mass Curve Analysis	55
Table 3.1	Beard's Coefficient.....	73
Table 3.2	Confidence Limits of Regional Flood Homogeneity Test	74
Table 3.3	Data For Homogeneity Test	76
Table 3.4	The Mean Annual Maximum Flood Discharge at each Gaging Station in Upper Benue River Basin.....	81
Table 3.5	The Logarithm to Base Ten of Mean Discharge, Q_m , Drainage Area, A and Length of Main Channel, L	82
Table 3.6	The Variance V and $\log_{10} V$ of the Annual Maximum Flood Discharges of the Selected Gaging Station in Upper Benue River Basin.....	84
Table 3.7	The Mean of the Natural Logarithm of the Annual Maximum Discharges at the Various Stations.....	85
Table 3.8	The Natural Logarithm of Discharge Q_i , Drainage Area, and Main Channel Length, L_i	86
Table 3.9	Variance, V of $\ln Q$ and $\ln V$ for the Selected Gauging Stations in Upper Benue River Basin.....	87
Table 3.10	The Mean Exceedance Discharge, Drainage Area, and Main Stream Length for Six Gaging Stations.....	89
Table 3.11	The Logarithm of Mean Exceedance Discharge, Drainage Area and Length of Main Stream for Six Gaging Stations	89
Table 3.12	Variance V and Log V of the Exceedance Flood Discharges for the Six Selected Gaging Stations in Upper Benue River Basin	91
Table 3.13	Summary of the Regression Models	92

Table 4.1	Calculation of Covariance and Correlation Coefficient	95
Table 4.2	Expectations and Variances of μ , and σ^2 for Normal Model	99
Table 4.3	Expectations and Variances of μ and σ^2 for Log-Normal Model	101
Table 4.4	Calculation of Average Arrival Rate, V for Some Gaging Stations Using Maximum Likelihood Method	103
Table 4.5	Sample Parameters	106
Table 4.6	Prior Parameters	106
Table 4.7	Posterior Parameters	107
Table 6.1	Comparison of Classical, Bayesian and Regional Discharges with observed discharges.....	130

<u>LIST OF FIGURES</u>	<u>PAGE</u>
Figure 3.1a - Map of Upper Benue River Basin showing the gaging stations, relief and drainages	48
Figure 3.1b - Location Map of Hydrological and Meteorological Stations	49
Figure 3.2 Double Mass Curve Analysis.....	58
3.3 - Frequency distribution for discharge of streamflow at Mayo Kam.....	60
Figure 3.4 - Frequency distribution for discharge of streamflow at Beli	61
Figure 3.5 - Frequency distribution for discharge of streamflow at Tela	62
Figure 3.6 - Frequency distribution for discharge of streamflow at Suntai	63
Figure 3.7 - Frequency distribution for discharge of streamflow at Suntai Bakundi	64
Figure 3.8 - Frequency distribution for discharge of streamflow at Gindin Dorowa	65
Figure 3.9 - Frequency distribution for discharge of streamflow at Donga	66
Figure 3.10 - Frequency distribution for discharge of streamflow at Manya	67
Figure 3.11 - Frequency distribution for discharge of streamflow at Abong	68
Figure 3.12 - Frequency distribution for discharge of streamflow at Sevav	69

	<u>PAGE</u>
Figure 3.13 - Frequency distribution for discharge of streamflow at Katsina Ala	70
Figure 3.14 - Frequency distribution for discharge of streamflow at Gassol	71
Figure 3.15 - Homogeneity Test Chart.....	77
Figure 4.0 - Correlation between drainage area and length of main channel for the stations.....	96
Figure 4.1 - Comparison of Classical and Bayesian Frequency Curves for Normal Model.....	110
Figure 4.2 - Comparison of Classical and Bayesian Frequency Curves for Log-Normal Model.....	111
Figure 4.3 - Comparison of Classical, Regional and Bayesian Frequency Curves for Exceedance Model.....	112

	<u>LIST OF APPENDICES</u>	<u>PAGE</u>
Appendix A - 1	Exceedance Probability Analysis for Station 1	141
Appendix A - 2	Exceedance Probability Analysis for Station 2	142
Appendix A - 3	Exceedance Probability Analysis for Station 3	143
Appendix A - 4	Exceedance Probability Analysis for Station 4	144
Appendix A - 5	Exceedance Probability Analysis for Station 5	145
Appendix A - 6	Exceedance Probability Analysis for Station 6	146
Appendix A - 7	Exceedance Probability Analysis for Station 7	147
Appendix A - 8	Exceedance Probability Analysis for Station 8	148
Appendix A - 9	Exceedance Probability Analysis for Station 9	149
Appendix A - 10	Exceedance Probability Analysis for Station 10	150
Appendix A - 11	Exceedance Probability Analysis for Station 11	151
Appendix A - 12	Exceedance Probability Analysis for Station 12	152

CHAPTER 1INTRODUCTION1.1 DEFINITION

A flood is an unusually high stage in a river that can cause damage to adjacent area. Large monetary damages, loss of lives and properties are often encountered whenever a flood occurs.

In water resources planning the hydrologist's main function is analysis that lead to engineering decisions. Such analysis is usually connected with the frequency of occurrence, the maximum stage reached, the volume of flood water, the area inundated and the duration of floods. The decision variable is not a hydrologic variable, but a general engineering variable like the height of a dike or the size of a spillway. Thus the separation of hydrologic analysis and economic analysis cannot occur if efficient designs are to be obtained.

1.2

FLOOD FREQUENCY ANALYSIS

The flood frequency analysis (the estimation of the relationship between the magnitude Q and the return period T) which is a central problem in applied hydrology is usually carried out using the following methods:

- (A) analysis using a historical record of discharge at the site in question.
- (B) analysis using regional information or data from other stations in the same homogeneous basin.
- (C) relationship between rainfall and discharge.
- (D) Two station comparison.

The first of the above methods is the most popular and is referred to here as the classical method of analysis. The procedure normally used in this method is summarized (Wood and Rodriguez-Iturbe, 1975a) as follows:

- (i) Observe a historical record of flood events, the length of such record averages 40-50 years and in many cases is much less.
- (ii) Pick a generating process or probability density function that seems reasonable. For example, Gumbel distribution, normal distribution, log-normal distribution, Pearson type III distribution, log-Pearson type III distribution, Poisson process, Binomial and multinomial distributions.
- (iii) Estimate the parameters of this underlying process from historical record. For example, mean discharge, variance and skewness using maximum likelihood estimate (MLE) method or method of moments.
- (iv) Make inferences about the occurrences of future flood events.

However, there exist a number of sources of errors in numbers (ii) and (iii) which are also grouped into the following three categories (Benjamin and Cornell, 1970):

- (a) Natural uncertainty which is determined as the uncertainty inherent in the natural generating process, due to the randomness of the stream flow process.
- (b) Parameter (statistical) uncertainty which is associated with the estimation of the parameters of the model of the stochastic process due to limited data.
- (c) Model uncertainty which is associated with the uncertainty that a particular probabilistic model of the stochastic process may not be the 'true' or best method.

Often the classical analyst takes care of these uncertainties by means of some artificial mechanisms such as adding a safety factor to the design variable, designing for the largest possible event or using some other methods which cannot be related to meaningful economic (or social) preference criterion.

The Bayesian procedure which allows some sources of information (for example, historical record and regional record) to be used is therefore proposed so as to reduce both parameter and model uncertainties.

In Regional frequency analysis, a statistically homogeneous basin is determined. Within such basins, the results of point-data analysis can be averaged to represent the frequency characteristic of the entire basin. Thus, the average probability curve so obtained is applicable to the whole basin (Chow, Ven-Te, 1964). Usually multiple regression analysis is employed. Thus, from the multiple regression equation so obtained, the discharges at other stations where there are scanty or short record are estimated. This method is attractive in its empiricism, but estimates obtained are usually subject to large sampling variance.

The rainfall and discharge relationship, despite its popularity in the early days of scientific hydrology, is still inadequately developed even in principle (Cunnane and Nash, 1973). The synthesis of a rainfall-amount-duration frequency relationship and a method of converting rainfall into discharge is by no means trivial, except when the model for the latter operation is the rational method.

In the two station comparison (Viessman et al, 1977), regression technique is frequently employed to extend short records if significant correlation exists between the gaging station of short record and a nearby gaging station of longer record.

The reliability of such method depends on the correlation coefficient and the length of the concurrent records. If the correlation coefficient is low and the concurrent record is too short, the method gives unreliable results.

1.3

LITERATURE REVIEW

Shane and Graver,(1970),were the first to propose the use of both regional information from regression models and historical at-site information in flood frequency analysis and the resulting engineering decisions. The objective in that work was to combine estimators from these two sources.

Wood,(1973),compared the Bayesian procedure with the classical method (Maximum Likelihood estimate) of estimating parameters of an exceedance model in a case where the historical data was short. The comparison was based on the historical data obtained for a river in the North-eastern United State.

The prior parameters of the exceedance model were estimated from regional frequency analysis carried out in 36 other U.S. River Basins. The result of his studies indicated that the Bayesian procedure is superior to the classical method in reducing parameter uncertainty.

Wood and Rodriguez-Iturbe,(1975a),applied the Bayesian proceedure to the Normal, log-Normal and exceedance distributions. The result of their analysis indicated that both the normal and log-normal distributions are transformed into Student t - distributions. In case of exceedance distribution an approximate exceedance model was developed using Bayesian procedure. These models take into consideration parameter uncertainty.

The Bayesian procedure was also applied to the same three distributions (normal, log-normal and exceedance) as if they were competing probabilistic models. Wood and Rodriguez-Iturbe, (1975b) arrived at a composite model which was obtained by combining the above three models. The composite model reduces the model uncertainty.

These analytical derivations were applied to the Blackstone River at Woonsocket, Rhode Island in U.S.A. The result of the studies indicated that it is safer to base engineering decisions on Bayesian procedure than the classical method.

Vicens et al, (1975) also applied the Bayesian procedure to the normal distribution and showed that the Normal distribution is transformed into a Student - t model. They also showed that the marginal distributions of the mean discharge, μ , and the variance, σ^2 are Student - t and inverted gamma - 1 distributions respectively. It was also shown in the paper that the prior parameters could be estimated from the first and second moments of the marginal distributions of μ and σ^2 . Analytical expressions for the parameters obtained by using the moments of these marginal distributions of μ and σ^2 , resulted into equations involving the expectations and variances of μ and σ^2 given the regional information.

The analytical results were applied to Pemigewasset River at Plymouth, New Hampshire, U.K. The result of the studies of Vicens et al, (1975) indicated that the Bayesian procedure, which allows the combination of both historical at-site record and the regional information significantly reduces the parameter uncertainty of the model when the historical record is short.

Odusote (1989) performed regional frequency analysis on Ogun-Oshun - River Basin in Nigeria, using regression analysis and he arrived at an equation of the form:

$$Q_{mi} = a A_i^b L_i^c \dots\dots\dots(1.1)$$

relating Mean Maximum Annual flood discharge, Q_m , to the physiographic information - drainage area, A , and the main stream length, L .

In equation (1.1) i is the station number while a , b , c are regression constants to be determined.

Benjamin and Cornell, (1970) stated that when g is a function of variables $x_1, x_2, \dots, x_n$, and if $x_1, x_2, \dots, x_n$ have a high linear dependence, then the expectation and variance of g should be found using Taylor's first order and second order approximations series respectively.

1.4

OBJECTIVE

In some of the gaging stations of the river basins in West Africa, there are virtually no historical records and if they exist, the records are often scanty and of short duration. Classical method of flood frequency analysis cannot be used at these stations due to limited data. The use of Bayesian analysis or procedure which can combine both regional information or other sources of information and the limited historical data will give a reliable result. That is, both parameter and model uncertainties are greatly reduced.

It is hoped that this work will complement the work earlier initiated on the area (Upper Benue River Basin in Nigeria) by Ogbuli (in print) on the use of regional frequency analysis.

The objective of the analysis (Classical or Bayesian) is to predict flood discharges with return period greater or equal to ten years.

The areas covered by this thesis are arranged as follow:-

In section (1.2) of this chapter, various traditional methods of flood frequency analysis were discussed together with their limitations. The author finally proposed the use of Bayesian method of analysis since it can combine the historical data with other sources of information.

The major part of this thesis have been devoted to the derivation of analytical expressions for reducing parameter and model uncertainties as well as their applications to River Taraba at Cassol in the Upper Benue River Basin in Nigeria.

The author has chosen regional information as his other source of information apart from the historical data and therefore an auxiliary work has been carried out on regional frequency analysis.

In chapter two models have been derived for reducing parameter uncertainty for three selected distributions, namely: normal, log-normal, and exceedance distributions. For each model, the estimation of the posterior parameter involves as a first step, the estimation of sample parameters and secondly, the estimation of prior parameters using regional information. Also the classical method of estimating the parameters of the models (using maximum likelihood estimate method) have been presented.

In chapter three, description of the Taraba River Basin and Cassol gaging station, frequency analysis of the twelve gaging stations, homogeneity test and theoretical multiple regression analysis have been presented. The application of the multiple regression analysis for the estimation of the mean and variance of the flood discharges for the three chosen distributions using

the regional information from some selected gaging station records in Upper Benue River Basin in Nigeria are also presented.

Chapter four deals with the application of the analytical models developed in chapter two to River Taraba at Gassol gaging station in the Upper Benue River Basin in Nigeria. Throughout the calculations, the Gassol gaging station record has been used as the sample record. Also for each of the three models a graphical comparison of the Bayesian method of flood frequency analysis with the classical method of flood frequency analysis have been presented.

In chapter five, sections 5.1 and 5.2, the analytical derivation of expressions for reducing the model uncertainty have been presented. These result into a composite model which reduces both parameter and model uncertainties. Section 5.3 of the chapter deals with the applications of the analytical expressions derived in sections 5.1 and 5.2 to the River Taraba at Gassol in Upper Benue River Basin. The composite model which consists of the three selected models was found to result into only the exceedance model.

The discussions and conclusions/recommendations drawn from the studies are presented in chapter 6 and 7 respectively.

CHAPTER 2

PARAMETER UNCERTAINTY FORMULATIONS2.1 DEFINITIONS OF DISTRIBUTIONS AND STANDARD FUNCTIONS

The following distributions have been widely used in this work and their definitions and moments are as given (Raiffa and Schlaifer, 1961) below.

2.1.1. Gamma - 1 Distribution

The gamma-1 normalized density function is defined by:

$$f_{\gamma 1}(z|r, y) = \frac{\exp(-yz)(yz)^{r-1}}{(r-1)!} y, \quad \left. \begin{array}{l} z \geq 0 \\ r, y \geq 0 \end{array} \right\} \dots\dots (2.1.01)$$

where r, y are the parameters and z , the dependent variable.

Its first two moments are:

$$\mu_1 = \frac{r}{y} \dots\dots\dots (2.1.02)$$

$$\mu_2 = \frac{r}{y^2} \dots\dots\dots (2.1.03)$$

2.1.2. Gamma - 2 Distribution

The gamma-2 normalized density function is defined by:

$$f_{\gamma 2}(z|y, v) = \frac{\exp(-\frac{1}{2}vyz)(\frac{1}{2}vyz)^{\frac{1}{2}v-1}}{(\frac{1}{2}v-1)!} (\frac{1}{2}vy), \quad \left. \begin{array}{l} z \geq 0 \\ v, y \geq 0 \end{array} \right\} \dots\dots (2.1.04)$$

where y and v are the parameters, and z , the dependent variable.

Its first two moments are:

$$\mu_1 = \frac{1}{y} \dots\dots\dots (2.1.05)$$

$$\mu_2 = \frac{1}{\frac{1}{2}vy^2} \dots\dots\dots (2.1.06)$$

2.1.3. Inverted Gamma - 1 Function

The inverted -gamma- 1 normalized function is defined

as:

$$f_{i\gamma 1}(t|r, y) = \frac{\text{Exp}(-y/t) (y/t)^{r+1}}{(r-1)!} \frac{1}{y}, \quad \left. \begin{array}{l} t \geq 0 \\ r, y > 0 \end{array} \right\} \dots\dots\dots (2.1.07)$$

where r, y are the parameters and t is the dependent variable.

Its first two moments are:

$$\mu_1 = \frac{y}{r-1} \dots\dots\dots (2.1.08)$$

$$\mu_2 = \frac{y^2}{(r-1)^2 (r-2)} \dots\dots\dots (2.1.09)$$

2.1.4. Inverted Gamma - 2 Functions

The inverted gamma - 2 normalized density function is

defined by:

$$f_{i\gamma 2}(v^2|s, v) = \frac{2 \exp(-\frac{1}{2}vs^2/v^2) (\frac{1}{2}vs^2/v^2)^{\frac{1}{2}v+\frac{1}{2}}}{(\frac{1}{2}v-1)! (\frac{1}{2}vs^2)^{\frac{1}{2}}}, \quad \left. \begin{array}{l} v \geq 0 \\ s, v > 0 \end{array} \right\} \dots\dots\dots (2.1.10)$$

where s, v are the parameters.

Its first two moments are:

$$\mu_1 = \frac{s \sqrt{\frac{1}{2}v} (\frac{1}{2}v-3/2)!}{(\frac{1}{2}v-1)!} \quad v > 1 \quad \dots\dots\dots (2.1.11)$$

$$\mu_2 = s^2 \frac{v}{v-2} - \mu_1^2 \quad v > 2 \quad \dots\dots\dots (2.1.12)$$

2.1.5. General Student Function

The general or non-standardized student density function

is defined by:

$$f_S(z|m, H, v) = \int_0^\infty f_N(z|m, hH) f_{\gamma 2}(h|1, v) dh \dots\dots (2.1.13a)$$

$$= \frac{v^{\frac{1}{2}v}}{B(\frac{1}{2}, \frac{1}{2}v)} \left[v + H(z - m)^2 \right]^{-\frac{1}{2}(v+1)} \sqrt{H} \dots\dots\dots (2.1.13b)$$

where m, h are the parameters (that is, mean and precisions)

v is the No. of degrees of freedom.

$$-\infty < z < \infty$$

$$H, v > 0$$

The equation (2.1.13b) can be transformed to give

$$f_S(z|m, H, v) = \frac{1}{B(\frac{1}{2}, \frac{1}{2}v)} \left[1 + \frac{H}{v} (z - m)^2 \right]^{-\frac{1}{2}(v+1)} \left[\frac{v}{H} \right]^{-\frac{1}{2}} \dots\dots\dots (2.1.13c)$$

The first two moments of the Student density function

are:

$$\mu_1 = m \quad v > 1 \quad \dots\dots\dots (2.1.14)$$

$$\mu_2 = \frac{1}{H} \cdot \frac{v}{v-2} \quad v > 2 \quad \dots\dots\dots (2.1.15)$$

2.1.6. Poisson Distribution

The Poisson normalized mass function is defined by:

$$f_p(r|m) = \frac{\exp(-m) m^r}{r!}, \quad r = 0, 1, 2, \dots\dots\dots (2.1.16)$$

$m > 0$

where m is a parameter and r is a variable.

Its first two moments are:

$$\mu_1 = m \quad \dots\dots\dots (2.1.17)$$

$$\mu_2 = m \quad \dots\dots\dots (2.1.18)$$

2.1.7. General Normal Function

The general or non-standardized normal normalized

density function is defined by:

$$F_N(z|m, H) = \frac{1}{\sqrt{2\pi}} \exp \left[-\frac{1}{2}H(z-m)^2 \right] \sqrt{H} \dots\dots\dots (2.1.19)$$

where m, H are parameters.

The first two complete moments of Normal Function are:

$$\mu_1 = m \quad \dots\dots\dots (2.1.20)$$

$$\mu_2 = \frac{1}{H} \quad \dots\dots\dots (2.1.21)$$

2.1.8. Standardized Beta Function

The standardized beta normalized density function is defined by:

$$f_{B^*}(z|p, q) = \frac{1}{B(p, q)} z^{p-1} (1-z)^{q-1}, \quad \left. \begin{array}{l} 0 < z < 1 \\ p, q > 0 \end{array} \right\} \dots\dots\dots (2.1.22)$$

where p, q are parameters and $B(p, q)$ is the complete beta function defined as:

$$B(p, q) = \frac{(p-1)! (q-1)!}{(p+q-1)!} \quad \dots\dots\dots (2.1.23)$$

The first two moments of the Beta function are:

$$\mu_1 = \frac{p}{p+q} \quad \dots\dots\dots (2.1.24)$$

$$\mu_2 = \frac{pq}{(p+q)^2 (p+q-1)} \quad \dots\dots\dots (2.1.25)$$

Gamma Function

2.1.9. The gamma function is defined as:

$$\Gamma(k) = \int_0^{\infty} e^{-u} u^{k-1} du \quad \dots\dots\dots (2.1.26)$$

For the values of k in the above integral, the gamma function is defined as:

$$\Gamma(k) = (k-1)! \quad \dots\dots\dots (2.1.27)$$

In the special case when $k = \frac{1}{2}$, then

$$\Gamma\left(\frac{1}{2}\right) = \left(-\frac{1}{2}\right)! = \pi^{\frac{1}{2}} \quad \dots\dots\dots (2.1.28)$$

2.2

PARAMETER UNCERTAINTY - GENERAL PRINCIPLE

Bayesian inference lays its foundation upon the fact that states of nature can be treated as random variables (Wood et al, 1974). Considering annual flood discharges, the mean μ , is a random variable distributed with a sample mean m and variance s^2 estimated from the available information. Usually the available information is the historical record. However, other sources of information may be used. An example of this is the regional information.

Supposing one has a set of historical data, Q of annual flood discharges which is assumed to come from the probability model, $f_Q(q|\theta)$, which is conditional upon the uncertain set of parameters θ .

Thus Q and θ are defined respectively as:

$$Q = \{ q_1, q_2, q_3 \quad \dots \quad q_n \} \quad \dots\dots\dots (2.2.1)$$

$$\theta = \{ \theta_1, \theta_2, \theta_3 \quad \dots \quad \theta_n \} \quad \dots\dots\dots (2.2.2)$$

Baye's theorem provides a means by which the prior distribution of the parameter set θ can be updated by the data set Q to provide posterior distribution of θ (Wood and Rodriguez-Iturbe, 1975a) namely:

$$f''(\theta) = f''(\theta|Q, I_0) = f(Q|\theta) \cdot f'(\theta|I_0) \frac{1}{k} \quad \dots\dots\dots (2.2.3)$$

where:

$$f''(\theta) = f''(\theta|Q, I_0) \quad - \quad \text{Posterior density function}$$

for θ conditional upon a set of data Q and information I_0 .

$$f(Q|\theta) = L(\theta|Q) \quad - \quad \text{sample likelihood function of}$$

the parameter set, conditional upon the observations (or the probability of the observations given the parameters)

$f'(\theta/I_0)$ - Prior density function for θ , conditional upon the initial information.

K - normalizing constant.

If future observations Q_F are available, Baye's theorem can be used to update the probability density function on θ . In this case the former posterior density function for θ now becomes the prior density function, since it is prior to the new observations or the utilization of new data. The new posterior density function would have been obtained if the two samples Q and Q_F had been observed sequentially as one set of data.

One is concerned about making a specific engineering decision involving the frequency of the flood discharge q . Thus, the inferences about q should reflect the uncertainty in θ . This is done by applying compound distribution theory in a Bayesian procedure. This procedure results in the Bayesian transformed form of the distribution (Benjamin and Cornell, 1970) for flood discharges q . This transformed distribution is found by:

$$\tilde{f}(q) = \int_{\theta} f_Q(q|\theta) \cdot f''(\theta) d\theta \dots\dots\dots(2.2.4)$$

where:

$\tilde{f}(q)$ - Bayesian transformed form of the distribution of the flood discharges now parameter free.

This resulting probability density function includes the natural or inherent randomness of the stream-flow process and the uncertainty about the parameters (Vicens et al, 1975). It can be viewed as a weighted average of the model probability density function with the posterior probability density function of the parameters as the weights.

In making inferences on flood discharges, it is important to use the Bayesian transformed form of the distribution for q , $f(q)$, as opposed to the probabilistic distribution for q , $f(q|Q^*)$ with the Bayes estimator (or any other point estimator) for the parameter set θ because using point estimators for uncertain parameters underestimates the variance in the flood discharges q due to those parameters (Wood and Rodriguez-Iturbe, 1975a).

The Bayesian transformed form of the distribution for flood discharges will be derived for three selected probability distributions Namely: (1) Normal distribution (2) Log-normal distribution and (3) Exceedance distribution.

To obtain this aim three relevant models were constructed.

The normal and log-normal models consider the annual series of flood discharges, which is made up of the largest flood discharges in each year. The exceedance model considers the partial duration series of flood peaks above a 'base' level. However, since partial duration series and annual flood series are virtually identical above a frequency of about the 10-year flood (See Table 2.1)

Langbein, (1949) and Wood, (1973), the use of the annual series for the prior was considered to be adequate for the example studied. Thus the base-level discharge corresponding to a return period of about 10-years is chosen for the exceedance model.

Table 2.1: Relationship Between Return Period By the
Partial - Duration Series and the Annual Series

Recurrence Interval in Years by Partial-Duration Series	Recurrence Interval in Years by Annual Series
0.50	1.16
1.00	1.58
1.45	2.00
2.00	2.54
5.00	5.52
10.00	10.50
20.00	20.50
50.00	50.50
100.00	100.50

After (Langbein, 1949)

2.3 PARAMETER UNCERTAINTY FOR NORMAL MODEL

2.3.1 Density Function and Likelihood Function

Supposing a normal distribution generates random variables say $q_1, q_2 \dots q_i$, then referring to equation (2.1.19) the probability density function is given by:

$$f_N(q|\mu, h) = (2\pi)^{-\frac{1}{2}} \exp \left[-\frac{1}{2}h(q - \mu)^2 \right] \cdot h^{\frac{1}{2}} \quad -\infty < q < \infty \dots (2.3.1) \\ -\infty < \mu < \infty \\ h > 0$$

where μ is the mean annual flood discharge and h is the precision.

Referring to equations (2.1.20) and (2.1.21) the mean flood discharge, μ and the precision h are given respectively by:

$$E(q|\mu, h) = \mu \dots (2.3.2)$$

$$V(q|\mu, h) = \frac{1}{h} \dots (2.3.3)$$

The likelihood of observations $q_1, q_2, \dots, q_{n-1}, q_n$ on a normal distribution is

$$L(\mu, h|Q) = \prod_{i=1}^n f(q_i|\mu, h) \dots (2.3.4)$$

$$\text{where } \prod_{i=1}^n f(q_i|\mu, h) = f(q_1|\mu, h) \cdot f(q_2|\mu, h) \dots f(q_n|\mu, h)$$

$$= (2\pi)^{-\frac{1}{2}} \exp \left[-\frac{1}{2}h(q_1 - \mu)^2 \right] h^{\frac{1}{2}} \cdot (2\pi)^{-\frac{1}{2}} \exp \left[-\frac{1}{2}h(q_2 - \mu)^2 \right] h^{\frac{1}{2}} \dots$$

$$(2\pi)^{-\frac{1}{2}} \exp \left[-\frac{1}{2}h(q_{n-1} - \mu)^2 \right] h^{\frac{1}{2}} \cdot (2\pi)^{-\frac{1}{2}} \exp \left[-\frac{1}{2}h(q_n - \mu)^2 \right] h^{\frac{1}{2}}$$

$$L(\mu, h|Q) = (2\pi)^{-n/2} \exp \left[-\frac{1}{2}h \sum_{i=1}^n (q_i - \mu)^2 \right] \cdot h^{n/2} \dots (2.3.5)$$

where n - number of observations in the sample Q .

Let us introduce the following statistics:

m - the mean of the sample

w - the variance of the sample

v - the number of degrees of freedom, defined as:

$$m = \frac{1}{n} \sum_{i=1}^n q_i \dots (2.3.6)$$

$$w = \frac{1}{n-1} \sum_{i=1}^n (q_i - m)^2 \dots (2.3.7)$$

$$v = n - 1 \dots (2.3.8)$$

The expression in exponent in equation (2.3.5) can be expanded as:

$$\begin{aligned} \exp \left[-\frac{1}{2}h \sum (q_i - \mu)^2 \right] &= \exp \left\{ -\frac{1}{2}h \left[\left(\sum q_i \right)^2 - 2 \left(\sum q_i \right) \mu + n \mu^2 \right] \right\} \\ &= \exp \left\{ -\frac{1}{2}h \left[n^2 \cdot \frac{1}{n^2} \left(\sum q_i \right)^2 - 2n \cdot \frac{1}{n} \sum q_i \mu + n \mu^2 \right] \right\} \dots (2.3.9) \end{aligned}$$

Substituting equation (2.3.6) into the right side of equation (2.3.9) gives:

$$\exp \left\{ -\frac{1}{2}h \sum (q_i - \mu)^2 \right\} = \exp \left\{ -\frac{1}{2}h \left[(nm)^2 - 2(nm) \mu + n \mu^2 \right] \right\} \dots (2.3.10)$$

On addition of $2nm^2$ and $-2nm^2$ to the right side of equation (2.3.10) and rearranging

$$\begin{aligned} \exp \left\{ -\frac{1}{2}h \sum (q_i - \mu)^2 \right\} &= \exp \left\{ -\frac{1}{2}h \left[(nm)^2 - 2nm^2 + 2nm^2 - 2nm\mu + n\mu^2 \right] \right\} \\ &= \exp \left\{ -\frac{1}{2}h \left[(nm)^2 - 2m(nm) + nm^2 + nm^2 - 2m\mu n + n\mu^2 \right] \right\} \dots (2.3.11) \end{aligned}$$

From equation (2.3.6) $mn = \sum q_i$, thus substituting this in the above expression gives:

$$\begin{aligned} \exp \left\{ -\frac{1}{2}h \sum (q_i - \mu)^2 \right\} &= \exp \left\{ -\frac{1}{2}h \left[\left(\sum q_i \right)^2 - 2m \sum q_i + nm^2 \right] + (nm^2 - 2m\mu n + n\mu^2) \right\} \\ \exp \left\{ -\frac{1}{2}h \sum (q_i - \mu)^2 \right\} &= \exp \left\{ -\frac{1}{2}h \left[\left(\sum (q_i - m) \right)^2 + n(m - \mu)^2 \right] \right\} \dots (2.3.12) \end{aligned}$$

The right side of equation (2.3.12) can be further modified by multiplying and dividing the first term by $n - 1$ to give:

$$\exp \left\{ -\frac{1}{2}h \sum (q_i - \mu)^2 \right\} = \exp \left\{ -\frac{1}{2}h(n-1) \sum \frac{(q_i - m)^2}{n-1} - \frac{1}{2}hn(m - \mu)^2 \right\} \dots (2.3.13)$$

Substituting equation (2.3.13) into equation (2.3.5) gives:

$$L(u, h|Q) = (2\pi)^{-n/2} \exp \left[-\frac{1}{2}h(n-1) \sum_{i=1}^n \frac{(q_i - m)^2}{(n-1)} \right] \cdot \exp \left[-\frac{1}{2}hn(m - \mu)^2 \right] h^{n/2} \dots (2.3.14)$$

Substituting equations (2.3.7) and (2.3.8) into equation (2.3.14) gives:

$$L(u, h|Q) = (2\pi)^{-n/2} \exp(-\frac{1}{2}hvw) \cdot \exp \left[-\frac{1}{2}hn(m - \mu)^2 \right] \cdot h^{n/2} \dots (2.3.15)$$

The Kernel of this function consisting of terms which vary only with the unknown parameters (μ and h) is:

$$\exp(-\frac{1}{2}hvw) \exp \left[-\frac{1}{2}hn(m - \mu)^2 \right] h^{n/2} = h^{\frac{1}{2}} \exp(-\frac{1}{2}hvw) \cdot \exp \left[-\frac{1}{2}hn(m - \mu)^2 \right] \cdot h^{(n-1)/2}$$

Substituting equation (2.3.8) in the power of last term of the Kernel function.

$$L(\mu, h|Q) = (2\pi)^{-n/2} h^{\frac{1}{2}} \exp(-\frac{1}{2}h\omega v) \cdot \exp[-\frac{1}{2}hn(m-\mu)^2] \cdot h^{v/2} \dots (2.3.16)$$

Thus from equation (2.3.16) results that likelihood function is proportional to the expression:

$$L(\mu, h|Q) \propto h^{\frac{1}{2}} \cdot \exp[-\frac{hn}{2}(m-\mu)^2] \cdot \exp[-\frac{h\omega v}{2}] \cdot h^{v/2} \dots (2.3.17)$$

From equation (2.3.17) it can be deduced that (a) the Kernel of the marginal likelihood of (m, n) is:

$$h^{\frac{1}{2}} \cdot \exp[-\frac{1}{2}hn(m-\mu)^2] \dots (2.3.18)$$

(b) While the Kernel of the marginal likelihood of (w, v) is:

$$h^{v/2} \cdot \exp[-\frac{1}{2}h\omega v] \dots (2.3.19)$$

2.3.2 Distribution of (μ, h)

When both parameters of a normal model are unknown and are to be treated as random variables μ and h , the most convenient joint distribution of the two variables, that is, the natural conjugate of equation (2.3.16) is a Normal-gamma distribution (Raiffa and Schlaifer, 1961), defined as a product of f_N and $f_{\gamma 2}$:

$$f_{N\gamma}(\mu, h|m, w, n, v) = f_N(\mu|m, hn) \cdot f_{\gamma 2}(h|w, v) \dots (2.3.20)$$

Following from equations (2.1.19) and (2.1.04) respectively.

$$f_N(\mu|m, hn) = (2\pi)^{-\frac{1}{2}} \exp[-\frac{1}{2}hn(m-\mu)^2] \cdot \sqrt{hn} \dots (2.3.21)$$

Where product $hn = H$. That is, we say that the information H is equivalent to n observations on the normal model.

$$f_{\gamma 2}(h|w, v) = \frac{\exp[-\frac{1}{2}whv] \cdot (\frac{1}{2}whv)^{\frac{1}{2}v-1} \cdot \frac{1}{2}wv}{(\frac{1}{2}v-1)!} \dots (2.3.22)$$

Thus from equation (2.3.21) by considering only terms involving h , it can be deduced that:

$$f_N(\mu|m, hn) \propto h^{\frac{1}{2}} \cdot \exp\left[-\frac{1}{2}hn(m-\mu)^2\right] \dots\dots\dots (2.3.23)$$

Similarly considering only terms involving h it can be deduced from equation (2.3.22) that:

$$f_{N2}(h|w, v) \propto h^{\frac{1}{2}v-1} \cdot \exp\left[-\frac{1}{2}hwv\right] \dots\dots\dots (2.3.24)$$

Substituting relations (2.3.23) and (2.3.24) into equation (2.3.20) gives:

$$f_{N\delta}(\mu, h|m, w, n, v) \propto h^{\frac{1}{2}} \cdot \exp\left[-\frac{h}{2}n(m-\mu)^2\right] \cdot h^{\frac{1}{2}v-1} \cdot \exp\left[-\frac{1}{2}hwv\right] \dots\dots\dots (2.3.25)$$

Thus if the prior distribution of (μ, h) is Normal-gamma with parameters (m', n', w', v') , then equation (2.3.25) implies that:

$$f'_{N\delta}(\mu, h|m', w', n', v') \propto h^{\frac{1}{2}} \cdot \exp\left[-\frac{1}{2}hn'(m'-\mu)^2\right] \cdot h^{\frac{1}{2}v'-1} \cdot \exp\left[-\frac{1}{2}hw'v'\right] \dots\dots\dots (2.3.26)$$

If the prior distribution of (μ, h) is Normal-gamma with parameters (m', n', w', v') and if a sample then yields a sufficient statistics (m, w, n, v) , the posterior distribution of (μ, h) will be Normal-gamma with parameters (m'', n'', w'', v'') .

Following the same form as in equation (2.3.25).

$$f''_{N\delta}(\mu, h|m'', w'', n'', v'') \propto h^{\frac{1}{2}} \cdot \exp\left[-\frac{1}{2}hn''(m''-\mu)^2\right] \cdot h^{\frac{1}{2}v''-1} \cdot \exp\left[-\frac{1}{2}hw''v''\right] \dots\dots\dots (2.3.27)$$

Collecting similar terms of eq (2.3.27) together gives:

$$f''_{N\delta}(\mu, h|m'', w'', n'', v'') \propto h^{\frac{1}{2}} \cdot h^{\frac{1}{2}v''-1} \cdot \exp\left\{-\frac{1}{2}h\left[n''(m''^2-2m''\mu+\mu^2)+w''v''\right]\right\} \dots\dots\dots (2.3.28)$$

But from equation (2.2.3)

$$f''_{N\delta}(\mu, h|m'', w'', n'', v'') = L(\mu, h|Q) \cdot f'_{N\delta}(\mu, h|m', w', n', v')/k \dots\dots\dots (2.3.29)$$

Substituting relations (2.3.17) and (2.3.26) into equation (2.3.29):

$$f''_{N\delta}(\mu, h|m'', w'', n'', v'') \propto \left\{ h^{\frac{1}{2}} \cdot \exp\left[-\frac{1}{2}hn(m-\mu)^2\right] \cdot \exp\left[-\frac{1}{2}hwv\right] \cdot h^{v/2} \right\} \cdot \left\{ h^{\frac{1}{2}} \cdot \exp\left[-\frac{1}{2}hn'(m'-\mu)^2\right] \cdot h^{\frac{1}{2}v'-1} \cdot \exp\left[-\frac{1}{2}hw'v'\right] \right\} \dots\dots\dots (2.3.30)$$

Collecting similar terms together, the right side of relation (2.3.30)

becomes:

$$h^{\frac{1}{2} + \frac{1}{2} + v/2 + v'/2 - 1} \exp \left\{ -\frac{1}{2} h n (m - \mu)^2 - \frac{1}{2} h v w - \frac{1}{2} h n' (m' - \mu)^2 - \frac{1}{2} h w' v' \right\}$$

or

$$h^{\frac{1}{2}(v+v')} \exp \left\{ -\frac{1}{2} h \left[n(m^2 - 2m\mu + \mu^2) + vw + n'(m'^2 - 2m'\mu + \mu^2) + w'v' \right] \right\}$$

Thus relation (2.3.30) becomes:

$$f''_{N\delta}(\mu, h | m'', w'', n'', v'') \propto h^{\frac{1}{2}(v+v')} \exp \left\{ -\frac{1}{2} h \left[n(m^2 - 2m\mu + \mu^2) + vw + n'(m'^2 - 2m'\mu + \mu^2) + w'v' \right] \right\} \dots (2.3.31)$$

Since the left side of relation (2.3.28) and (2.3.31) are the same,

it means the right sides of both equations are identically the same

and hence

$$h^{\frac{1}{2}(v''-1)} \exp \left\{ -\frac{1}{2} h \left[n''(m''^2 - 2m''\mu + \mu^2) + w''v'' \right] \right\} \\ = h^{\frac{1}{2}(v+v')} \exp \left\{ -\frac{1}{2} h \left[n(m^2 - 2m\mu + \mu^2) + vw + n'(m'^2 - 2m'\mu + \mu^2) + w'v' \right] \right\} \dots (2.3.32)$$

Comparing the coefficients of μ^2 in equation (2.3.32) implies:

$$n'' = n + n' \dots (2.3.33)$$

Comparing the powers of h in equation (2.3.32) implies:

$$\frac{1}{2}(v'' - 1) = \frac{1}{2}(v' + v) \\ v'' = v + v' + 1 \dots (2.3.34)$$

Comparing the coefficients of μ in equation (2.3.32) implies

$$n''m'' = nm + n'm' \\ m'' = \frac{nm + n'm'}{n''} \dots (2.3.35)$$

Comparing the constant powers of the exponent

$$n''m''^2 + w''v'' = nm^2 + vw + n'm'^2 + w'v' \\ w'' = 1/v'' \left[(nm^2 + vw) + (n'm'^2 + w'v') - n''m''^2 \right] \dots (2.3.36)$$

The corresponding equation obtained by (Wood and Rodriguez-Iturbe,

1975a) for the value of m'' was

$$m'' = 1/2n'' \left[n'm' + nm \right]$$

Thus an error has been detected here because the correct value for

m'' is as given by equation (2.3.35).

The Bayesian distribution of the annual maximum flood q can be determined from equation (2.2.4). If the posterior density function of the parameter μ and h is normal gamma as defined by equation (2.3.27) with posterior parameters as proved above and if the distribution is normal with mean μ and precision h , then

$$\tilde{f}_Q = \int_{\mu} \int_h f_N(q|\mu, h) \cdot f_{NG}(\mu, h|m'', w'', n'', v'') d\mu dh \dots\dots\dots(2.3.37)$$

(Wood et al, 1974) suggested that the Bayesian transformation of the normal distribution is the Student-t distribution with parameters as follows:

$$\int_{\mu} \int_h f_N(q|\mu, h) \cdot f_{NG}(\mu, h|m'', w'', n'', v'') d\mu dh = \tilde{f}_Q(q|m'', \frac{n''}{(n''+1)w''}, v'') \dots\dots\dots(2.3.38)$$

To prove this, the following procedure was adopted by the author of this thesis.

Applying equation (2.1.13a), the right side of equation (2.3.38) can be evaluated as:

$$\tilde{f}_Q(q|m'', \frac{n''}{(n''+1)w''}, v'') = \int_0^{\infty} f_N(q|m'', h \frac{n''}{(n''+1)w''}) \cdot f_{\gamma 2}(h|1, v'') dh \dots\dots\dots(2.3.39)$$

By applying the equations (2.1.19) and (2.1.04)

$$f_N(q|m'', hH) = (2\pi)^{-\frac{1}{2}} \exp \left[-\frac{1}{2} Hh (q-m'')^2 \right] \cdot \sqrt{Hh} \dots\dots\dots(2.3.40)$$

$$f_{\gamma 2}(h|1, v'') = \frac{1}{(\frac{1}{2}v''-1)!} \left\{ \exp \left[-\frac{1}{2} h v'' \right] \cdot (\frac{1}{2} h v'')^{\frac{1}{2}v''-1} \cdot \frac{1}{2} v'' \right\} \dots\dots\dots(2.3.41)$$

Substituting equation (2.3.40) and (2.3.41) into equation (2.3.39) and denoting $H = \frac{n''}{(n''+1)w''}$ gives:

$$\begin{aligned} \tilde{f}_Q(q|m'', H, v'') &= \int_0^{\infty} (2\pi)^{-\frac{1}{2}} h^{\frac{1}{2}} \cdot \exp \left[-\frac{1}{2} h H (q-m'')^2 \right] H^{\frac{1}{2}} \\ &\quad \frac{\exp \left[-\frac{1}{2} h v'' \right] \cdot (\frac{1}{2} h v'')^{\frac{1}{2}v''-1} \cdot \frac{1}{2} v'' dh}{(\frac{1}{2}v''-1)!} \dots\dots\dots(2.3.42) \end{aligned}$$

$$= \frac{H^{\frac{1}{2}} (\frac{1}{2}v'')^{\frac{1}{2}v''}}{(\frac{1}{2}v''-1)! (2\pi)^{\frac{1}{2}}} \int_0^{\infty} \exp \left\{ -\frac{1}{2} h [H(q-m'')^2 + v''] \right\} \cdot h^{\frac{1}{2}v''-1+\frac{1}{2}} dh \dots\dots\dots(2.3.43)$$

By making the following substitution.

$$u = \frac{1}{2}h [H(q-m'')^2 + v''] \quad \dots\dots\dots(2.3.44)$$

This implies that

$$\begin{aligned} du &= \frac{1}{2} [H(q-m'')^2 + v''] dh \\ \text{or } dh &= \left\{ \frac{1}{2} [H(q-m'')^2 + v''] \right\}^{-1} du \quad \dots\dots\dots(2.3.45) \end{aligned}$$

$$h = u \left\{ \frac{1}{2} [H(q-m'')^2 + v''] \right\}^{-1} \quad \dots\dots\dots(2.3.46)$$

Substituting equations (2.3.44), (2.3.45) and (2.3.46) into equation (2.3.43) gives:

$$\tilde{f}_Q = \frac{H^{\frac{1}{2}}(\frac{1}{2}v'')^{\frac{1}{2}v''}}{(\frac{1}{2}v''-1)!(2\pi)^{\frac{1}{2}}} \int_0^\infty \exp[-u] \cdot \left\{ \frac{1}{2} [H(q-m'')^2 + v''] \right\}^{-\frac{1}{2}v''+\frac{1}{2}} \cdot u^{\frac{1}{2}v''-\frac{1}{2}} \left\{ \frac{1}{2} [H(q-m'')^2 + v''] \right\}^{-1} du \dots\dots\dots(2.3.47)$$

$$\tilde{f}_Q = \frac{H^{\frac{1}{2}}(\frac{1}{2}v'')^{\frac{1}{2}v''}}{(\frac{1}{2}v''-1)!(2\pi)^{\frac{1}{2}}} \left\{ \frac{1}{2} [H(q-m'')^2 + v''] \right\}^{-\frac{1}{2}v''-\frac{1}{2}} \int_0^\infty e^{-u} u^{\frac{1}{2}v''-\frac{1}{2}} du \dots\dots\dots(2.3.48)$$

$$\tilde{f}_Q = \frac{H^{\frac{1}{2}}(v'')^{\frac{1}{2}v''} (\frac{1}{2})^{\frac{1}{2}v''-\frac{1}{2}v''-\frac{1}{2}}}{(\frac{1}{2}v''-1)!(2\pi)^{\frac{1}{2}}} [H(q-m'')^2 + v'']^{-\frac{1}{2}(v''+1)} \int_0^\infty e^{-u} u^{\frac{1}{2}v''-\frac{1}{2}} du \dots\dots\dots(2.3.49)$$

Multiplying numerator and denominator of the right side of equation (2.3.49) by $(\frac{1}{2}v''-\frac{1}{2})!$ gives:

$$\tilde{f}_Q = \frac{H^{\frac{1}{2}}(v'')^{\frac{1}{2}v''} (\frac{1}{2}v''-\frac{1}{2})!}{(\frac{1}{2}v''-1)!\pi^{\frac{1}{2}}} [H(q-m'')^2 + v'']^{-\frac{1}{2}(v''+1)} \frac{\int_0^\infty e^{-u} u^{\frac{1}{2}v''-\frac{1}{2}} du \dots\dots\dots(2.3.50)}{(\frac{1}{2}v''-\frac{1}{2})!}$$

Applying the equation (2.1.23) for $p = \frac{1}{2}$ and $q = \frac{1}{2}v''$, the following equation is obtained:

$$B(\frac{1}{2}, \frac{1}{2}v'') = \frac{(\frac{1}{2}-1)!(\frac{1}{2}v''-1)!}{(\frac{1}{2}+\frac{1}{2}v''-1)!} = \frac{(-\frac{1}{2})!(\frac{1}{2}v''-1)!}{(\frac{1}{2}v''-\frac{1}{2})!} \dots\dots\dots(2.3.51)$$

Substituting equation (2.1.28) into equation (2.3.51) gives:

$$B(\frac{1}{2}, \frac{1}{2}v'') = \frac{\pi^{\frac{1}{2}} (\frac{1}{2}v''-1)!}{(\frac{1}{2}v''-\frac{1}{2})!} \quad \dots\dots\dots(2.3.52)$$

Substituting equation (2.3.52) into equation (2.3.50) gives:

$$\tilde{f}_Q(q|m'', H, v'') = \frac{H^{\frac{1}{2}}(v'')^{\frac{1}{2}v''}}{B(\frac{1}{2}, \frac{1}{2}v'')} [H(q-m'')^2 + v'']^{-\frac{1}{2}(v''+1)} \int_0^\infty e^{-u} u^{\frac{1}{2}v'' - \frac{1}{2}} du \dots (2.3.53)$$

The integral $\int_0^\infty \frac{e^{-u} u^{\frac{1}{2}(v''-1)}}{(\frac{1}{2}v'' - \frac{1}{2})!} du$ can be evaluated by applying the

general equation (2.1.26) with $K = \frac{1}{2}v'' + \frac{1}{2}$, then $K-1 = \frac{1}{2}v'' - \frac{1}{2} = \frac{1}{2}(v''-1)$,
so that $\int_0^\infty \frac{e^{-u} u^{\frac{1}{2}(v''-1)}}{(\frac{1}{2}v'' - \frac{1}{2})!} du = \frac{\Gamma(\frac{1}{2}v'' + \frac{1}{2})}{(\frac{1}{2}v'' - \frac{1}{2})!} \dots (2.3.54)$

Applying the equation (2.1.27)

$$\Gamma(\frac{1}{2}v'' + \frac{1}{2}) = (\frac{1}{2}v'' + \frac{1}{2} - 1)! = (\frac{1}{2}v'' - \frac{1}{2})!$$

Substituting this in equation (2.3.54)

$$\int_0^\infty \frac{e^{-u} u^{\frac{1}{2}(v''-1)}}{(\frac{1}{2}v'' - \frac{1}{2})!} du = \frac{(\frac{1}{2}v'' - \frac{1}{2})!}{(\frac{1}{2}v'' - \frac{1}{2})!} = 1 \dots (2.3.55)$$

Substituting equation (2.3.55) into (2.3.53) gives

$$\tilde{f}_Q(q|m'', H, v'') = \frac{H^{\frac{1}{2}}(v'')^{\frac{1}{2}v''}}{B(\frac{1}{2}, \frac{1}{2}v'')} [v'' + H(q-m'')^2]^{-\frac{1}{2}(v''+1)} \dots (2.3.56)$$

$$= \frac{H^{\frac{1}{2}}(v'')^{\frac{1}{2}v''}}{B(\frac{1}{2}, \frac{1}{2}v'')} (v'')^{-\frac{1}{2}(v''+1)} \left[1 + \frac{H}{v''} (q-m'')^2 \right]^{-\frac{1}{2}(v''+1)}$$

$$= \frac{1}{B(\frac{1}{2}, \frac{1}{2}v'')} \left[1 + \frac{H}{v''} (q-m'')^2 \right]^{-\frac{1}{2}(v''+1)} \left[\frac{v''}{H} \right]^{-\frac{1}{2}} \dots (2.3.57)$$

Substituting for H

$$\tilde{f}_Q(q|m'', \frac{n''}{(n''+1)w''}, v'') = \frac{1}{B(\frac{1}{2}, \frac{1}{2}v'')} \left[1 + \frac{n''}{n''+1} \cdot \frac{1}{v''w''} (q-m'')^2 \right]^{-\frac{1}{2}(v''+1)}$$

$$\left[\frac{n''+1}{n''} \cdot w''v'' \right]^{-\frac{1}{2}} \dots (2.3.58)$$

Substituting $r = \frac{n''}{n''+1}$

$$\tilde{f}_Q(q|m'', \frac{r}{w''}, v'') = \frac{1}{B(\frac{1}{2}, \frac{1}{2}v'')} \left[1 + \frac{r(q-m'')^2}{w'' v''} \right]^{-\frac{1}{2}(v''+1)} \left[\frac{v'' w''}{r} \right]^{-\frac{1}{2}} \dots (2.3.59)$$

Since the equation (2.3.59) is a Student-t distribution applying the equations (2.1.14) and (2.1.15) remembering that

$$H = \left(\frac{n''}{n''+1} \cdot \frac{1}{w''} \right)$$

its first two moments are given by:

$$E(q) = m'' \quad v'' > 1 \quad \dots\dots\dots(2.3.60)$$

$$V(q) = w'' \left[\frac{v''}{v''-2} \right] \left[\frac{n''+1}{n''} \right] \quad v'' > 2 \quad \dots\dots\dots(2.3.61)$$

Thus inferences about future flood discharges modeled according to normal distribution with unknown mean and variance can be made from equation (2.3.59), the Student-t model which fully accounts for parameter uncertainty (Wood and Rodriguez-Iturbe, 1975a).

The Student-t density function is similar to Normal density function, but it has a larger dispersion than normal density function.

2.3.3 Estimation of Prior Parameters

A normal inverted gamma 2 pdf, which is the conjugate of equation (2.3.15) is the convenient prior pdf for μ and σ . This is because when it is combined with the likelihood function, equation (2.3.15), it yields an analytical tractable posterior pdf. Thus the prior pdf is a product of a normal and an inverted gamma 2 (IG2) pdf (Vicens et al, 1975).

$$f'(\mu, \sigma | I_R) = f'_N(\mu | m', \sigma | n'^{\frac{1}{2}}) \cdot f'_{IG2}(\sigma | w'^{\frac{1}{2}}, v') \quad \dots\dots\dots(2.3.62)$$

where the parameters m' , w' , n' , v' are estimated from prior information I_R .

The marginal pdf for either of the two unknown variables μ and σ can be obtained by 'integrating out' the other variables. (Vicens et al, 1975) showed that in the case of normal and inverted gamma-2 (ig2) distributions, the resulting marginal distribution for μ is a Student-t distribution.

$$f'(\mu | I_R) = \int_0^{\infty} f'(\mu, \sigma | I_R) d\sigma = f_S'(\mu | m', \frac{w'}{n'}, v') \dots (2.3.63)$$

where $f_S'(\mu | m', \frac{w'}{n'}, v')$ is a Student-t distribution with parameters, m' , w' , n' , and v' . The expected value of the student distribution for μ by applying the general equation

(2.1.14) and (2.1.15) are:

$$\begin{aligned} E[\mu | I_R] &= m' & v' > 1 & \dots (2.3.64) \\ V[\mu | I_R] &= \frac{w'v'}{n'(v'-2)} & v' > 2 & \end{aligned}$$

The marginal distribution of σ is inverted gamma-2

distribution

$$f'(\sigma | I_R) = \int_0^{\infty} f'(\mu, \sigma | I_R) d\mu = f_{IG2}'(\sigma | w'^{\frac{1}{2}}, v') \dots (2.3.66)$$

where $f_{IG2}'(\sigma | w'^{\frac{1}{2}}, v')$ is an inverted gamma 2 distribution with parameters (w', v') .

Applying the general equation of moments (2.1.11) and 2.1.12),

the first two moments are:

$$E[\sigma | I_R] = w'^{\frac{1}{2}} (\frac{1}{2} v')^{\frac{1}{2}} \cdot \frac{(\frac{1}{2} v' - 3/2)!}{(\frac{1}{2} v' - 1)!} \dots (2.3.67)$$

$$V[\sigma | I_R] = \frac{v'w'}{v'-2} - \{E[\sigma | I_R]\}^2 \quad v' > 2 \quad \dots (2.3.68)$$

Introducing gamma function given by equation (2.1.27), equation

(2.3.67) can be written as:

$$E[\sigma | I_R] = (\frac{1}{2} v')^{\frac{1}{2}} w'^{\frac{1}{2}} \frac{\Gamma[\frac{1}{2}(v'-1)]}{\Gamma(\frac{1}{2} v')} \quad v' > 1 \quad \dots (2.3.69)$$

From equation (2.3.66), the marginal pdf for the normal model can be obtained through the theory of derived distributions (Vicens et al, 1975)

$$f'(\sigma^2 | I_R) = f_{IG2}'(\sigma^2 | I_R) \cdot \left| \frac{1}{2} \sigma \right| = f_{IG1}'(\sigma^2 | \frac{1}{2} v', \frac{1}{2} v' w') \dots (2.3.70)$$

which is an inverted gamma-1 distribution. By applying the equations (2.1.08) and (2.1.09) with $r = \frac{1}{2}v'$, and $y = \frac{1}{2}v'w'$ the expected value and variance of σ^2 are respectively:

$$E[\sigma^2 | I_R] = \frac{\frac{1}{2}v'w'}{(\frac{1}{2}v'-1)} = \frac{v'w'}{2} \cdot \frac{2}{v'-2} = \frac{v'w'}{v'-2} \quad v' > 2 \quad \dots (2.3.71)$$

$$V[\sigma^2 | I_R] = \frac{(\frac{1}{2}v'w')^2}{(\frac{1}{2}v'-1)^2(\frac{1}{2}v'-2)} = \frac{(v'w')^2}{4} \cdot \frac{4}{(v'-2)^2} \cdot \frac{2}{(v'-4)}$$

$$V[\sigma^2 | I_R] = \frac{2(v'w')^2}{(v'-2)^2(v'-4)} \quad v' > 4 \quad \dots (2.3.72)$$

From equation (2.3.64), $m' = E(\mu | I_R)$

Dividing equation (2.3.71) by equation (2.3.65) gives:

$$\frac{E[\sigma^2 | I_R]}{v[\mu | I_R]} = \frac{v'w'}{v'-2} \cdot \frac{n'(v'-2)}{v'w'}$$

Thus

$$n' = \frac{E[\sigma^2 | I_R]}{v[\mu | I_R]} \quad \dots (2.3.73)$$

Multiplying the square of equation (2.3.71) by 2 and dividing by equation (2.3.72) gives:

$$\frac{2E^2[\sigma^2 | I_R]}{v[\sigma^2 | I_R]} = \frac{2(v'w')^2}{(v'-2)^2} \cdot \frac{(v'-2)^2(v'-4)}{2(v'w')^2}$$

$$v' = 4 + \frac{2E^2[\sigma^2 | I_R]}{v[\sigma^2 | I_R]} \quad \dots (2.3.74)$$

Multiplying both sides of equation (2.3.71) by $\frac{(v'-2)}{v'}$ gives:

$$E[\sigma^2 | I_R] \cdot \frac{v'-2}{v'} = \frac{v'w'}{v'-2} \cdot \left[\frac{v'-2}{v'} \right]$$

Thus

$$w' = \left[\frac{v'-2}{v'} \right] \cdot E[\sigma^2 | I_R] \quad \dots (2.3.75)$$

The corresponding value of v' obtained by (Vicens et al, 1975) was:

$$v' = \frac{2 E^2(\sigma^2 | I_R)}{v(\sigma^2 | I_R)}$$

which is wrong because 4 should be added to the right hand side.

Hence equation (2.3.74) should be used as the correct equation.

2.3.4 Estimation of Expectation and Variance of μ and σ^2

In order to solve equation (2.3.64) and (2.3.73) through (2.3.75), it is necessary to determine the expectations and variances of both μ and σ^2 . To do this, the approximate moment of a function for the multivariate case (Benjamin and Cornell, 1970) is used.

For the expectation of a function

$$Y = g(x_1, x_2, x_3, \dots, x_n) \quad \dots\dots\dots(2.3.76)$$

(Benjamin and Cornell, 1970) used the second order approximation of the Taylor's series expansion and obtained:

$$E[Y] = g(M_{x_1}, M_{x_2}, \dots, M_{x_n}) + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 g}{\partial x_i \partial x_j} \bigg|_m \cdot \text{cov}(x_i, x_j) \dots\dots(2.3.77)$$

where

$x_1, x_2, \dots, x_n$ are the independent variables

y is the dependent variable

$M_{x_1}, M_{x_2}, \dots, M_{x_n}$ are the mean value of the variables

$x_1, x_2, \dots, x_n$

$\frac{\partial^2 g}{\partial x_i \partial x_j} \bigg|_m$ is the mixed second order partial derivative of $g(x_1, x_2, \dots, x_n)$ with respect to x_i and x_j evaluated at $M_{x_1}, M_{x_2}, \dots, M_{x_n}$.

If the variables $x_1, x_2, \dots, x_n$, are uncorrelated then the second term of the right side of equation (2.3.77) vanishes, giving:

$$E[Y] = \varepsilon(M_{x_1}, M_{x_2}, \dots, M_{x_n}) \dots\dots\dots(2.3.78)$$

In order to determine the variance of Y , for the multivariate case, (Benjamin and Cornell, 1970) used the first - order approximation and obtained:

$$\text{Var}[Y] = \sum_{i=1}^n \sum_{j=1}^n \frac{\partial \varepsilon}{\partial x_i} \Big|_m \cdot \frac{\partial \varepsilon}{\partial x_j} \Big|_m \cdot \text{cov}[x_i, x_j] \dots\dots\dots(2.3.79)$$

If x_i and x_j are uncorrelated, then the variance of Y is simply:

$$\text{Var}[Y] = \sum_{i=1}^n \left[\frac{\partial \varepsilon}{\partial x_i} \Big|_m \right]^2 \cdot \text{var}[x_i] \dots\dots\dots(2.3.80)$$

For a special case where Y is a function of two variables x_1 and x_2 , equations (2.3.77) and (2.3.79) can be written respectively as:

$$\begin{aligned} E[Y] &= \varepsilon(M_{x_1}, M_{x_2}) + \frac{1}{2} \left\{ \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 \varepsilon}{\partial x_i \partial x_j} \Big|_m \cdot \text{cov}[x_i, x_j] \right\} \\ &= \varepsilon(M_{x_1}, M_{x_2}) + \frac{1}{2} \left\{ \frac{\partial^2 \varepsilon}{\partial x_1^2} \Big|_m \cdot \text{Var}[x_1] + \frac{\partial^2 \varepsilon}{\partial x_1 \partial x_2} \Big|_m \cdot \text{Cov}[x_1, x_2] + \right. \\ &\quad \left. \frac{\partial^2 \varepsilon}{\partial x_2 \partial x_1} \Big|_m \cdot \text{Cov}[x_2, x_1] + \frac{\partial^2 \varepsilon}{\partial x_2^2} \Big|_m \cdot \text{Var}[x_2] \right\} \dots\dots\dots(2.3.81) \end{aligned}$$

$$\begin{aligned} \text{Var}[Y] &= \sum_{i=1}^2 \sum_{j=1}^2 \frac{\partial \varepsilon}{\partial x_i} \Big|_m \cdot \frac{\partial \varepsilon}{\partial x_j} \Big|_m \cdot \text{Cov}[x_i, x_j] \\ &= \left[\frac{\partial \varepsilon}{\partial x_1} \Big|_m \right]^2 \cdot \text{Var}[x_1] + \frac{\partial \varepsilon}{\partial x_1} \Big|_m \cdot \frac{\partial \varepsilon}{\partial x_2} \Big|_m \cdot \text{Cov}[x_1, x_2] + \end{aligned}$$

$$\frac{\partial \varepsilon}{\partial x_2} \Big|_m \cdot \frac{\partial \varepsilon}{\partial x_1} \Big|_m \cdot \text{Cov}[x_2, x_1] + \left[\frac{\partial \varepsilon}{\partial x_2} \Big|_m \right]^2 \text{Var}[x_2] \dots\dots\dots(2.3.82)$$

From regional frequency analysis (Odusote, 1989), the mean annual discharge Q_m is a function of both drainage area, A and the length of the main stream L . Substituting in equations (2.3.81) and (2.3.82) gives:

$$E [\mu_Q] \approx Q(A_m, L_m) + \frac{1}{2} \left\{ \frac{\partial^2 Q}{\partial A^2} \Big|_m \cdot \text{var}[A] + \frac{\partial^2 Q}{\partial A \partial L} \Big|_m \cdot \text{cov}[A, L] + \frac{\partial^2 Q}{\partial L^2} \Big|_m \cdot \text{var}[L] + \frac{\partial^2 Q}{\partial L \partial A} \Big|_m \cdot \text{cov}[L, A] \right\} \dots\dots\dots(2.3.83)$$

where A_m and L_m are the mean values of A and L respectively.

$$\text{Var} [\mu_Q] \approx \left[\frac{\partial Q}{\partial A} \Big|_m \right]^2 \cdot \text{var}[A] + \frac{\partial Q}{\partial A} \Big|_m \cdot \frac{\partial Q}{\partial L} \Big|_m \cdot \text{cov}[A, L] + \frac{\partial Q}{\partial L} \Big|_m \cdot \frac{\partial Q}{\partial A} \Big|_m \cdot \text{cov}[L, A] + \left[\frac{\partial Q}{\partial L} \Big|_m \right]^2 \cdot \text{var}[L] \dots\dots(2.3.84)$$

From the regional frequency analysis (Odusote, 1989), the relation between the discharge, Q , the drainage area, A and main stream length, L is given by equation (1.1) as:

$$Q = a A^b L^c,$$

where a , b and c are constants to be determined from multiple regression analysis.

When this equation is substituted into equations (2.3.83) and (2.3.84) the following equations are obtained:

$$E [\mu_Q] \approx a A_m^b L_m^c + \frac{1}{2} \left\{ ab(b-1) A_m^{b-2} L_m^c \cdot \text{var}[A] + abc A_m^{b-1} L_m^{c-1} \cdot \text{cov}[A, L] + abc A_m^{b-1} L_m^{c-1} \cdot \text{cov}[L, A] + ac(c-1) A_m^b L_m^{c-2} \cdot \text{var}[L] \right\} \dots\dots\dots(2.3.85)$$

$$\text{Var} [\mu_Q] \approx (ab A_m^{b-1} L_m^c)^2 \cdot \text{var}[A] + 2a^2 bc A_m^{2b-1} L_m^{2c-1} \cdot \text{cov}[A, L] + (ac A_m^b L_m^{c-1})^2 \cdot \text{var}[L] \dots\dots\dots(2.3.86)$$

2.3.5 Posterior Parameters

Having evaluated the prior parameters as described above, the posterior parameters n'' , v'' , m'' and w'' can be calculated from equations (2.3.33) through (2.3.36).

These values obtained are then substituted into equation (2.3.59) to obtain the Student- density function from which the return period can be calculated.

2.3.6 Classical Procedure for Normal Model

The Maximum Likelihood Estimate (MLE) method is used to estimate the parameters μ and h . The normal distribution is given by equation (2.3.1) and the likelihood function is given by equation (2.3.5) as:

$$L(\mu, h|Q) = (2\pi)^{n/2} \exp \left[-\frac{1}{2h} \sum_{i=1}^n (q_i - \mu)^2 \right] \cdot h^{n/2}$$

where $h = \sigma^{-2}$

Taking the natural logarithm of both sides of equation (2.3.5) gives:

$$\ln [L(\mu, h|Q)] = -\frac{n}{2} \ln(2\pi) - \frac{1}{2h} \sum_{i=1}^n (q_i - \mu)^2 + \frac{n}{2} \ln(h) \dots (2.3.87)$$

Differentiating both sides of equation (2.3.87) with respect to μ and equating to zero.

$$\begin{aligned} \frac{\partial \ln [L(\mu, h|Q)]}{\partial \mu} &= h \sum_{i=1}^n (q_i - \hat{\mu}) + 0 = 0 \\ \sum_{i=1}^n (q_i - \hat{\mu}) &= 0 \\ \sum_{i=1}^n q_i - n \hat{\mu} &= 0 \\ \hat{\mu} &= 1/n \sum_{i=1}^n q_i \dots \dots \dots (2.3.88) \end{aligned}$$

This is an unbiased estimate of μ .

Differentiating both sides of equation (2.3.87) with respect to h and equating to zero.

$$\frac{\partial \ln [L(\mu, h|Q)]}{\partial h} = -\frac{1}{2} \sum_{i=1}^n (q_i - \mu)^2 + \frac{n}{2} \cdot \frac{1}{h} = 0$$

$$\hat{h} = \frac{n}{\sum_{i=1}^n (q_i - \hat{\mu})^2} \dots\dots\dots(2.3.89)$$

This is a biased estimate of h .

$$\text{The unbiased estimate is } \hat{h} = \frac{n-1}{\sum_{i=1}^n (q_i - \hat{\mu})^2} \dots\dots\dots(2.3.90)$$

The values of the estimators $\hat{\mu}$ and $\hat{h}$ are substituted into equation (2.3.1.) to give:

$$f_N(q | \mu, h) = (2\pi)^{-\frac{1}{2}} \exp \left[-\frac{1}{2} \hat{h} (q - \hat{\mu})^2 \right] \cdot \hat{h}^{\frac{1}{2}} \dots\dots\dots(2.3.91)$$

Equation (2.3.91) is the probability density function of the normal model utilized in the classical method of analysis.

2.4 PARAMETER UNCERTAINTY FOR LOG-NORMAL MODEL2.4.1 Density Function, Likelihood Function and Distribution of (μ, h)

An appropriate model which seems to fit the sample information of many rivers is one which satisfies the following criteria:

- (i) the probability distribution takes values greater than zero only for positive flood discharges ($q > 0$)
- (ii) the distribution is positively skewed (long upper tail).

An example of such a distribution is the log normal distribution. In this distribution, we define

$$X_i = \ln q_i \quad \dots\dots\dots(2.4.1)$$

which is normally distributed with mean μ and precision h . Then q_i is distributed log-normally by definition. The probability density function for q is (Wood and Rodriguez-Iturbe, 1975b):

$$f(q | \mu, h) = q^{-1} (2\pi)^{-\frac{1}{2}} h^{\frac{1}{2}} \exp \left[-\frac{1}{2} h(x-\mu)^2 \right] \dots\dots\dots(2.4.2)$$

The likelihood function for μ and h , given set of independent observations, Q is

$$L(\mu, h | Q) = \frac{1}{q_1} (2\pi)^{-\frac{1}{2}} h^{\frac{1}{2}} \exp \left[-\frac{1}{2} h(x_1 - \mu)^2 \right] \cdot \frac{1}{q_2} (2\pi)^{-\frac{1}{2}} h^{\frac{1}{2}} \exp \left[-\frac{1}{2} h(x_2 - \mu)^2 \right] \dots \frac{1}{q_n} (2\pi)^{-\frac{1}{2}} h^{\frac{1}{2}} \exp \left[-\frac{1}{2} h(x_n - \mu)^2 \right] \dots\dots(2.4.3)$$

$$\text{where } Q = q_1, q_2 \dots q_n$$

Simplifying the right side of equation (2.4.3) gives:

$$L(\mu, h | Q) = \left[\prod_{i=1}^n q_i \right]^{-1} (2\pi)^{-n/2} \cdot h^{n/2} \exp \left[-\frac{1}{2} h \sum (x_i - \mu)^2 \right] \dots\dots(2.4.4)$$

$$\text{where } \prod_{i=1}^n q_i = q_1 \cdot q_2 \cdot q_3 \dots q_n.$$

Following the same procedure as for the normal distribution, the Bayesian transformed form of the log-normal model can be shown to be (Wood and Rodriguez-Iturbe, 1975a) log Student - t distribution of the form:

$$f_Q(q|m'', \frac{w''}{r}, v'') = \frac{1}{q} \cdot \frac{1}{B(\frac{1}{2}, \frac{1}{2}v'')} \left[1 + r \frac{(\ln q - m'')^2}{w'' v''} \right]^{-\frac{1}{2}(v''+1)} \left[\frac{r}{w'' v''} \right]^{\frac{1}{2}} \quad \dots(2.4.5)$$

$$\text{where } r = \frac{n''}{n'' + 1}$$

Prior and Posterior Parameters of the log-Normal distribution are calculated in the same way as for Normal distribution.

2.4.2 Classical Procedure for Log-Normal Model

The maximum likelihood method is used to estimate the parameters μ and h . The log-normal distribution is given by equation (2.4.2) and the likelihood function is given by equation (2.4.4) as:

$$L(\mu, h|Q) = \left[\prod_{i=1}^n q_i \right]^{-1} (2\pi)^{-n/2} \cdot h^{n/2} \exp \left[-\frac{1}{2}h \sum_{i=1}^n (\ln q_i - \mu)^2 \right]$$

Taking the natural logarithm of both sides of equation (2.4.4) gives:

$$\ln L(\mu, h|Q) = -\ln \prod_{i=1}^n q_i - \frac{n}{2} \ln(2\pi) + \frac{n}{2} \ln h - \frac{1}{2}h \sum_{i=1}^n (\ln q_i - \mu)^2 \quad \dots\dots(2.4.6)$$

Differentiating both sides of equation (2.4.6) partially with respect to μ and equating to zero gives:

$$\frac{\partial \ln [L(\mu, h|Q)]}{\partial \mu} = h \sum_{i=1}^n (\ln q_i - \mu) = 0 \quad \dots\dots(2.4.7)$$

$$\sum_{i=1}^n \ln q_i - n \hat{\mu} = 0$$

$$\text{or } \hat{\mu} = \frac{\sum_{i=1}^n (\ln q_i)}{n} \quad \dots\dots\dots(2.4.8)$$

Differentiating both sides of equation (2.4.6) partially with respect to h and equating to zero gives:

$$\frac{\partial \ln [L(\mu, h|Q)]}{\partial h} = \frac{n}{2} \cdot \frac{1}{h} - \frac{1}{2} \sum_{i=1}^n (\ln q_i - \hat{\mu})^2 = 0 \quad \dots\dots(2.4.9)$$

$$\text{or } \hat{h} = \frac{n}{\sum_{i=1}^n (\ln q_i - \hat{\mu})^2} \quad \dots\dots\dots(2.4.10)$$

This is the biased estimate of the precision h . The unbiased estimate of h is:

$$\hat{h} = \frac{n-1}{\sum_{i=1}^n (\ln q_i - \hat{\mu})^2} \dots\dots\dots(2.4.11)$$

The values of the estimators μ and h are substituted into equation (2.4.2) to give:

$$f_{IN}(q | \hat{\mu}, \hat{h}) = q^{-1} (2\pi)^{-\frac{1}{2}} \exp \left[-\frac{1}{2\hat{h}}(q - \hat{\mu})^2 \right] \cdot \hat{h}^{\frac{1}{2}} \dots\dots\dots(2.4.12)$$

Equation (2.4.12) is the probability density function of the log-normal model utilized in the classical method of analysis.

2.5 PARAMETER UNCERTAINTY FOR EXCEEDANCE MODEL

2.5.1 Density Function and Likelihood Functions

While the Normal and log-normal distribution models of the flood events were based on annual series, that is, a series consisting of largest flood annually, the exceedance model is based on partial duration series, that is, a series consisting of flood peaks above a certain level Q_b .

The probability distribution of the entire series may be unknown. (Wood and Rodriguez-Iturbe, 1975a) suggested that the base level flood discharge, Q_b , be chosen large enough so that the probability distribution of the flood events greater than Q_b will be assumed to be exponentially distributed. That is:

$$f_Q(q) = \alpha \cdot \exp[-\alpha(q - Q_b)] \quad \text{for } q > Q_b \quad \dots\dots\dots(2.5.1)$$

where α is the event magnitude distribution defined as:

$$\alpha = \frac{1}{q - Q_b}$$

$$\text{Let } z = q - Q_b \quad \dots\dots\dots(2.5.2)$$

where z is the exceedance discharge, then substituting equation (2.5.2) into equation (2.5.1) gives:

$$f_z(z) = \alpha \exp(-\alpha z), \quad \text{for } z \geq 0 \quad \dots\dots\dots(2.5.3)$$

The probability that an exceedance discharge will be greater than Z is then:

$$\begin{aligned} P(z) &= \int_z^{\infty} \alpha \exp[-\alpha z] dz = \left[-e^{-\alpha z} \right]_z^{\infty} = \left[-e^{-\alpha \infty} \right] - \left[-e^{-\alpha z} \right] \\ &= 0 + e^{-\alpha z} \end{aligned}$$

$$\therefore P(z) = \exp[-\alpha z] \quad \dots\dots\dots(2.5.4)$$

If it is assumed that time between independent flood peaks larger than the base level flood discharge Q_b is exponentially distributed with an average arrival rate v , then the occurrence of exceedance flows will be governed by Poisson Process.

The occurrence of floods above some exceedance level z is also a Poisson process but one with an average arrival rate $v \cdot P_z$.

Thus the probability that in time t , n_z exceedances
(n exceedances above a flood exceedance level z) is just
(Wood and Rodriguez-Iturbe, 1975a);

$$p_N(N = N_z) = (v p_z)^{n_z} \cdot \frac{\exp[-v p_z t]}{n_z!} \dots\dots\dots(2.5.5.)$$

But if z is such that no exceedances occurs in time t , then
 $p_N(N = 0)$ is the cumulative density function for z , $F_z(z)$.

Substituting $N_z = 0$ in equation (2.5.5) gives:

$$F_z(z) = (v p_z)^0 \frac{\exp[-v p_z t]}{0!}$$

$$F_z(z) = \exp[-vt e^{-\alpha z}] \quad \text{for } z \geq 0 \quad \dots\dots\dots(2.5.6)$$

The probability that $Z = 0$ is the probability that q is less
than the base value, Q_b . If z is large so that the probability
of exceeding it is small and the arrival rate of such event is
small, (Wood et al, 1974), have developed an approximate equation
to equation (2.5.6) as:

$$F_z(z) \approx 1 - vt \exp[-\alpha z] \quad \dots\dots\dots(2.5.7)$$

or

$$G_z(z) \approx 1 - F_z(z) = vt \exp[-\alpha z] \quad \dots\dots\dots(2.5.8)$$

Equation (2.5.7) represents the underlying probabilistic
distribution of the exceedance flood. This distribution serves
both the classical and Bayesian analysts but in slightly different
ways.

The Bayesian analyst obtains the posterior pdf for a parameter
by applying equation (2.2.3) and assuming that α and v are
independent.

To obtain the approximate density function of the exceedance model, equation (2.5.7) is differentiated with respect to z to give:

$$f_z(z) = \frac{d F_z(z)}{dz} = \alpha v t \exp(-\alpha z) \dots\dots\dots(2.5.9)$$

The likelihood function for the exceedance parameter " α " is gamma.

$$L[\alpha | z(n, T)] = \prod_{z_n} f(z) \dots\dots\dots(2.5.10)$$

$$L[\alpha | z(n, T)] = v^n t^n \{ \alpha \cdot e^{-\alpha z_1} \cdot \alpha e^{-\alpha z_2} \dots \alpha e^{-\alpha z_n} \}..$$

$$L[\alpha | z(n, T)] = v^n t^n \alpha^n \exp \alpha \sum_{i=1}^n z_i \dots\dots\dots(2.5.11)$$

Thus from equation (2.5.11) it can be deduced that:

$$L[\alpha | z(n, T)] \propto \alpha^n \cdot \exp \left[-\alpha \sum_{i=1}^n z_i \right] \dots\dots\dots(2.5.12)$$

where n is the number of exceedance discharges above the base level Q_b in record of T years, $\propto$ means "proportional to"

The natural conjugate prior is gamma-1 distributed with parameters w' and l'

Thus by applying the equation (2.1.1)

$$f'_{\gamma 1}(\alpha | w', l') = \frac{\exp[-\alpha l'] \cdot (\alpha l')^{w'} \cdot l'}{w'!} \dots\dots\dots(2.5.13)$$

Also by applying equations (2.1.02) and (2.1.03) the first two moments are:

$$E[\alpha] = \frac{w' + 1}{l'} \dots\dots\dots(2.5.14)$$

$$V[\alpha] = \frac{w' + 1}{l'^2} \dots\dots\dots(2.5.15)$$

From equation (2.5.13) it can be deduced that:

$$f'_{\gamma_1}(\alpha | w', l') \propto \alpha^{w'} \exp(-\alpha l') \dots\dots\dots (2.5.16)$$

Similarly, the posterior density function of α will be gamma-1 distributed and by applying equations (2.1.01), (2.1.02) and (2.1.03), the distribution and first two moments are respectively given by:

$$f''_{\gamma_1}(\alpha | w'', l'') = \frac{\exp(-\alpha l'') (\alpha l'')^{w''}}{w''!} \dots\dots\dots (2.5.17)$$

$$E[\alpha] = \bar{\alpha} = \frac{w'' + 1}{l''} \dots\dots\dots (2.5.18)$$

$$V[\alpha] = \frac{w'' + 1}{l''^2} \dots\dots\dots (2.5.19)$$

From equation (2.5.17) it can be deduced that:

$$f''_{\gamma_1}(\alpha | w'', l'') \propto \alpha^{w''} \exp(-\alpha l'') \dots\dots\dots (2.5.20)$$

Applying equation (2.2.3) gives:

$$f''_{\gamma_1}(\alpha | w'', l'') = L[\alpha | z(n, T)] \cdot f'_{\gamma_1}(\alpha | w', l') \dots\dots\dots (2.5.21)$$

Substituting equations (2.5.12), (2.5.16) and (2.5.20) into equation (2.5.21) gives:

$$\begin{aligned} \alpha^{w''} \exp(-\alpha l'') &\propto \alpha^n \exp(-\alpha \sum_{i=1}^n z_i) \cdot \alpha^{w'} \exp(-\alpha l') \\ \text{or} \\ \alpha^{w''} \exp(-\alpha l'') &\propto \alpha^{(w'+n)} \exp\left[-\alpha\left(l' + \sum_{i=1}^n z_i\right)\right] \dots\dots\dots (2.5.22) \end{aligned}$$

equating the powers of the exponents

$$l'' = l' + \sum_{i=1}^n z_i \dots\dots\dots (2.5.23)$$

Equating powers of α :

$$w'' = w' + n \dots\dots\dots (2.5.24)$$

To obtain the density of the exact exceedance model, the equation (2.5.6) is differentiated with respect to z

$$f_z(z) = \frac{dF_z(z)}{dz} = vt \alpha e^{-\alpha z} \exp\left[-vt e^{-\alpha z}\right] \dots\dots\dots (2.5.25)$$

The likelihood function for the exceedance model parameter

"v" is gamma.

$$L [v|z(n, T)] = \prod_{z_n} f(z)$$

$$\begin{aligned} L [v|z(n, T)] &= VT \alpha e^{-\alpha z_1} \cdot e^{-VT e^{-\alpha z_1}} VT \alpha e^{-\alpha z_2} \cdot e^{-VT e^{-\alpha z_2}} \dots VT \alpha e^{-\alpha z_n} \cdot e^{-VT e^{-\alpha z_n}} \\ &= v^n T^n \alpha^n \cdot e^{-\alpha(z_1 + z_2 + \dots + z_n)} e^{-VT(e^{-\alpha z_1} + e^{-\alpha z_2} + \dots + e^{-\alpha z_n})} \end{aligned}$$

$$L[v|z(n, T)] = v^n T^n \alpha^n \cdot \exp \left[-\alpha \sum_{i=1}^n z_i \right] \cdot \exp \left\{ -VT \sum_{i=1}^n \exp[-\alpha z_i] \right\} \dots (2.5.26)$$

$\sum e^{-\alpha z_i}$ is the sum of the probabilities that the exceedance discharges $z_1, z_2, \dots, z_n$ will be greater than z . This is approximately equal to 1.

That is

$$\sum e^{-\alpha z_i} \simeq 1.$$

Substituting this into equation (2.5.26) gives:

$$L [v|z(n, T)] = v^n T^n \alpha^n \exp \left[-\alpha \sum_{i=1}^n z_i \right] \cdot e^{-VT} \dots (2.5.27)$$

From equation (2.5.27) it can be deduced that:

$$L [v|z(n, T)] \propto e^{-VT} \cdot v^n \dots (2.5.28)$$

The natural conjugate is also a gamma-1 distribution with parameters u' and s' . Substituting $r-1 = u'$, $y = s'$ and $z = v$ into equation (2.1.01) gives:

$$f'_{\gamma_1}(v|u', s') = \frac{e^{-vs'} (vs')^{u'} \cdot s'}{u'!} \dots (2.5.29)$$

The corresponding first two moments of the above gamma-1 distribution are:

$$E [V] = \frac{u' + 1}{s'} \dots (2.5.30)$$

$$V [V] = \frac{u' + 1}{s'^2} \dots (2.5.31)$$

From equation (2.5.29), it can be deduced that:

$$f'_{\gamma_1}(v|u', s') \propto e^{-vs'} \cdot v^{u'} \dots\dots\dots (2.5.32)$$

The posterior density function for v is also gamma-1 distributed and by similar substitutions as above, the following equation is obtained:

$$f''_{\gamma_1}(v|u'', s'') = \frac{\exp(-vs'') (vs'')^{u''} \cdot s''}{u''!} \dots\dots\dots (2.5.33)$$

and the corresponding first two moments are:

$$E[v] = \bar{v} = \frac{u'' + 1}{s''} \dots\dots\dots (2.5.34)$$

$$V[v] = \frac{u'' + 1}{s''^2} \dots\dots\dots (2.5.35)$$

From equation (2.5.33) it can be deduced that:

$$f''_{\gamma_1}(v|u'', s'') \propto e^{-vs''} \cdot v^{u''} \dots\dots\dots (2.5.36)$$

Applying the equation (2.2.3)

$$f''_{\gamma_1}(v|u'', s'') = L[v|z(n, T)] \cdot f'_{\gamma_1}(v|u', s') \dots\dots\dots (2.5.37)$$

Substituting equations (2.5.28), (2.5.32) and (2.5.36) into equation (2.5.37) gives:

$$\begin{aligned} e^{-vs''} \cdot v^{u''} &= e^{-vT} \cdot v^n \cdot e^{-vs'} \cdot v^{u'} \\ e^{-vs''} \cdot v^{u''} &= e^{-v(T + s')} \cdot v^{(n + u')} \dots\dots\dots (2.5.38) \end{aligned}$$

Equating coefficients of the exponential terms gives:

$$s'' = s' + T \dots\dots\dots (2.5.39)$$

Equating powers of v gives:

$$u'' = u' + n \dots\dots\dots (2.5.40)$$

2.5.2 Distribution (α, v)

To find the Bayesian transformed form of the exceedance distribution, equation (2.2.4) is applied as for the normal distribution.

$$\tilde{f}_z(z) = \int \int_{\alpha, v} f_z(z|\alpha, v) \cdot f_{\gamma_1}''(\alpha|w'', l'') \cdot f_{\gamma_1}''(v|u, s'') dv d\alpha \quad \dots (2.5.41)$$

where $f_z(z|\alpha, v)$ is given by equation (2.5.9)

(Wood et al, 1974) performed the integration and obtained the following as a solution to the equation (2.5.41).

$$\tilde{f}_z(z) = \bar{\alpha} \bar{v} t \left[1 + \frac{\bar{\alpha} z}{w''+1} \right]^{-(w''+2)} \quad \dots (2.5.42)$$

where $\bar{\alpha}$, $\bar{v}$ are as defined by equations (2.5.18) and (2.5.34) respectively; that is, they are the expected values of α and v respectively.

In the same manner the exceedance probability $\hat{G}_z(z) = 1 - F_z(z)$ can be calculated from:

$$\hat{G}_z(z) = \int \int_{v, \alpha} vt \exp(-\alpha z) \cdot f_{\gamma_1}''(v|u'', s'') \cdot f_{\gamma_1}''(\alpha|w'', l'') dv d\alpha \quad \dots (2.5.43)$$

(Wood et al, 1974), also integration equation (2.5.43) and obtained the following:

$$\hat{G}_z(z) = \bar{v} t \left[1 + \frac{\bar{\alpha} z}{w''+1} \right]^{-(w''+1)} \quad \dots (2.5.44)$$

$$\text{From equation (2.5.18) } \frac{\bar{\alpha}}{w''+1} = \frac{1}{l''} \quad \dots (2.5.45)$$

Substituting equation (2.5.45) into (2.5.44) gives:

$$\hat{G}_z(z) = \bar{v} t \left[1 + \frac{z}{l''} \right]^{-(w''+1)} \quad \dots (2.5.46)$$

Equation (2.5.46) is the Bayesian transformed form of the exceedance distribution which accounts for the parameter uncertainty.

2.5.3 Estimation of Prior Parameter of Exceedance Model

The prior parameters of the exceedance model are grouped into two classes; (a) those associated with event magnitude distribution, α , and (b) those associated with average arrival rate, v .

2.5.3 (a) Prior Parameters for α

The mean exceedance flood, Z_m can be related to the event magnitude distribution, (Wood, 1973) by:

$$Z_m = \frac{1}{\alpha} \quad \dots\dots\dots (2.5.47)$$

Since α is gamma-1 distributed as discussed above, it follows that Z_m may be assumed to be inverted gamma-1 distributed. By making the following substitutions:

$$t = Z_m, \quad r - 1 = w' \quad \text{and} \quad y = l'$$

into the general equation (2.1.07), the following equation results:

$$f_{ig1}(Z_m | w', l') = \frac{\exp(-l'/Z_m)(l'/Z_m)^{w'+2}}{w'!} \cdot \frac{1}{l'} \quad \begin{matrix} Z_m \geq 0 \\ w', l' > 0 \end{matrix} \quad \dots\dots\dots (2.5.48)$$

The first two moments corresponding to the above gamma-1 function are:

$$E[Z_m] = \frac{l'}{w'} \quad \dots\dots\dots (2.5.49)$$

$$V[Z_m] = \frac{l'^2}{(w')^2(w'-1)} \quad \dots\dots\dots (2.5.50)$$

Dividing the square of equation (2.5.49) by equation (2.5.50) gives:

$$\frac{E^2[Z_m]}{V[Z_m]} = \frac{l'^2}{w'^2} \cdot \frac{w'^2(w'-1)}{l'^2}$$

or

$$w' = 1 + \frac{E^2[Z_m]}{V[Z_m]} \quad \dots\dots\dots (2.5.51)$$

From equation (2.5.49)

$$l' = w' E[Z_m] \quad \dots\dots\dots (2.5.52)$$

The expectation $E[Z_m]$ and the variance $V[Z_m]$ are evaluated from regression model obtained from regional information as described in chapter three.

2.5.3 (b) Prior Parameters for V

The estimation of prior information on V, the average arrival rate involves, as a first step, the estimation of the first two central moments of the arrival rate of a peak flow that will exceed the base level discharge, Q_b .

There are some probabilistic or statistical methods one may use to approach the problem. For example, the Maximum Likelihood Estimate (MLE) method may be applied to determine V at each station and the results tabulated. The tabulated values of V may be used by the engineer as a basis to make statements such as (Wood and Rodriguez-Iturbe, 1975a) "based upon my experience in the area, my best estimate of V is ' ϕ ' and there is a 50-50 chance that V could be plus or minus t of ϕ ".

The implication of the statement is that the variance,

$\text{Var}[V]$ is:

$$\text{Var}[V] = \frac{1}{n-2} (2 e^2) \quad \dots\dots\dots(2.5.53)$$

and the expectation $E[V]$ is

$$E[v] = \phi \quad \dots\dots\dots(2.5.54)$$

Substituting equations (2.5.54) and (2.5.53) into equations (2.5.30) and (2.5.31) give respectively:

$$E[V] = \phi = \frac{u' + 1}{s'} \quad \dots\dots\dots(2.5.55)$$

$$\text{Var}[V] = \frac{1}{n-2} (2 e^2) = \frac{u' + 1}{s'^2} \quad \dots\dots\dots(2.5.56)$$

Dividing equation (2.5.28) by equation (2.5.29) gives:

$$\frac{E[V]}{\text{Var}[V]} = \frac{u' + 1}{s'} \times \frac{s'^2}{u' + 1}$$

$$s' = \frac{E[V]}{\text{Var}[V]} \dots\dots\dots(2.5.57)$$

Rearranging equation (2.5.55),

$$u' = s' E[V] - 1 \dots\dots\dots(2.5.58)$$

2.5.4 Posterior Parameters of Exceedance Model

The posterior parameter for α can be evaluated, using equations (2.5.23) and (2.5.24) while the posterior parameters for V can be evaluated using equations (2.5.39) and (2.5.40).

2.5.5 Classical Procedure for the Exceedance Model

Estimators for both the parameter of the event magnitude distribution, α , and the average arrival rate, V , can be obtained by application of the Maximum Likelihood Estimate (MLE) criterion.

From equation (2.5.12), the likelihood function for α may be written as:

$$L = l[\alpha|z(n, T)] = k_1 \alpha^n \exp \left[-\alpha \sum_{i=1}^n z_i \right] \dots\dots\dots(2.5.59)$$

where k_1 is a constant of proportionality

The estimator of α , $\hat{\alpha}$ may be evaluated by taking the natural logarithm of both sides of equation (2.5.59).

$$\ln[L] = \ln k_1 + n \ln \alpha - \alpha \sum_{i=1}^n z_i$$

Differentiating partially with respect to α and equating to zero gives:

$$\frac{\partial \ln L}{\partial \alpha} = 0 + \frac{n}{\hat{\alpha}} - \sum_{i=1}^n z_i = 0$$

$$\hat{\alpha} = \frac{n}{\sum_{i=1}^n z_i} \dots\dots\dots(2.5.60)$$

Also from equation (2.5.28), the likelihood function for V may be written as:

$$L = 1 [V | z(n, T)] = k e^{-VT} V^n \dots\dots\dots(2.5.61)$$

where k is a constant of proportionality.

The estimator of V, $\hat{V}$ may be evaluated by taking the natural logarithm of both sides of equation (2.5.61).

$$\ln[L] = \ln[K] + n \ln[V] - VT$$

Differentiating partially with respect to V and equating to zero gives:

$$\frac{\partial \ln L}{\partial V} = 0 + \frac{n}{\hat{V}} - T$$

from which

$$\hat{V} = \frac{n}{T} \dots\dots\dots(2.5.62)$$

The values of $\hat{\alpha}$ and $\hat{V}$ obtained from equation (2.5.60) and (2.5.62) respectively are then substituted into equation (2.5.8) to give:

$$G_z(z) = 1 - F_z(z) = \hat{V} t \exp \left[- \hat{\alpha} z \right] \dots\dots\dots(2.5.63)$$

Equations (2.5.46) and (2.5.63) is the probability density function of the exceedance model utilized in the classical method of analysis.

CHAPTER 3

REGIONAL FREQUENCY ANALYSIS

3.1 DESCRIPTION OF STUDY AREA

ALTS and LIDECO ITALY, (1983) have the description of the Taraba river basin and Gassol gaging station as follows:

3.1.1 TARABA RIVER - DESCRIPTION OF THE BASIN

The Taraba is the first large tributary on the left side of the Benue in Nigeria. Its source is on the slopes of the Gotel Mountains and Mambilla Plateau, beginning as various streams which join into one in the large basin at the foot of the Mambilla Plateau escarpment, generally known as the Gashaka Plains (Figure 3.1a).

The Taraba then proceeds north until it joins the Yim River, half-way between Beli and Serti. The Yim, too, has its source in the Gotel Mountains and runs north, and is separated from the Taraba by the Filinga Plateau. As it proceeds it collects a few tributaries which come from the region of the Cameroun border; subsequently it turns west and joins the Taraba (Figure 3.1b).

After having been joined by the Yim, the Taraba runs towards the west, collecting only minor tributaries as far as Beli.

Then, as it proceeds it is joined by the Gazabu, a tributary flowing from the left which descends from the extreme ramifications North-West of the Mambilla Plateau, and by the Kam on the right. The Kam drains a vast mountainous area and stretches out in an East-North-East direction as far as the Shebshi Mountains.

Having passed the gorge of Garin Dali, the valley opens out, carved in modest relief with old deposits from the Benue valley until, having left at Gassol the last ramifications of this area in relief, it pours into the Benue divided into several branches. The last reach is elevated with respect to the alluvial plain of the Benue, and there rest in it traces of various ancient water courses.



FIG.3-1a Map of Upper Benue River Basin showing the aquing stations, relief and topographic

The watershed starting from Gassol follows the gentle ridge which rises as far as Garba-Chede then runs eastwards to the Shebshi Mountains, and from this point, turning towards South-East, it goes as far as the Cameroun border, which it follows more or less to the Gotel Mountains; then it follows the ridge of the Mambilla Plateau as far as Nguroje and Kafshishi Hill overlooking Mai Samari. From here it descends into the valley of the Benue towards Dendirdi (see Figure 3.1a), North-North-West.

In the area of the junction with the Benue the watershed is lost in the swampy ground.

From a geological point of view, it is important to distinguish the area downstream of Garin Dali, on soils deriving from sedimentary rocks, with low relief and gentle morphology, from the area upstream from this village deriving from the crumbling of rocks which constitute the basement complex. In this section are frequent outcroppings of the substratum in the form of steep mountains or hills made up of gneiss and granite, which rise in the broad valley areas.

The valley bottoms are in general wide, constituted by recent alluvium. In them, water courses flow between natural banks which cause the formation of lakes at the end of secondary valleys or at the bends of the primary.

These lakes and the recent alluvium which are relatively permeable constitute a considerable storage, because of which the flood waves are delayed and attenuated significantly.

In the depletion phase, these catchment areas provide a significant fraction of the discharges which are found again in the channel.

The high part of the basin is characterized by relatively steep slopes of the versant and dendritic drainage. The climate of the basin is characterized by a strong North-West/South-East rainfall gradient, caused in part by the existing elevation gradient in West-East direction, and corresponding for the other part to the North-South gradient common to sub-tropical and Sahel Africa. This gradient is due to the fact that the humid air masses originating on the Atlantic move in an East-North-East direction as far as the Inter-Tropical Continental Zone (ITCZ), which is shifted between about the Equator in the winter, and about 15° latitude in the summer.

It follows from this that the air masses become drier as they move northward and that the length of the rainy season is shorter. The rains have a stormy character, and an unpredictable small scale spatial distribution. Moreover, this character of the rains determines a minor importance of the orography as far as the orientation of slopes is concerned: the air masses rise and produce condensations and rain much more in daily thermal convection cells than in the form of frontal disturbances or masses of moist air forced upward by the orography encountered along its path.

Nevertheless, the peaks facilitate the formation of convection cells and with the increase of elevation relative humidity increases, thus it can be clearly observed that in the high zones the rainy season begins significantly earlier and ends significantly later than in the Benue valley.

The annual rainfall varies from 1,065mm at Gassol to 1,984mm at Mayo Ndaga. Since the effective evaporation is reasonably constant on the basin — about 900mm/year, the unitary discharges vary from 165mm to 1,080 mm/year. The southeastern ends of the basin seem to contribute in a determinative manner to the low discharge of water, because the dry season is briefer here.

3.1.2 GASSOL STATION

The Gassol station is situated on the right bank of the Taraba, in exact correspondence with the village of the same name, at the point where the villagers go down to the river.

The Station was installed by Nederland Development Company (NEDECO) in 1955 and in 1958 was passed on to Inland Waterways Division (IWD), which has operated it until the present (1991). It consists of five staff attached to driven metal tubes and by two crest gauge staff.

The range of the station is 0.7 m; at the inspection on 13-14 September, 1979 errors appeared in the levelling of the various staff, with a difference of from 4 to 9 cm.

The Station is located at km 1,112.30 km upstream from the present confluence with the Benue, near a limestone outcropping which extends into the channel. Downstream from the Station the spilled waters probably do not return to the Taraba channel but reach the Benue by crossing the swamps and the low lands which are characteristic of the zone. Further downstream, secondary branches are to be found. Therefore, the Station represents the natural closure section of the Taraba basin.

The river channel consists of one channel that runs near the right slope of the valley and is bordered by natural banks which separate it from the lowlands, lakes and swamps on both shores.

The valley at maximum high waters level is about 4 km wide, while the channel has a mean width of 400 m. The wide low strip beyond the left bank of the Taraba is drained by the Wurio river, which flows for over 50 km in sinuous meanderings, among oxbow lakes — the remains of old meanders — separated from the Taraba by the natural bank.

The channel bed consists of coarse granitic sand, size about 2-3 mm; with frequent nodules having a dimension of about 1 cm. The banks are stabilised by luxuriant vegetation and silt.

The local slope of the channel is 40 cm/km and tends to grow slightly as the channel proceeds downstream.

The gauge zero drawn from the 1979 ALTS levelling along the entire axis of the bund of the reclamation area, is at an elevation of 106.95 above Federal Survey of Nigeria (FSN) datum.

The basin has a surface of 22,354 km², half of which falls within the Beli station range.

3.2 DATA

3.2.1 Data Verification

The Flood discharge record obtained at Gassol gaging station was verified for consistency using double mass curve analysis. The cumulative discharges at Gassol gaging station was plotted against the cumulative mean discharges of five stations. The five stations used were, River Katsina Ala at Katsina Ala, River Katsina Ala at Sevav, River Donga at Donga, River Taraba at Beli and River Donga at Manyu.

The double mass curve analysis was carried out with the aid of Table 3.0 and the Figure 3.2. From Figure 3.2, it can be seen that the Gassol gaging station discharge records are consistent.

Table: 3.0 Double Mass Curve Analysis

Year	Taraba at Gassol	K/Ala at Sevay	Donga at Donga	Taraba at Beli	Donga at Manye	Taraba at Gassol Cumulative discharges	Cumulative mean discharges of 5 stations
1955	1930	2670	1776	2322	(1050)	1,930	2,084
1956	2038	2826	2047	2382	(1130)	3,968	4,503
1957	2000	2610	1800	2180	(1080)	5,968	6,617
1958	1710	1910	1632	1612	(920)	7678	8,327
1959	1120	1635	1250	1100	602	8,798	9,608
1960	1885	2428	1890	2426	735	10,683	11,679
1961	1725	3658	1645	2125	726	12,408	13,868
1962	1635	1809	1548	1642	982	14,043	15,504
1963	2280	3000	2272	2830	970	16,323	17,934
1964	1680	2050	1580	2220	858	18,003	19,766
1965	1880	2095	1730	2467	1050	19,883	21,744
1966	1785	2329	1850	(1940)	1030	21,668	23,697
1967	1720	2145	(1560)	1860	1000	23,388	25,572
1968	2195	(2900)	1750	2550	947	25,583	27,851
1969	2400	3164	2150	2700	1040	27,983	30,302

Table: 3.0 Contn'd

Year	Taraba at Gassol	K/Ala at Sevav	Donga at Donga	Taraba at Beli	Donga at Manyu	Taraba at Gassol Cumulative discharges	Cumulative mean discharges of 5 stations
1970	2017	2598	1680	2000	835	30,000	32,173
1971	1188	1820	1350	1372	840	31,188	33,649
1972	1675	2007	1510	2100	864	32,863	35,281
1973	1700	2480	(1720)	2150	850	34,563	37,172
1974	1730	2122	1910	2010	837	36,293	39,052
1975	1410	1700	1475	1798	846	37,703	40,628
1976	1605	1920	1690	1789	945	39,308	42,357
1977	1670	1950	1700	1737	584	40,978	44,035
1978	1650	2246	2010	1500	852	42,628	45,829
1979	1690	2246	1670	1957	965	44,318	47,737
1980	2047	(2650)	2100	2350	940	46,365	50,426
1981	2500	3450	2400	1767	1400	48,865	53,019
1982	1865	2465	1740	1897	1550	50,730	55,066
1983	1460	2324	1600	1415	1020	52,190	56,768
1984	1552	2241	1490	1398	1100	53,742	58,484

Table: 3.0 Contin'd

Year	Taraba at Gassol	K/Ala at Sevay	Donga at Donga	Taraba at Beli	Donga at Manyu	Taraba at Gassol Cumulative discharges	Cumulative mean discharges of 5 - stations
1985	1685	2200	1650	2095	(900)	55,427	60,478
1986	1890	(2450)	1930	2200	(1010)	57,317	62,596
1987	1575	2150	1615	(1650)	(840)	58,892	64,327
1988	1275	1550	1420	(1340)	(660)	60,167	65,665

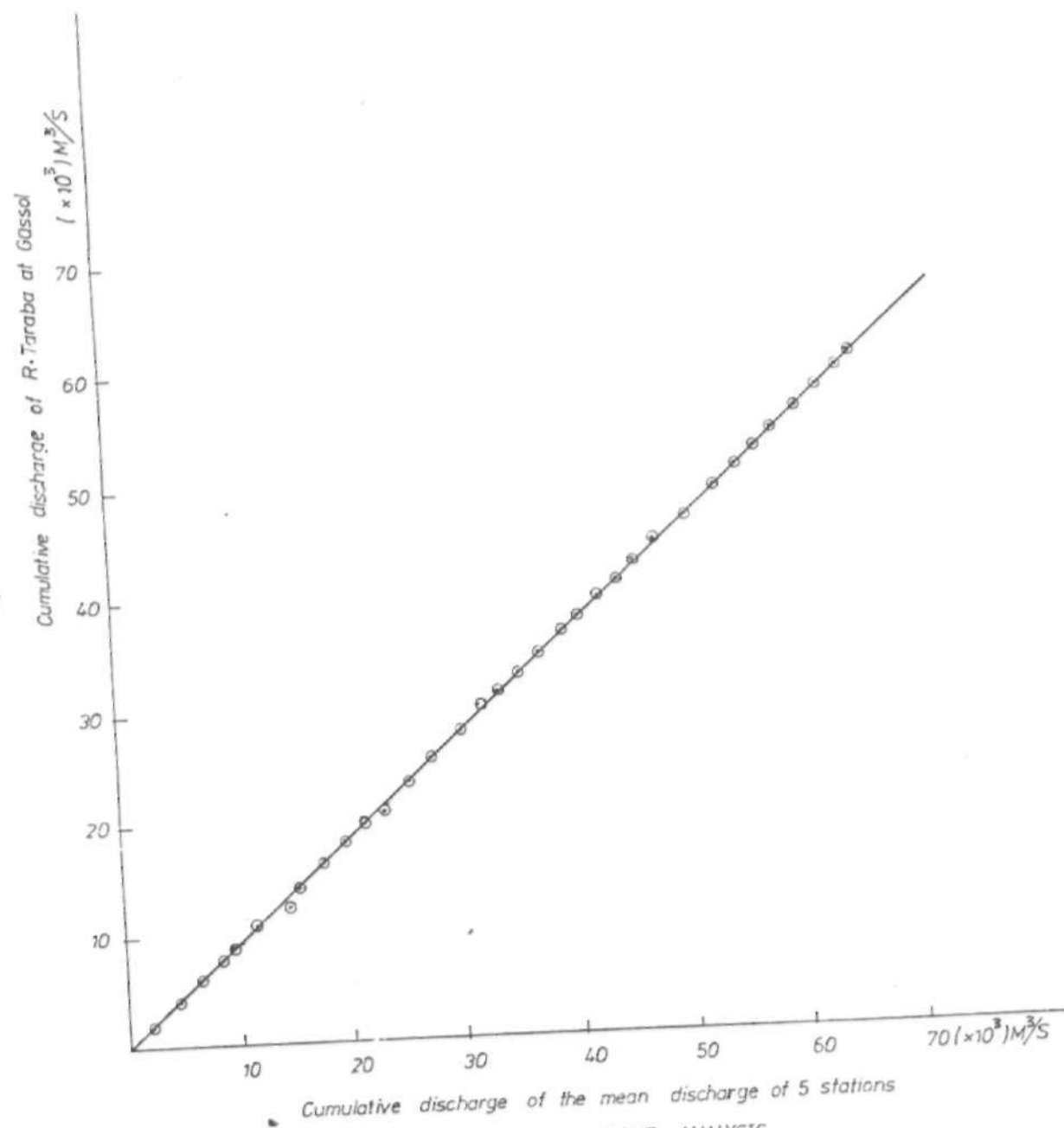


FIG. 3.2 DOUBLE MASS CURVE ANALYSIS

3.2.2 Data Analysis

The data shown in Appendix A-1 through A-12 were collected from twelve selected gaging stations in Upper Benue River Basin by Ogbuli. The peak flood discharges are shown in column (2) of each table. Those discharges shown in bracket in column (2) were generated by Ogbuli, using random generation technique. In column (3) the discharges have been rearranged in descending order of magnitude with the generated data omitted. The generated data only assist in estimating the order number, M (column 4) of the collected discharges of each station. Thus, in the frequency analysis of all the stations, the generated discharges have been omitted.

In column (5) the recurrence intervals have been calculated using Weibull's formula:

$$T_r = \frac{N+1}{m} \dots\dots\dots(3.2.1)$$

where T_r is the recurrence interval, N is the length of record in years and m is the order number.

The exceedance probability shown in column (6) have been calculated as the reciprocal of the recurrence interval expressed in percentages.

3.3 FITTING OF FREQUENCY CURVES

Several probability distributions were fitted to the data shown in Appendices A - 1 through A - 12 by the author of this thesis. It was found that only the log-normal distribution was the most suitable for all the twelve stations considered. These frequency curves are shown in Figures 3.3 through 3.14.

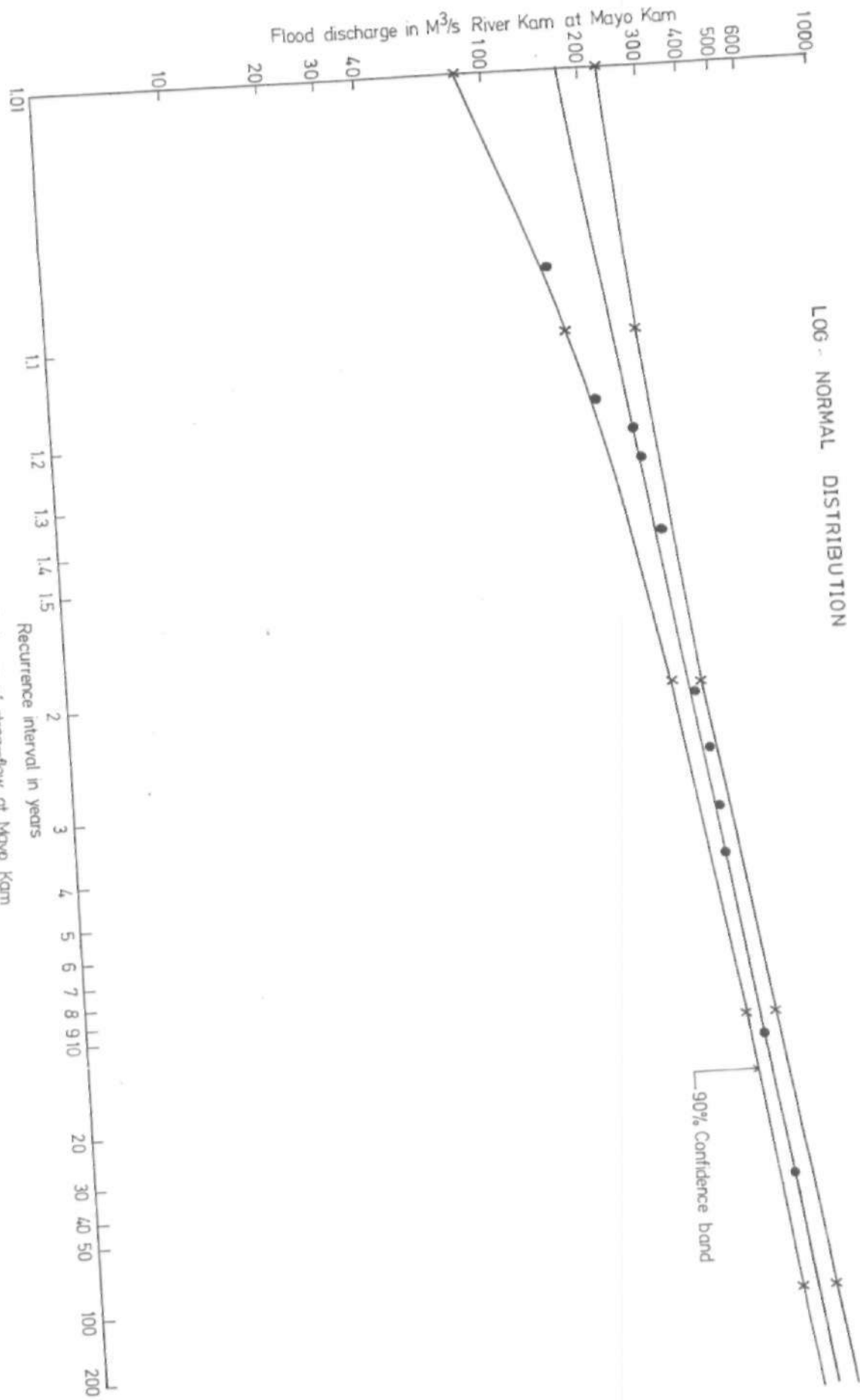


Fig 3.3 Frequency distribution for discharge of streamflow at Mayo Kam

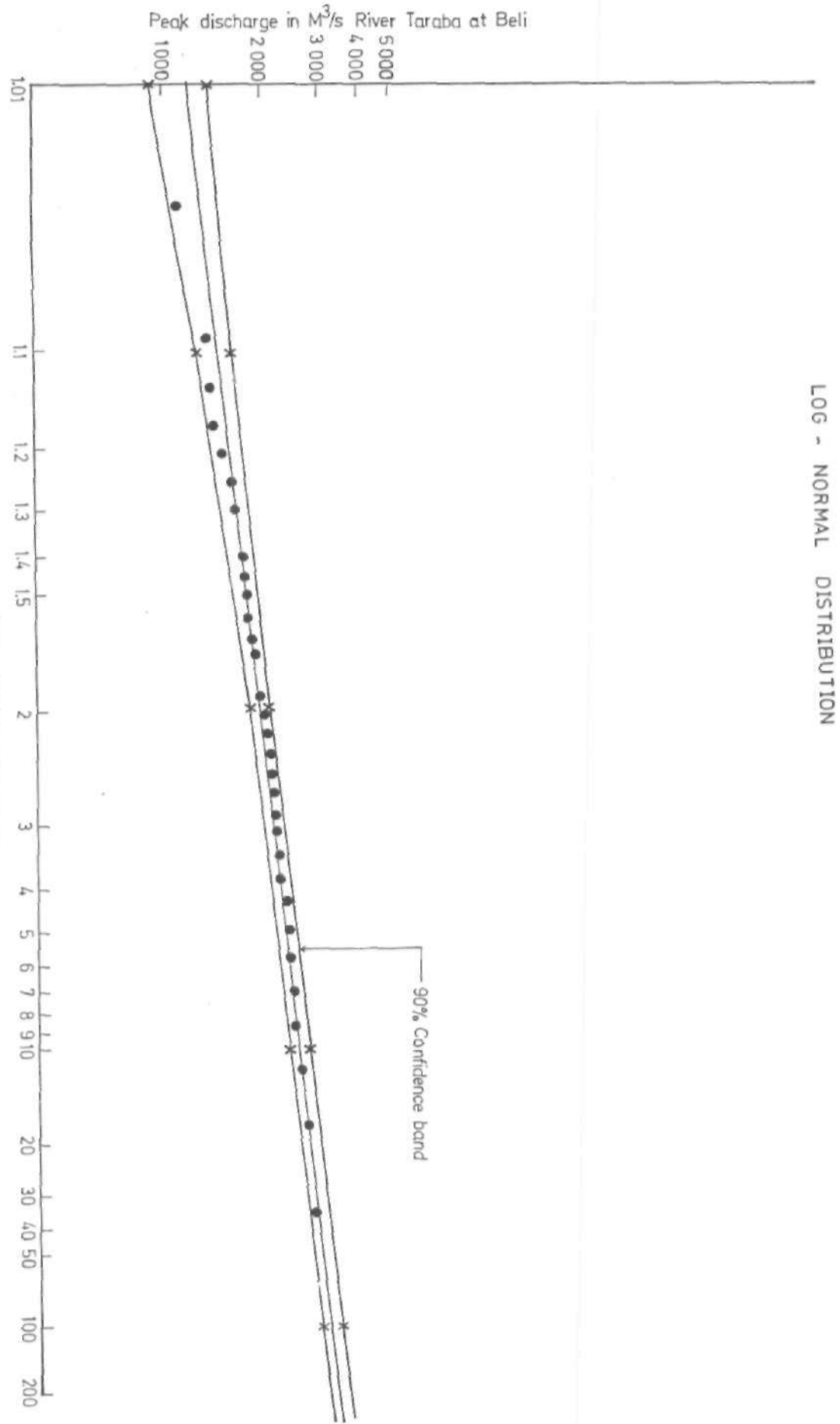


Fig. 3.4. Frequency distribution for discharge of streamflow at Beli

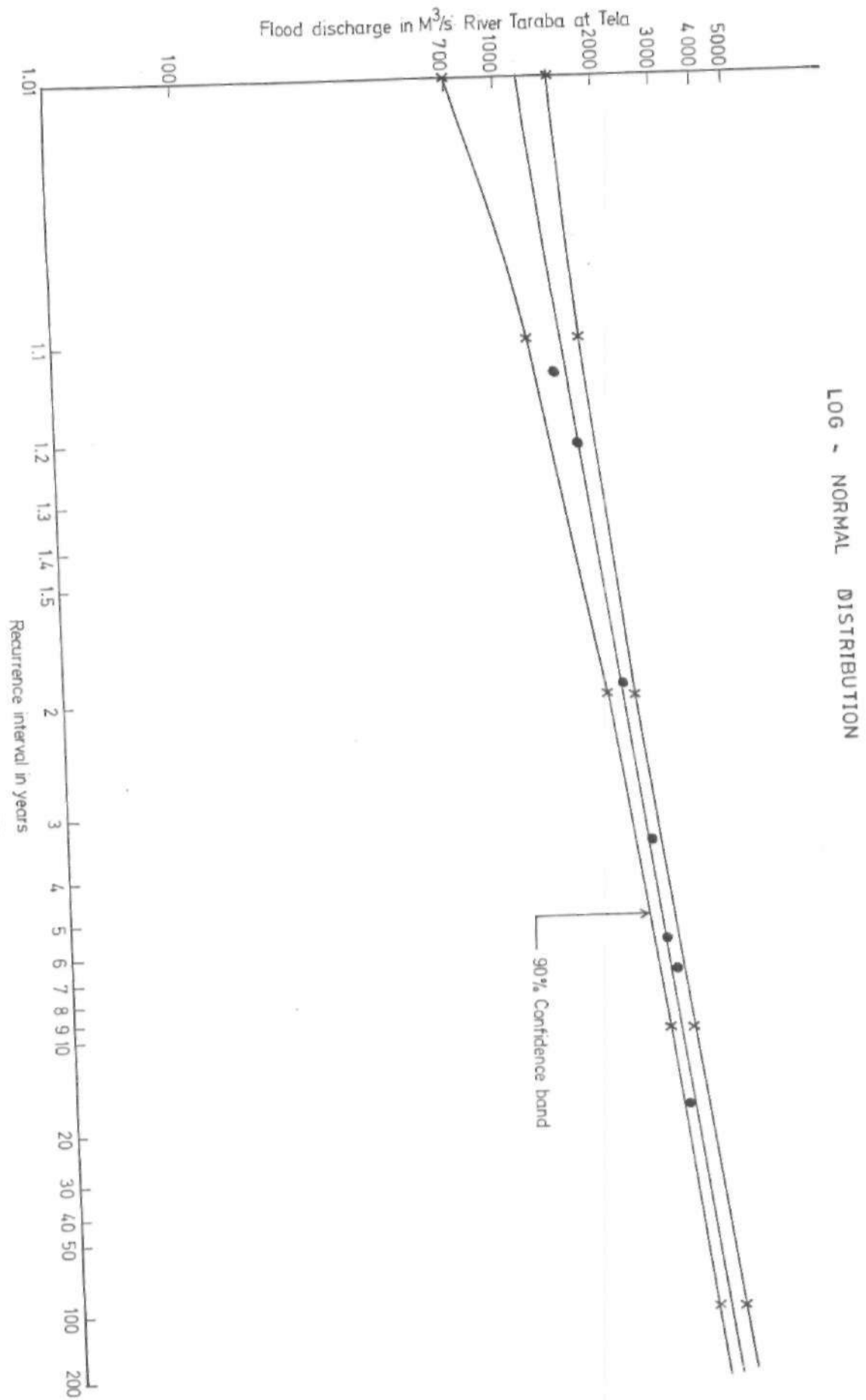


Fig.3.5. Frequency distribution for discharge of streamflow at Tela

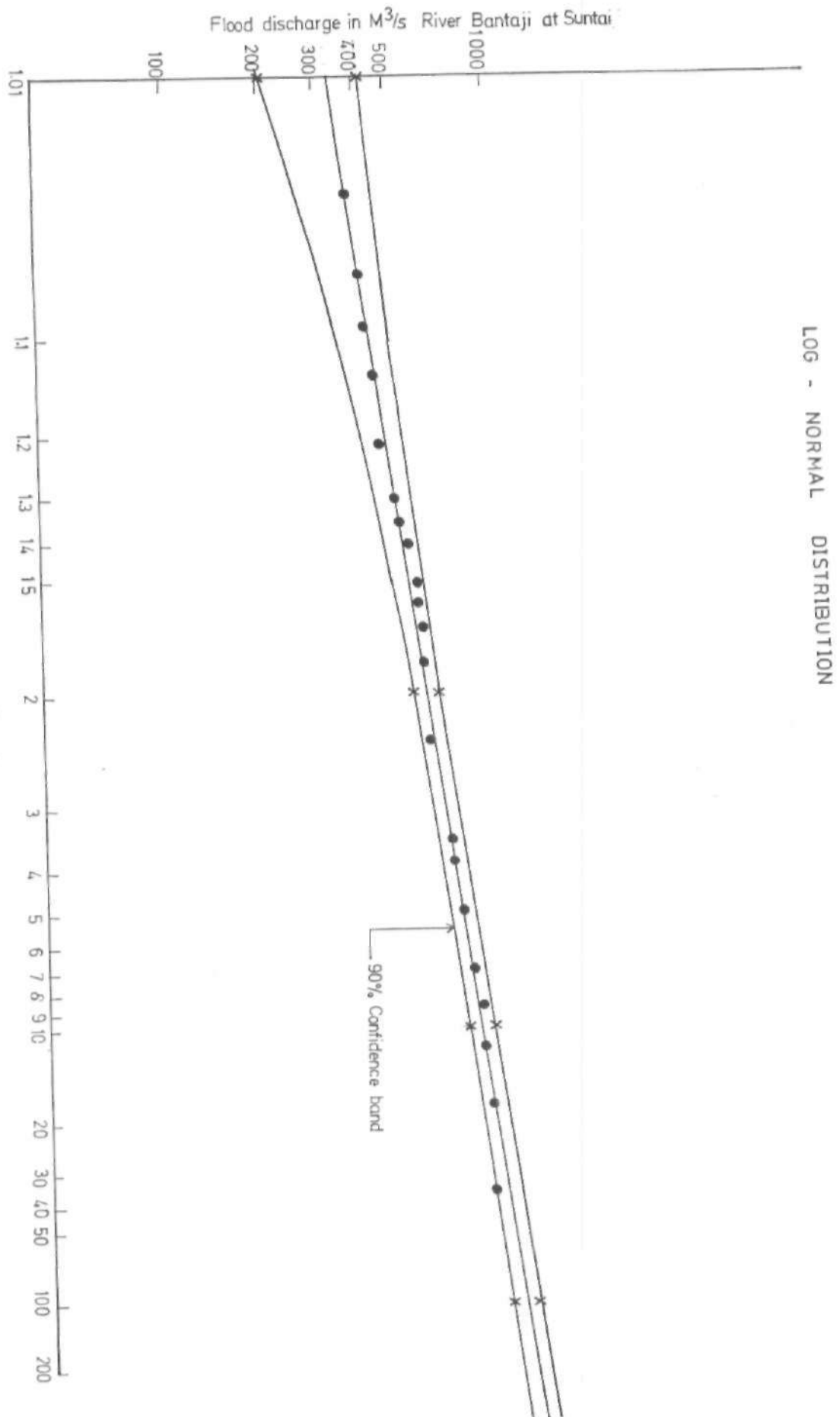


Fig. 3.6. Frequency distribution for discharge of streamflow at Suntai

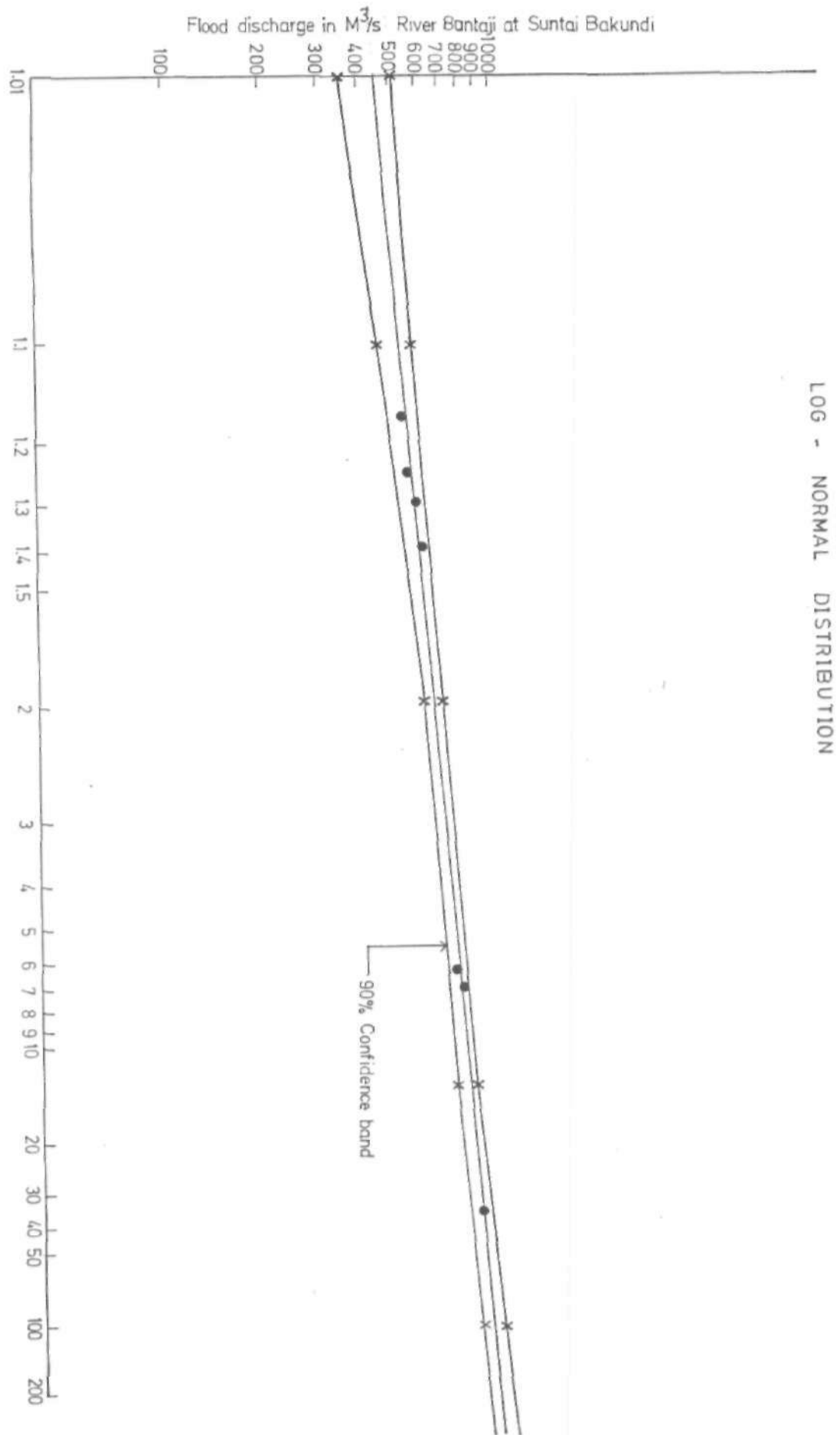


Fig. 3.7 Frequency distribution for discharge of streamflow at Suntai Bakundi

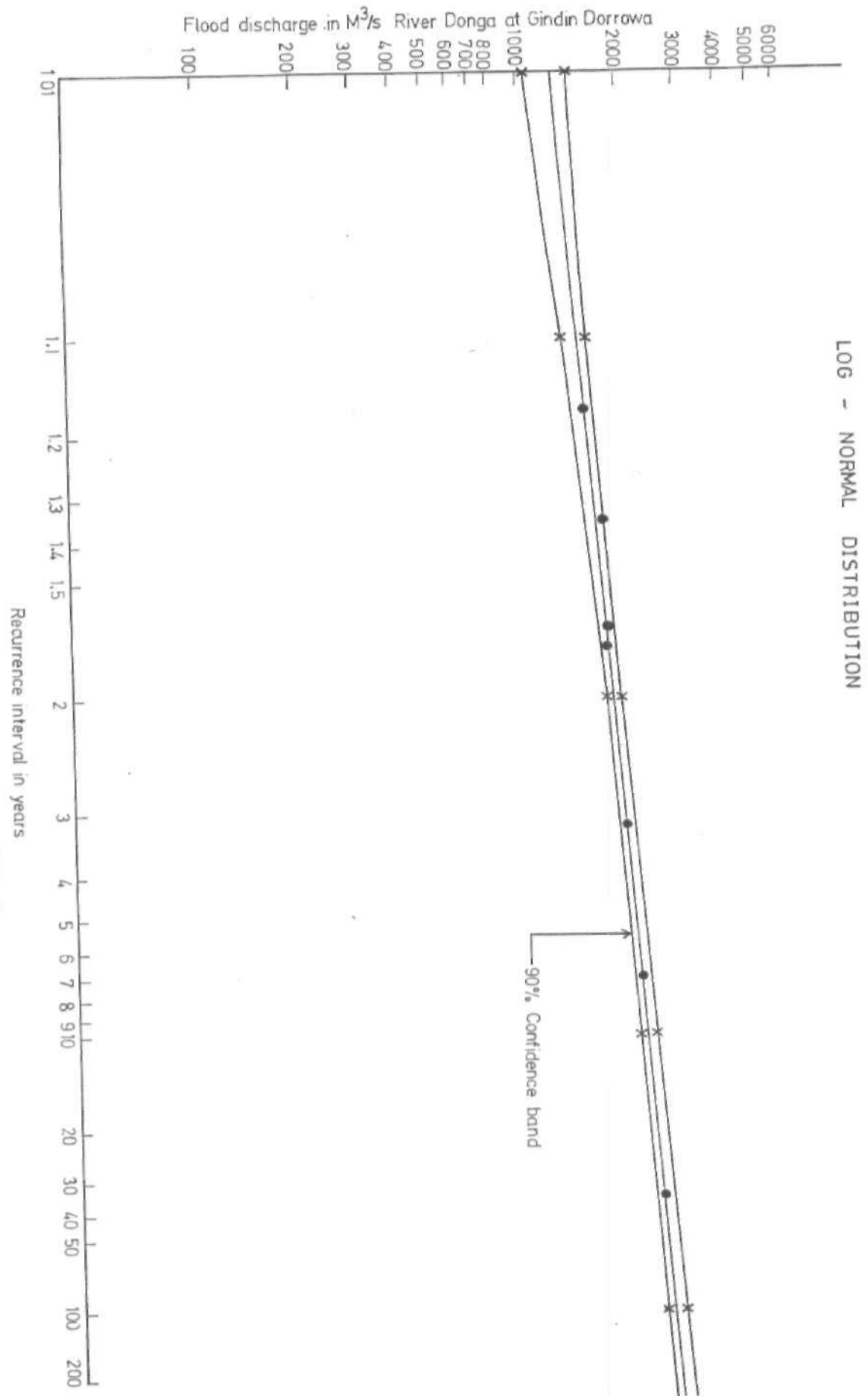


Fig.3.8 Frequency distribution for discharge of streamflow at Gindin Dorrowa

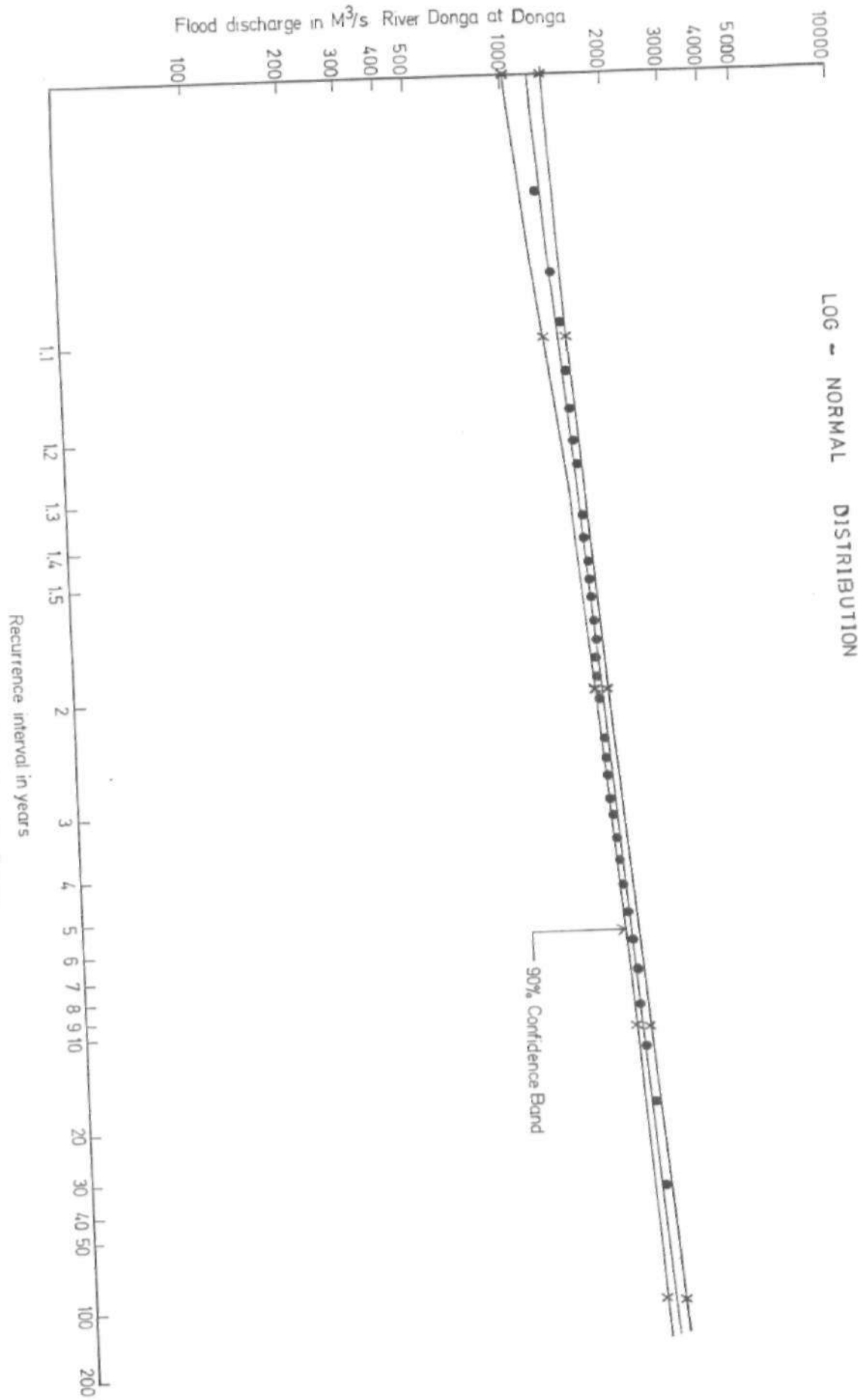


Fig. 3.9 Frequency distribution for discharge of streamflow at Donga

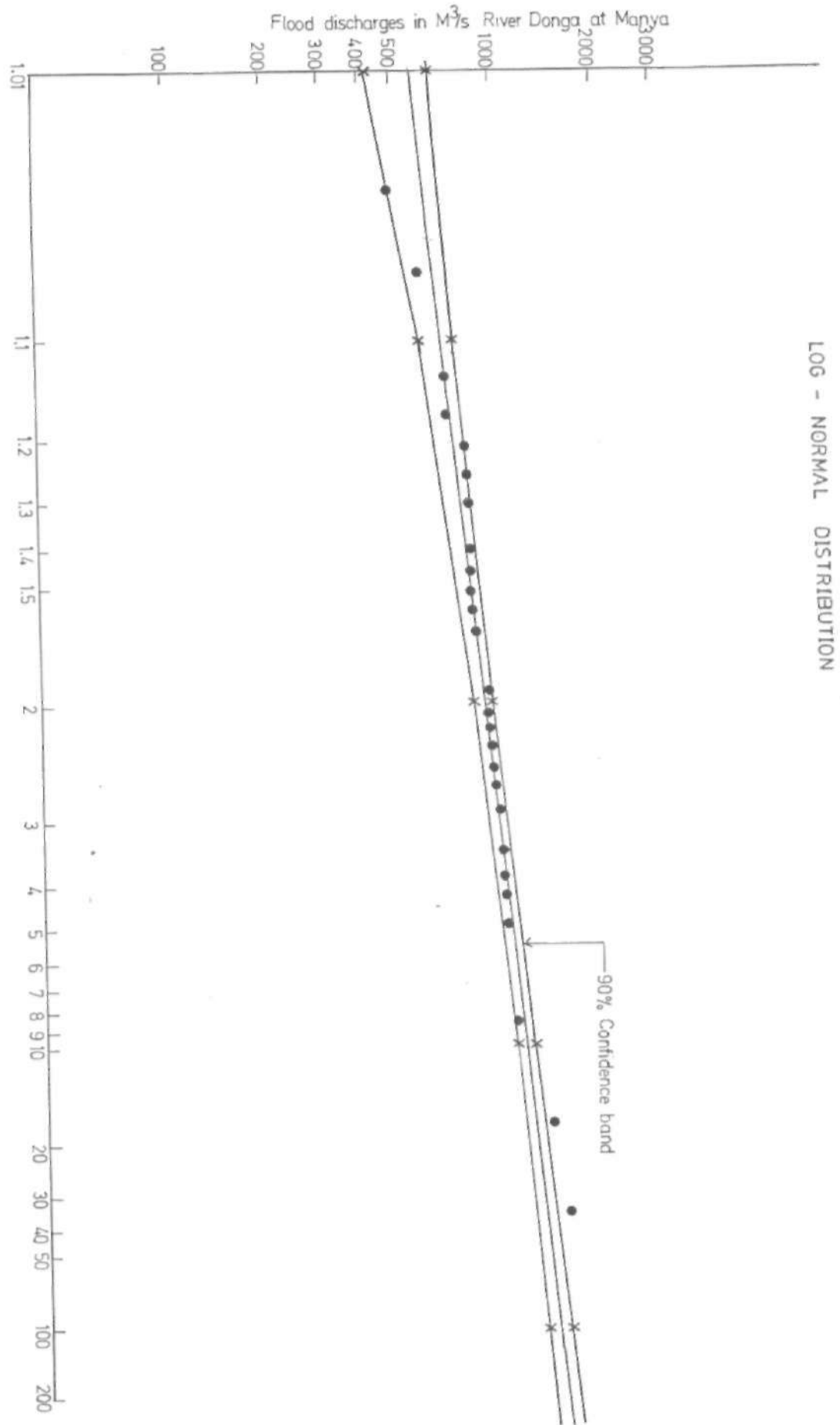


Fig. 3.10 Frequency distribution for discharge of streamflow at Manya

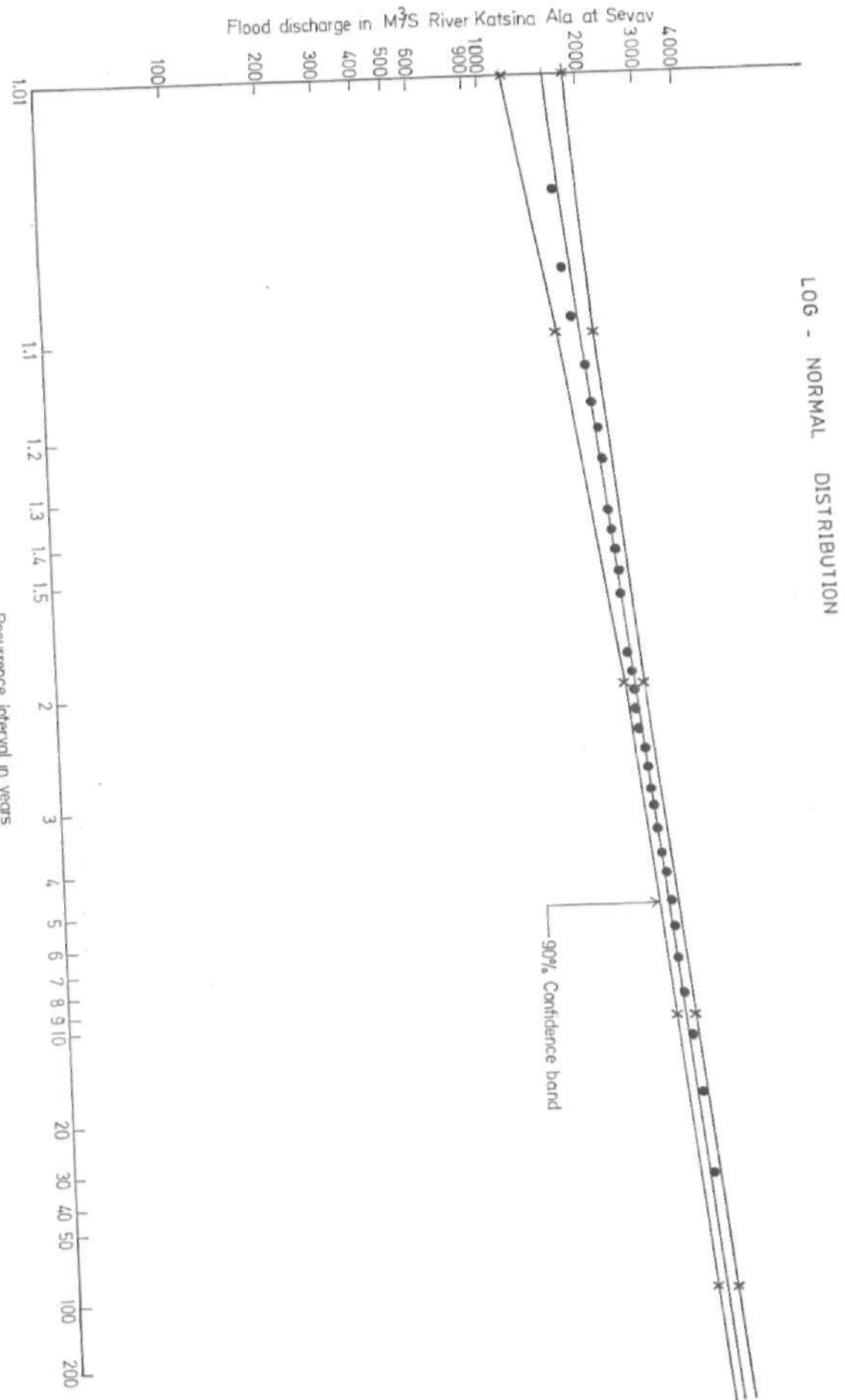


Fig. 3.12 Frequency distribution for discharge of streamflow at Seva

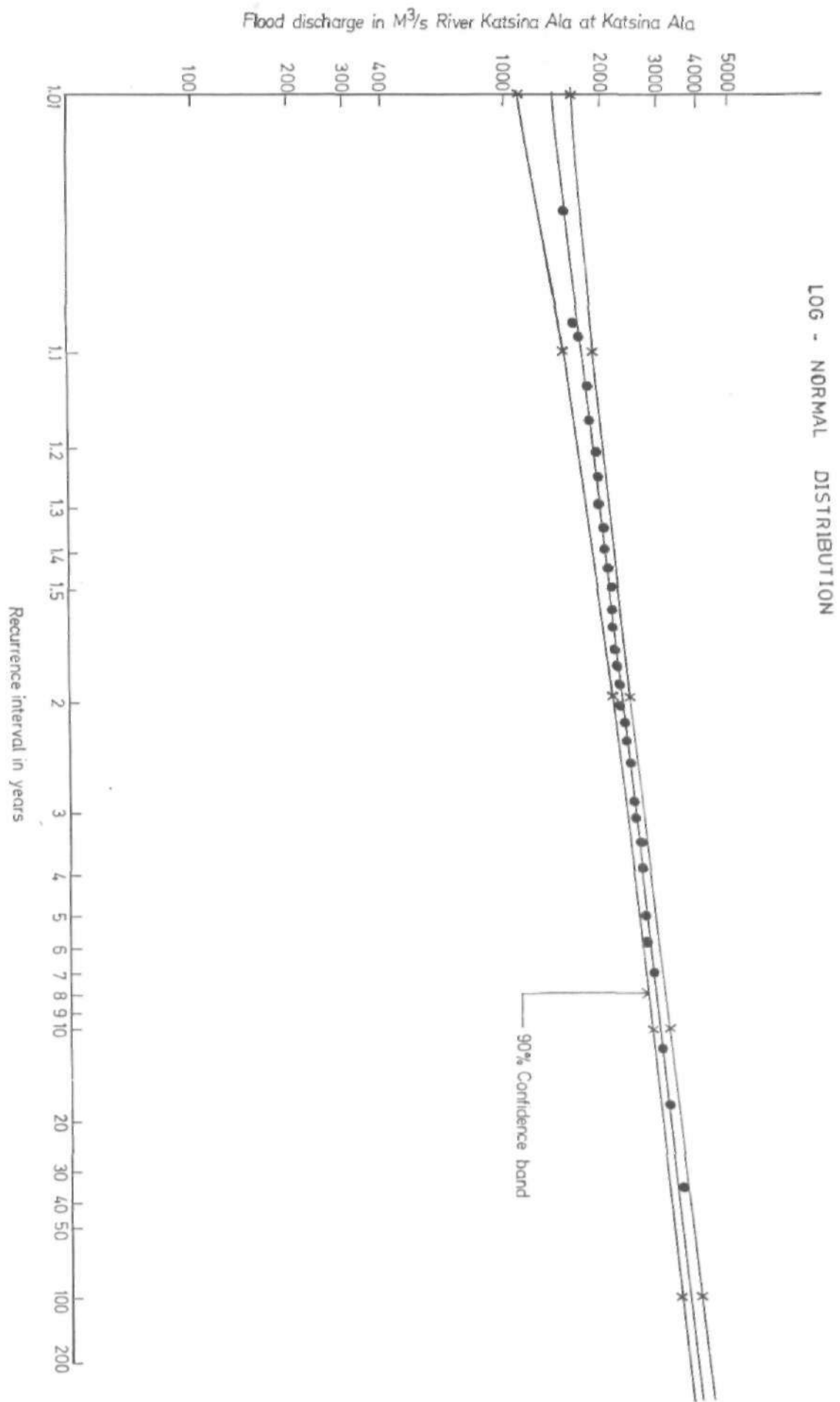


Fig. 3.13 Frequency distribution for discharge of streamflow at Katsina Ala

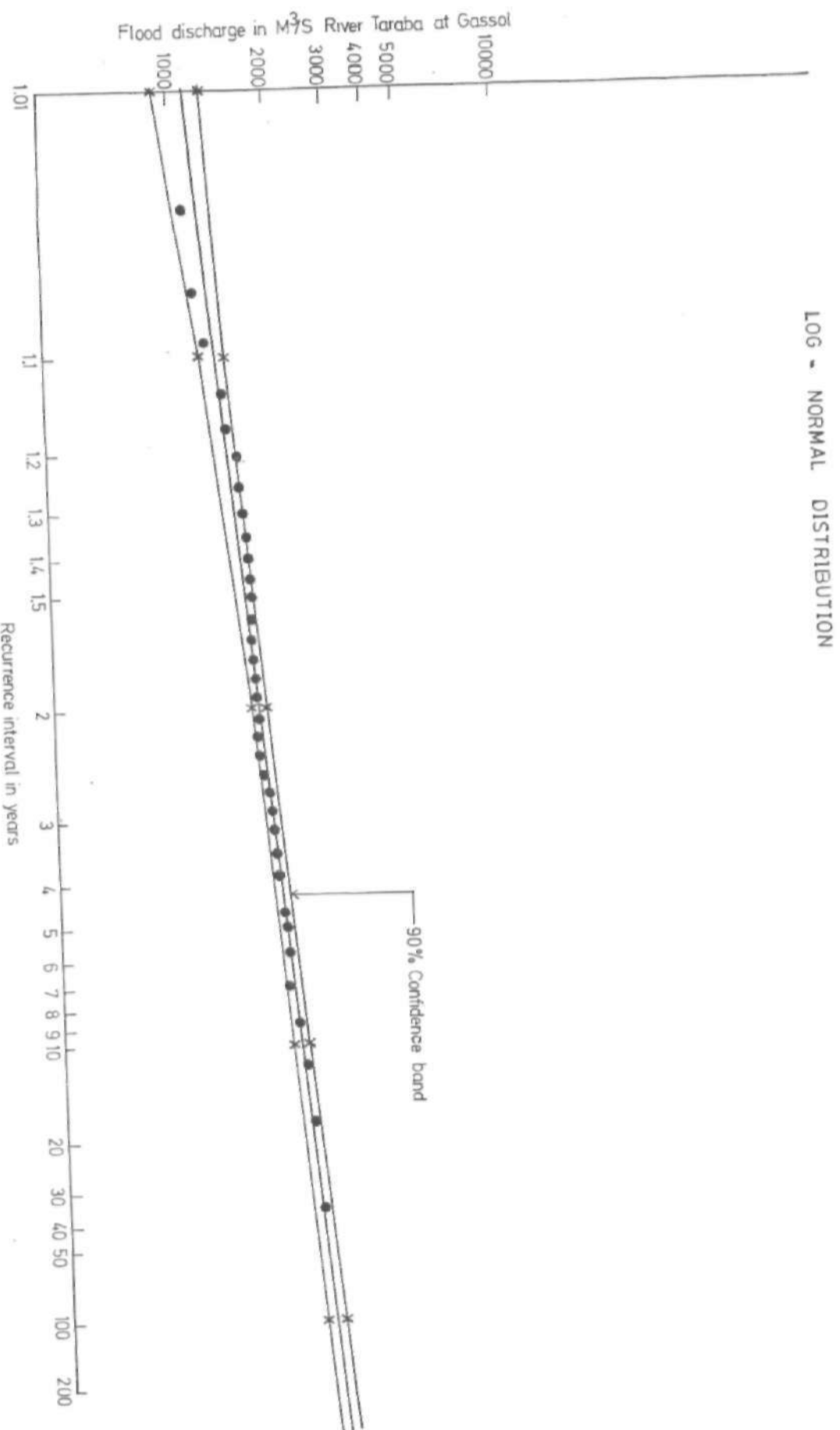


Fig 3.14 Frequency distribution for discharge of streamflow at Gassol

The confidence band have been plotted with the aid of Beard's coefficient. 95% confidence band were initially tried and it was found unsuitable for most of the stations. Therefore 90% confidence band was used for all the twelve gaging stations and found to be satisfactory.

The upper confidence bands were plotted by adding the product of the Beard's coefficient and the standard deviation of the gaging station discharge record to the theoretical discharge at the station for a particular return period.

The lower confidence bands were plotted by subtraction of the product of the Beard's coefficient and the standard deviation of the gaging station discharge record from the theoretical discharge at the station for a particular return period.

The Beard's coefficients for the various return periods for the different number of years of records are presented in Table 3.1.

Table 3.1: Coefficients for the calculation of the 90% confidence intervals around the flood frequency curve Beard (1962)

Confidence Band	Year of Record	Recurrence Interval, T_r (Years)					
		1000	100	10	2	1.1	1.0
90% Upper	5	4.41	3.41	2.12	0.95	0.76	1.00
	10	2.11	1.65	1.07	0.58	0.57	0.76
	15	1.52	1.19	0.79	0.46	0.48	0.65
	20	1.23	0.97	0.64	0.39	0.42	0.58
	30	0.93	0.74	0.50	0.31	0.35	0.49
	40	0.77	0.61	0.42	0.27	0.31	0.43
	50	0.67	0.54	0.36	0.24	0.28	0.39
	70	0.55	0.44	0.30	0.20	0.24	0.34
	100	0.45	0.36	0.25	0.17	0.21	0.29
90% Lower	5	-1.22	-1.00	-0.76	-1.95	-2.12	-3.41
	10	-0.94	-0.76	-0.57	-0.58	-1.07	-1.65
	15	-0.80	-0.65	-0.48	-0.46	-0.76	-1.19
	20	-0.71	-0.58	-0.42	-0.39	-0.64	-0.97
	30	-0.6-	-0.49	-0.35	-0.31	-0.50	-0.74
	40	-0.53	-0.43	-0.31	-0.27	-0.42	-0.61
	50	-0.49	-0.39	-0.28	-0.24	-0.36	-0.54
	70	-0.42	-0.34	-0.24	-0.20	-0.30	-0.44
	100	-0.37	-0.29	-0.21	-0.17	-0.25	-0.36

3.4

HOMOGENEITY TEST

In order to check for the regional hydrologic homogeneity of the flood discharge records of the twelve gauging stations from Upper Benue River Basin used for the analysis, the procedure described by Dalrymple (1960) was followed.

According to Dalrymple (1960), (Odusote, 1989), the confidence limits, assuming 95% of the estimate lie within $\pm 2 \sigma_x$ of the most probable value, is given by:

$$C. L. = 2.25 \pm 6.33/\sqrt{n} \quad \dots\dots\dots(3.4.1)$$

for a ten year flood.

n is the sample size, and σ_x is the standard deviation.

Table 3.2 below after Dalrymple gives the upper and lower limits with the corresponding return periods for various values of n .

Table 3.2 Confidence Limits of Regional Flood
Homogeneity Test Dalrymple (1960).

Sample Size	Lower	Limit	Upper	Limit
n	$X - 2 \sigma_x$	T_L	$X + 2 \sigma_x$	T_u
5	- 0.59	1.20	5.09	160
10	0.25	1.85	4.25	70
20	0.83	2.60	3.67	40
50	1.35	4.40	3.15	24
100	1.62	5.60	2.88	18
200	1.80	6.50	2.70	15
500	1.97	7.70	2.53	13
1000	2.05	8.30	2.45	12

The procedure used for the test is as follows:

- (1) The values of T_L and T_U obtained from Table 3.2 were plotted against n on a semi-log scale paper to give the lower and upper regional confidence limits respectively as shown in Figure 3.15
- (2) The mean annual flood discharge, $\bar{Q}$, at each gauging station was calculated and are tabulated in column 5 of Table 3.3
- (3) The 10-year flood, Q_{10} , at each gauging station was estimated by interpolation from the frequency curve of each station to give the values in column 6 of Table 3.3.
- (4) The ratio $Q_{10}/\bar{Q}$ for each gauging station was calculated and the mean value of $Q_{10}/\bar{Q} = C$ computed. (see column 7 Table 3.3).
- (5) The mean annual flood discharge value, $\bar{Q}$, at each station was multiplied by the constant C to give values shown in column 8, Table 3.3.
- (6) By interpolation from the frequency curve of each station, the return period, T_E , corresponding to the modified 10-year flood discharges was estimated to give values in column 9, Table 3.3.
- (7) The effective period, N_E , of each of the gauging station was determined as the number of recorded annual events plus one half the number of events computed for that station by inter-correlation (Odusote, 1989). These values are shown in column 10, Table 3.3.
- (8) The co-ordinates (N_E, T_E) for each station was plotted on the same semi-log paper used in step (1) above. The plotted points are shown in Figure 3.15 for all the twelve stations.

From Figure 3.15, it could be seen that all the twelve stations. points lie within the regional confidence limits; showing that all the flood discharge records are from the same homogeneous basin.

Table 3.3: DATA FOR HOMOGENEITY TEST

1	2	3	4	5	6	7	8	9	10
Station no.	Stream (Station)	Drainage Area (km ²)	Length of Main Channel (km)	Mean Annual Flood, $\bar{Q}$ (m ³ /s)	10-year Flood, Q_{10} (m ³ /s)	Ratio $Q_{10}/\bar{Q}$	$\bar{Q} \times 1.2605$ (m ³ /s)	T_E for Q of column 8 (years)	Period of record, N_E adjusted (Years)
1.	Mayo Kam (at Mayo Kam)	2,986	120	345	405	1.406	435.7	5.3	22.5
2.	Paraba (at Bali)	10,850	192	1,998.4	2520	1.261	2524	10.1	32.5
3.	Paraba (at Pela)	19,979	264	2,433.7	3060	1.263	3061	9.8	20.5
4.	Suntai (Bantaji)	5,815	185	652	890	1.365	823	6.3	27.5
5.	Suntai (Suntai Bakundi)	3,350	106	684.3	812	1.187	864.3	18.5	20.5
6.	Donga (Gindin Dorowa)	20,390	305	1,937.1	2280	1.177	2446.6	24.0	20.5
7.	Donga (Donga)	11,909	257	1,745.6	2120	1.214	2204.7	16.2	33
8.	Donga (Many)	9,040	196	937.2	1170	1.248	1183.7	11.0	30
9.	Donga (Abong)	6,424	121	2,416.8	2950	1.221	3052.4	16	25
10.	Katsina Ala (Sevav)	22,000	291	2,619.4	3400	1.298	3308.3	8.1	32.5
11.	Katsina Ala (Katsina Ala)	16,700	227	2,316.1	2950	1.274	2925.2	9.5	32.5
12.	Paraba (at Gassol)	20,513	291	1769.6	2200	1.243	2235	12	34

Mean, C
= 1.263

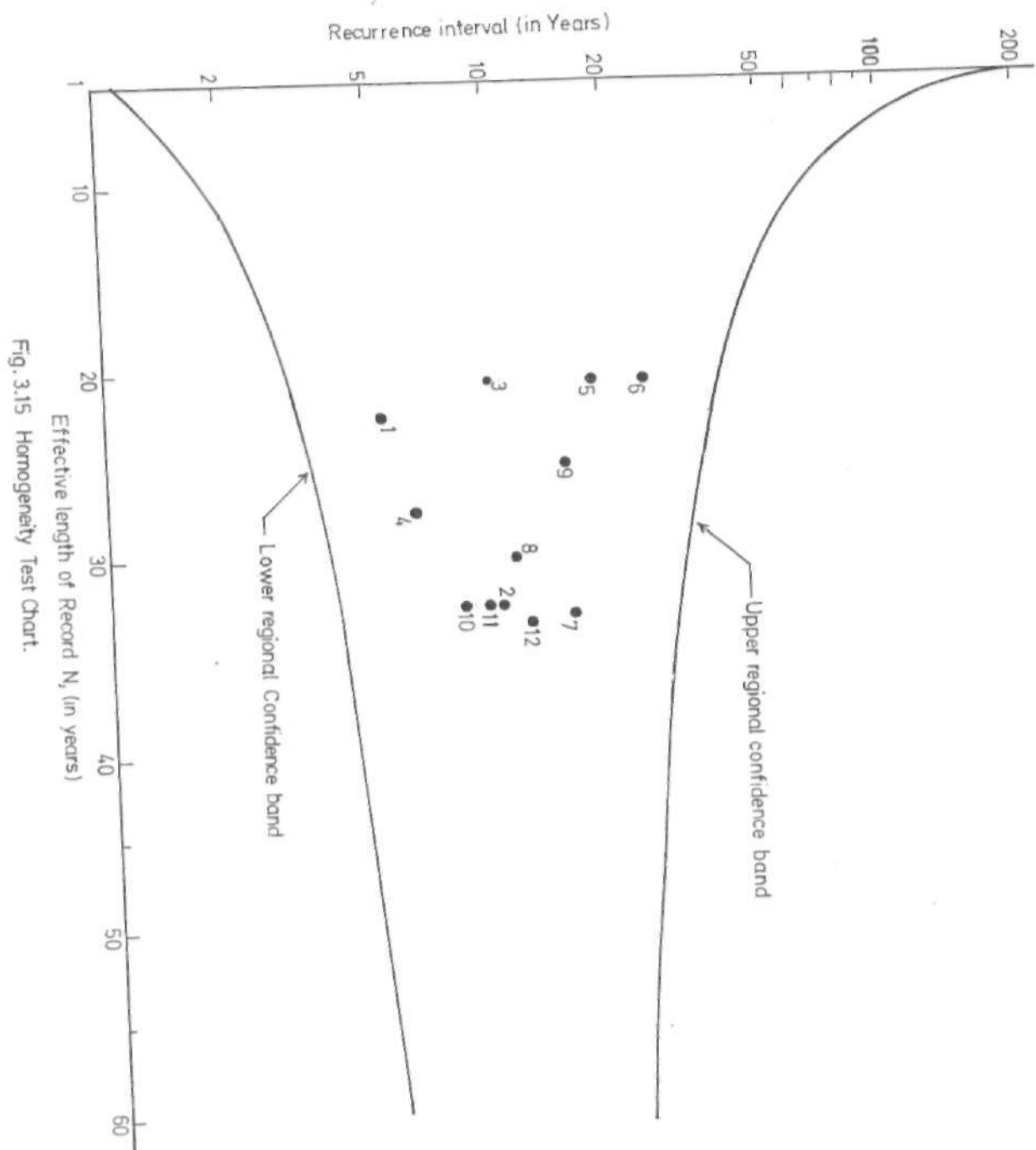


Fig. 3.15 Homogeneity Test Chart.

3.5

REGRESSION ANALYSIS

3.5.1

Review of the Previous Work on Regional Analysis

The analytical expressions for the prior parameters of models derived in chapter 2 involve expectations and variances of the variables (discharges). These expectations and variances can be estimated using regional frequency analysis.

Odusote (1989) performed regional frequency analysis on Ogun-Oshun - River Basin in Nigeria, using regression analysis and he arrived at equation (1.1) given by:

$$Q_{mi} = a A_i^b L_i^c$$

relating mean Maximum annual flood discharge, Q_m , to the physiographic information - drainage area, A , and the main stream length, L .

In equation (1.1) i is the station number while a , b , c are regression constants to be determined.

Ogbuli (in print) also carried out the regional frequency analysis on the Upper Benue River Basin in Nigeria using regression analysis and arrived at an equation of the same form as equation (1.1). For the above reasons, the author of this thesis have performed regression analysis for the mean and variance of the flood discharges assuming the same form of equation as equation (1.1) throughout this work.

3.5.2

General Procedure:

The general method of the regression analysis is formulated as follows:

Taking the logarithm of both sides of equation (1.1)

gives:

$$\log Q_{mi} = \log a + b \log A_i + c \log L_i \dots\dots\dots(3.5.1)$$

There is an error, E_i , between the actual value of Q_{mi} and the estimated value of Q_{mi} from equation (3.5.1). Taking the error into consideration equation (3.5.1) can be written as:

$$\log Q_{mi} = \log a + b \log A_i + c \log L_i + E_i \quad \dots\dots\dots (3.5.2)$$

Equation (3.5.2) implies that

$$E_i^2 = (\log Q_{mi} - \log a - b \log A_i - c \log L_i)^2 \quad \dots\dots\dots (3.5.3)$$

If we substitute P to be the sum of the square of the errors,

$$\begin{aligned} \text{then} \\ P &= \sum_{i=1}^n E_i^2 = \sum_{i=1}^n (\log Q_{mi} - \log a - b \log A_i - c \log L_i)^2 \dots\dots (3.5.4) \end{aligned}$$

In order to minimise the square of the errors, the partial derivatives

$$\frac{\partial P}{\partial (\log a)}, \frac{\partial P}{\partial b}, \frac{\partial P}{\partial c},$$

must each be zero. This implies that:

$$\begin{aligned} \frac{\partial P}{\partial (\log a)} &= -2 \sum_{i=1}^n (\log Q_{mi} - \log a - b \log A_i - c \log L_i) = 0 \\ \sum_{i=1}^n (\log Q_{mi}) &= n \log a + b \sum_{i=1}^n \log A_i + c \sum_{i=1}^n \log L_i = 0 \dots (3.5.5) \end{aligned}$$

Similarly

$$\begin{aligned} \frac{\partial P}{\partial b} &= -2 \log A_i \cdot \sum_{i=1}^n (\log Q_{mi} - \log a - b \log A_i - c \log L_i) = 0 \\ \sum_{i=1}^n (\log Q_{mi}) \cdot (\log A_i) &= \log a \cdot \sum_{i=1}^n \log A_i + b \sum_{i=1}^n (\log A_i)^2 + \\ &\quad c \sum_{i=1}^n (\log L_i)(\log A_i) \dots (3.5.6) \end{aligned}$$

Similarly

$$\begin{aligned} \frac{\partial P}{\partial c} &= -2 \log L_i \cdot \sum_{i=1}^n (\log Q_{mi} - \log a - b \log A_i - c \log L_i) = 0 \\ \sum_{i=1}^n (\log Q_{mi}) \cdot (\log L_i) &= \log a \cdot \sum_{i=1}^n \log L_i + b \sum_{i=1}^n (\log A_i)(\log L_i) \\ &\quad + c \sum_{i=1}^n (\log L_i)^2 \dots\dots (3.5.7) \end{aligned}$$

The equations (3.5.5), (3.5.6) and (3.5.7) are called the normal equations. Since there are three unknown constants (a, b, c) to be determined in equation (1.1), the three normal equations are solved simultaneously using method of successive elimination to obtain the constants a, b, c.

3.6 APPLICATIONS OF REGRESSION ANALYSIS

3.6.1 General Procedure

The flood record at Gassol gaging station has been used by the author of this thesis as the historical record (being the longest and continuous record) while the eleven other gaging stations records from Upper Benue River Basin were used for the regression analysis.

Regression analysis was performed for both the mean annual maximum flood discharges and the variance for all the three selected distributions (normal, log-normal and exceedance distributions).

In the case of exceedance distribution, it was found that only six out of the eleven gaging stations have records with flood discharges greater than the chosen base level discharge of $2,200\text{m}^3/\text{s}$.

Hence regression analysis was based on these six gaging stations records. The regression analysis for the three distributions are presented as follows:

3.6.2 Normal Model

(i) Regression Model for Mean Annual Maximum Flood Discharges

The mean of the annual maximum flood discharges have been calculated by the author of this thesis from Appendix (A-1) through (A-11), using equation (2.3.6) and are as shown in

Table (3.4) together with the corresponding drainage area, A and main stream length, L .

Table (3.4): The Mean annual Maximum flood discharge at each gaging station in Upper Benue River Basin.

River	Station	Station Number i	Mean Annual Maximum Flood discharge $Q(m^3/s)$	Drainage Area $A(km^2)$	Main Stream Length $L(KM)$
Mayo Kam	Mayo Kam	1	345	2,986	120
Taraba	Beli	2	1,998.4	10,850	192
Taraba	Tela	3	2,423.7	19,979	264
Bantaji	Suntai	4	652	5,815	185
Bantaji	Suntai Bakundi	5	684.3	3,350	106
Donga	Gindin Dorrowa	6	1,937.1	20,390	305
Donga	Donga	7	1,745.6	11,909	257
Donga	Manya	8	937.2	9,040	196
Donga	Abong	9	2,416.8	6,424	121
Katsina Ala	Sevav	10	2,619.4	22,000	291
Katsina Ala	Katsina Ala	11	2,316.1	16,700	227

In order to be able to use the normal equations (3.5.5) through (3.5.7), the logarithm to base ten of the mean annual flood discharges Q_{mi} , the drainage area, A_i , and the main stream length L_i were taken and are shown in Table (3.5).

Table (3.5): The logarithm to base ten of Q_{mi} , A_i and L_i

Station Number	$\log Q_{mi}$	$\log A_i$	$\log L_i$
1	2.5378	3.4751	2.0792
2	3.3007	4.0354	2.2833
3	3.3845	4.3006	2.4216
4	2.8142	3.7645	2.2672
5	2.8352	3.5250	2.0253
6	3.2872	4.3094	2.4843
7	3.2419	4.0759	2.4099
8	2.9718	3.9562	2.2923
9	3.3832	3.8078	2.0828
10	3.4182	4.3424	2.4639
11	3.3648	4.2227	2.3560
SUM	34.5395	43.8150	25.1658

The sum of the products of the logarithms of Q_{mi} , A_i and L_i were calculated as:

$$\sum_{i=1}^{11} (\log Q_{mi})(\log A_i) = 138.3531; \quad \sum_{i=1}^{11} (\log A_i)(\log L_i) = 100.7038$$

$$\sum_{i=1}^{11} (\log A_i)^2 = 175.4754; \quad \sum_{i=1}^{11} (\log L_i)^2 = 57.8363$$

$$\sum_{i=1}^{11} (\log Q_{mi})(\log L_i) = 79.3189$$

The summation values were substituted into the normal equations (3.5.5) through (3.5.7) to give the following system of equations:

$$34.5395 = 11 \log a + 43.8150b + 25.1658 c \quad \dots\dots\dots(3.6.1)$$

$$138.3531 = 43.8150 \log a + 175.4754 b + 100.7038 c \quad \dots\dots\dots(3.6.2)$$

$$79.3189 = 25.1658 \log a + 100.7038 b + 57.8363 c \quad \dots\dots\dots(3.6.3)$$

The equations (3.6.1), (3.6.2) and (3.6.3) were solved simultaneously using method of successive elimination to obtain the following solutions:

$$a = 4.41, \quad b = 1.89 \quad \text{and} \quad c = -2.20.$$

These values of a, b, c obtained above were then substituted into equation (3.6.1) to give the following equation for the mean annual maximum flood discharge:

$$Q_{mi} = 4.41 A_i^{1.89} L^{-2.20} \quad \dots\dots\dots(3.6.4)$$

3.6.2 (b) Regression Model for Variances of the annual Maximum Flood discharges

Several forms of multiple regression linear equations were tried by the author to relate the variance, V_i to the variables A_i , L_i and it was found that the regression equation of the form given by equation (3.6.1) was the most suitable. Hence, regression analysis for the variance of the discharges was carried out by the author following the same procedure as for the mean annual maximum flood discharge.

The variances calculated by the author at each of the gaging station using equation (2.3.7) are shown in Table (3.6).

Table(3.6): The variance, V , and $\log_{10} V$ of the annual maximumflood discharges of the selected gauging stations in Upper Benue River Basin

River	Gaging Station	Station No.	Variance $V_i (m^6/s^2)$	$\log V_i$
Mayo Kam	Mayo Kam	1	11,522	4.0615
Taraba	Beli	2	170,320	5.2313
Taraba	Tela	3	433,070	5.6366
Bantaji	Suntai	4	35,456	4.5497
Bantaji	Suntai Bakundi	5	19,729	4.2951
Donga	Gindin Dorrowa	6	100,057	5.0002
Donga	Donga	7	67,950	4.8322
Donga	Manya	8	42,011	4.6234
Donga	Abong	9	119,610	5.0778
Katsina-Ala	Sevav	10	229,426	5.4763
Katsina-Ala	Katsina-Ala	11	253,806	5.4045

$$\text{SUM} \quad \sum (\log V_i) = 54.1886$$

The summation of the logarithms of the variables and their products were calculated as follows:

$$\sum_{i=1}^{11} (\log V_i)(\log A_i) = 217.2045; \quad \sum_{i=1}^{11} (\log A_i)(\log L_i) = 100.7038$$

$$\sum_{i=1}^{11} (\log A_i)^2 = 175.4754; \quad \sum_{i=1}^{11} (\log L_i)^2 = 57.8363$$

$$\sum_{i=1}^{11} (\log V_i)(\log L_i) = 124.5203; \quad \sum (\log A_i) = 43.8150,$$

$$\sum \log L_i = 25.1658$$

Following the same procedure as for the mean annual maximum flood discharge, Q_m , the variance, V_i , was found to be related to the drainage area, A_i , and the main stream length, L_i , by the equation:

$$V_i = 2.12 A_i^{3.01} L^{-3.24} \dots\dots\dots (3.6.5)$$

3.6.3 , Log - Normal Model

(a) Regression Model for the mean of the natural logarithm of the annual maximum flood discharges ($\ln Q$)

The mean of $\ln Q$ at each gaging station was calculated by the author using equation (2.3 6) and are shown in Table (3.7).

Table 3.7: The mean of the natural logarithm of the annual maximum discharges at the various stations.

River	Gaging Station	Station No.	Mean of $(\ln Q) = \overline{\ln Q_i}$ $\ln (m^3/s)$
Mayo Kam	Mayo Kam	1	5.7753
Taraba	Beli	2	7.5779
Taraba	Tela	3	7.7575
Bantaji	Suntai	4	6.4393
Bantaji	Suntai Bakundi	5	6.5108
Donga	Gindin Dorrowa	6	7.5580
Donga	Donga	7	7.4543
Donga	Manya	8	6.8214
Donga	Abong	9	7.7803
Katsina-Ala	Sevav	10	7.8496
Katsina-Ala	Katsina-Ala	11	7.7145

In order to use the above table in the normal equations (3.5.5) through (3.5.7), the natural logarithm of the mean of $\ln Q_i$, A_i , and L_i were calculated and are shown in Table (3.8).

Table 3.8: The natural logarithm of variable $\overline{\ln Q_i}$, A_i and L_i .

Station Number	$\ln (\overline{\ln Q_i})$	$\ln A_i$	$\ln L_i$
1	1.7536	8.0017	4.7875
2	2.0252	9.2919	5.2575
3	2.0487	9.9024	5.5759
4	1.8624	8.6682	5.2204
5	1.8735	8.1167	4.6634
6	2.0226	9.9228	5.7203
7	2.0088	9.3850	5.5491
8	1.9201	9.1094	5.2781
9	2.0516	8.7678	4.7958
10	2.0605	9.9988	5.6733
11	2.0431	9.7232	5.4250
SUM	21.6701	100.8879	57.9463

The sum of the products of the natural logarithm of the variables $\overline{\ln Q}$, A_i and L_i were calculated as:

$$\sum_{i=1}^{11} (\ln (\overline{\ln Q})) \cdot (\ln A_i) = 199.3565; \quad \sum_{i=1}^{11} (\ln \overline{\ln Q})(\ln L_i) = 114.3897;$$

$$\sum_{i=1}^{11} (\ln A_i)^2 = 930.3549; \quad \sum_{i=1}^{11} (\ln L_i)^2 = 306.6413;$$

$$\sum_{i=1}^{11} (\ln A_i) (\ln L_i) = 533.9212$$

Following the same procedure as for the normal model, the following regression equation was obtained:

$$\overline{\ln Q_i} = 3.04 + 0.266 A_i - 0.30 L_i \quad \dots\dots\dots(3.6.6)$$

3.6.3 (b) Regression Model for the Variance of the natural logarithm of the annual maximum flood discharges ($\ln Q$)

The variance of the $\ln Q$ at each gaging station was calculated by the author and are shown in Table (3.9) together with the natural logarithm.

River	Gaging Station	Station No.	Variance of $\ln Q_i$ $V_i \ln m^6/s^2$	$\ln V_i$
Mayo Kam	Mayo Kam	1	0.1594	- 1.8363
Taraba	Beli	2	0.0479	- 3.0386
Taraba	Tela	3	0.0879	- 2.4315
Bantaji	Suntai	4	0.0868	- 2.4441
Bantaji	Suntai Bakundi	5	0.0404	- 3.2089
Donga	Gindin Dorrowa	6	0.0250	- 3.6889
Donga	Donga	7	0.0216	- 3.8351
Donga	Manya	8	0.0441	- 3.1213
Donga	Abong	9	0.0217	- 3.8304
Katsina-Ala	Sevav	10	0.0439	- 3.1258
Katsina-Ala	Katsina-Ala	11	0.0360	- 3.3242
SUM				- 33.8851

Table (3.9): Variance, V of $\ln Q$, and $\ln V$ for the selected gauging stations in Upper Benue River Basin.

The sum of the variables V_i , A_i , and L_i as well as the sum of their products were calculated as:

$$\begin{aligned}\sum_{i=1}^{11} (\ln L_i) &= 57.9463; & \sum_{i=1}^{11} (\ln A_i) &= 100.8879; \\ \sum_{i=1}^{11} (\ln V_i) &= -33.8851; & \sum_{i=1}^{11} (\ln V_i)(\ln A_i) &= -312.4273; \\ \sum_{i=1}^{11} (\ln V_i)(\ln L_i) &= -179.0428; & \sum_{i=1}^{11} (\ln A_i)^2 &= 930.3549; \\ \sum_{i=1}^{11} (\ln L_i)^2 &= 306.6413; & \sum_{i=1}^{11} (\ln A_i)(\ln L_i) &= 533.9212\end{aligned}$$

Following the same procedure as for the mean annual flood discharges, Q_m , the natural logarithm of the variance, $\ln V_i$, was found to be related to the drainage area, A_i and the main stream length, L_i by the equation:

$$\ln V_i = 3.7 \times 10^{-9} A_i^{0.99} L_i^{1.37} \dots\dots\dots(3.6.7)$$

3.6.4 Exceedance Model

The exceedance discharges above the base level of 2,200 m³/s was chosen because by interpolation from frequency curve shown in Figure 3.14, this discharge corresponds to a return period of about 10-years which according to (Wood, 1973) is high enough. For the reason discussed earlier, only six gaging stations records were used for the regression analysis.

3.6.4 (a) Regression Model for Mean Exceedance Flood discharge

The mean exceedance flood discharge,

for each gaging station record was calculated by the author using equation (2.3.6) and results are shown in Table (3.19) together with the corresponding drainage area A_i , and main stream length, L_i .

Table (3.10): Mean exceedance discharge, drainage area, and main stream length for six gaging stations.

River	Gaging Station	Station No.	Mean Exceedance Discharge Z_m (m^3/s)	Drainage Area A_i (km^2)	Main Stream length L_i (km)
Taraba	Beli	2	271.9	10,850	192
Taraba	Tela	3	700	19,979	264
Donga	Donga	7	136	11,909	257
Donga	Abong	9	319.2	6,424	121
Katsina-Ala	Sevav	10	613.1	22,000	291
Katsina-Ala	Katsina-Ala	11	432.6	16,700	227

The logarithm to base ten of the mean exceedance discharge Z_m , the drainage area, A_i , and the main stream length, L_i , were calculated and are shown in Table (3.11).

Table (3.11): The logarithm of mean exceedance discharge, drainage area and length of main stream for six gaging stations.

Station No.	$\log Z_m$	$\log A_i$	$\log L_i$
2	2.4344	4.0354	2.2833
3	2.8451	4.3006	2.4216
7	2.1335	4.0759	2.4099
9	2.5041	3.8078	2.0828
10	2.7875	4.3424	2.4639
11	2.5041	3.8078	2.0828
SUM	15.3407	24.7848	14.0175

The sum of the products of the logarithm to base ten of

Z_m , A_i and L_i were calculated as:

$$\begin{aligned}\sum (\log Z_m) (\log A_i) &= 63.5264; \quad \sum (\log Z_m) (\log L_i) = 35.8840 \\ \sum (\log A_i)^2 &= 102.5795; \quad \sum (\log L_i)^2 = 32.8448\end{aligned}$$

$$\sum (\log A_i) (\log L_i) = 58.0297$$

Substituting these values into the normal equations (3.5.5) through (3.5.8) and following the same procedure as for the mean annual maximum flood discharge, Q_m , the mean exceedance discharge, Z_m was found to be related to the drainage area, A_i , and the main stream length, L_i , by the equation:

$$Z_m = 0.02 A^{2.961} L^{-3.414} \dots\dots\dots (3.6.7)$$

3.6.4 (b) Regression Model for Variance of Exceedance Flood Discharge

The variance of the exceedance flood discharge at each gaging station was calculated by the author using equation (2.3.7) from the six gaging stations records used for regression model for the mean exceedance discharge above. The variances of the exceedance flood discharges are shown in Table (3.12) together with their logarithm to base ten.

Variance, V and $\log V$ of the exceedance flood discharges for the six selected gauging stations in Upper Benue River Basin.

Table (3.12)

River	Gaging Station	Station No.	Variance of exceedance discharge V_i (m^6/s^2)	$\log V_i$
Taraba	Beli	2	37,170.1	4.5702
Taraba	Tela	3	66,666.7	4.8239
Donga	Donga	7	8,192	3.9134
Donga	Abong	9	75,874.3	4.8801
Katsina-Ala	Sevav	10	199,107.6	5.2991
Katsina-Ala	Katsina-Ala	11	122,151.5	5.0869
		SUM	-	28.5736

The sum of the logarithm of the variables, V_i , A_i and L_i and the sum of the product of their logarithms to base ten were calculated as:

$$\sum \log V_i = 28.5736 ; \quad \sum \log A_i = 24.7848, \quad \sum \log L_i = 14.0175$$

$$\sum (\log V_i) (\log A_i) = 118.2126 ; \quad \sum (\log V_i) (\log L_i) = 66.7531$$

$$\sum (\log A_i)^2 = 102.5795 ; \quad \sum (\log L_i)^2 = 32.8448$$

$$\sum (\log A_i) (\log L_i) = 58.0297.$$

Following the same procedure as for the mean annual maximum flood discharge, the variance, V_i , was found to be related to the drainage area, A_i , and the mainstream length L_i by the equation:

$$V_i = 0.0848 A_i^{5.476} L_i^{-7.185} \dots\dots\dots(3.6.9)$$

3.7

SUMMARY OF THE REGRESSION MODELS

The regression models derived above are summarized in Table (3.13) below:

Table (3.13): Summary of the regression models.

Distribution	Regression model for	
	Mean discharge	Variance of discharge
Normal	$Q_{mi} = 4.41 A_i^{1.89} L_i^{-2.20}$	$V_i = 2.12 A_i^{3.01} L_i^{-3.24}$
log-normal	$\ln Q_i = 3.04 A_i^{0.266} L_i^{-0.30}$	$\ln V_i = 3.7 \times 10^{-9} A_i^{0.99} L_i^{1.37}$
Exceedance	$Z_m = 0.02 A_i^{2.961} L_i^{-3.414}$	$V_{rm} = 0.0848 A_i^{5.476} L_i^{-7.185}$

These models were found to be suitable because the results obtained from them are within $\pm 10\%$ of the actual values.

CHAPTER 4APPLICATION OF FORMULATED MODELS TO RIVER TARABA AT GASSOL
GAUGING STATION IN UPPER BENUE RIVER BASIN, NIGERIA.**4.1** SAMPLE PARAMETER ESTIMATION**4.1.1** General Information

The models formulated in chapter 2 were applied to River Taraba at Gassol gaging station. River Taraba is a tributary of River Benue in the Upper Benue River Basin in Gongola State of Nigeria.

The historical record available was from 1955 through 1988 - and is shown in Appendix A-12. The drainage area is $20,513 \text{ (km)}^2$ and the main stream length is 291 Km. This sample information was used to calculate the sample parameters of the three selected distributions.

4.1.2 Normal Distribution

The mean maximum yearly flood discharge, m , calculated from equation (2.3.6) is:

$$m = 1769.6 \text{ m}^3/\text{s}$$

The sample variance, w , calculated from equation (2.3.7) is:

$$w = 95,160 \text{ m}^6/\text{s}^2$$

The number of years of record, n is:

$$n = 34$$

The number of degrees of freedom, V , is:

$$V = 34 - 1 = 33 \text{ from equation (2.3.8).}$$

4.1.3 Log - Normal distribution

The natural logarithm of all the annual maximum flood discharge in Appendix A-12 were calculated and the following parameters calculated:

$$n = 34 \text{ (years)}$$

$$m = 7.4635 \quad \ln (m^3/s) \quad \text{from equation (2.3.6)}$$

$$w = 0.0317 \quad \ln (m^6/s^2) \quad \text{from equation (2.3.7)}$$

$$v = 34 - 1 = 33 \text{ (years)} \quad \text{from equation (2.3.8)}$$

4.1.4 Exceedance distribution

The base level discharge was chosen as 2,200 m³/s which corresponds to 10-year recurrence interval (by interpolating from frequency curve in Figure 3.14).

Exceedance events were considered to be flows greater than 2,200 m³/s.

Also from Appendix A-12 it is seen that only 3 flows exceed the base level discharge and that the number of years of record is 34 (1955 through 1988). Therefore the following sample parameters are obtained for the exceedance distribution:

$$n = 3$$

$$T = 34 \text{ (years)}$$

$$\sum_{i=1}^n z = (2,500 - 2,200) + (2,400 - 2,200) + (2,280 - 2,200) = 580 \text{ m}^3/\text{s}.$$

4.2

PRIOR PARAMETERS ESTIMATION

4.2.1

General Principle

In chapter 3 the regression models obtained for the mean and variance of the series of annual flood discharges using regional information from other stations in the Upper Benue River Basin were all of the form:

$$Q = a A_i^b L_i^c$$

With the aid of the first and second order approximation of the Taylor's series expansion, the above regression models will be used in computing the prior parameters for the three selected distributions.

The covariance and correlation coefficient between A and L are calculated with the aid of the results obtained from Table (4.1).

Table 4.1: Calculation of Covariance and Correlation Coefficient

Station Number i	A_i (km) ²	L_i (km)	$A_i - A_m$ - 8,781.5	$L_i - L_m$
1	2,986	120	- 917.5	- 85.82
2	10,850	192	8,211.5	- 13.82
3	19,979	264	- 5,952.5	58.18
4	5,815	185	- 8,417.5	- 20.82
5	3,350	106	8,622.5	- 99.82
6	20,390	305	141.5	99.18
7	11,909	257	- 2,727.5	51.18
8	9,040	196	- 5,343.5	- 9.82
9	6,424	121	10,232.5	- 84.82
10	22,000	291	4,932.5	85.18
11	16,700	227		21.18
Sum			$\sum_{i=1}^{11} (A_i - A_m)(L_i - L_m) = 4,526,735.09$	
Mean	$A_m = 11,767.5$	$L_m = 205.82$	-	-
Standard deviation	$S_A = 7007.6$	$S_L = 69.85$	-	-

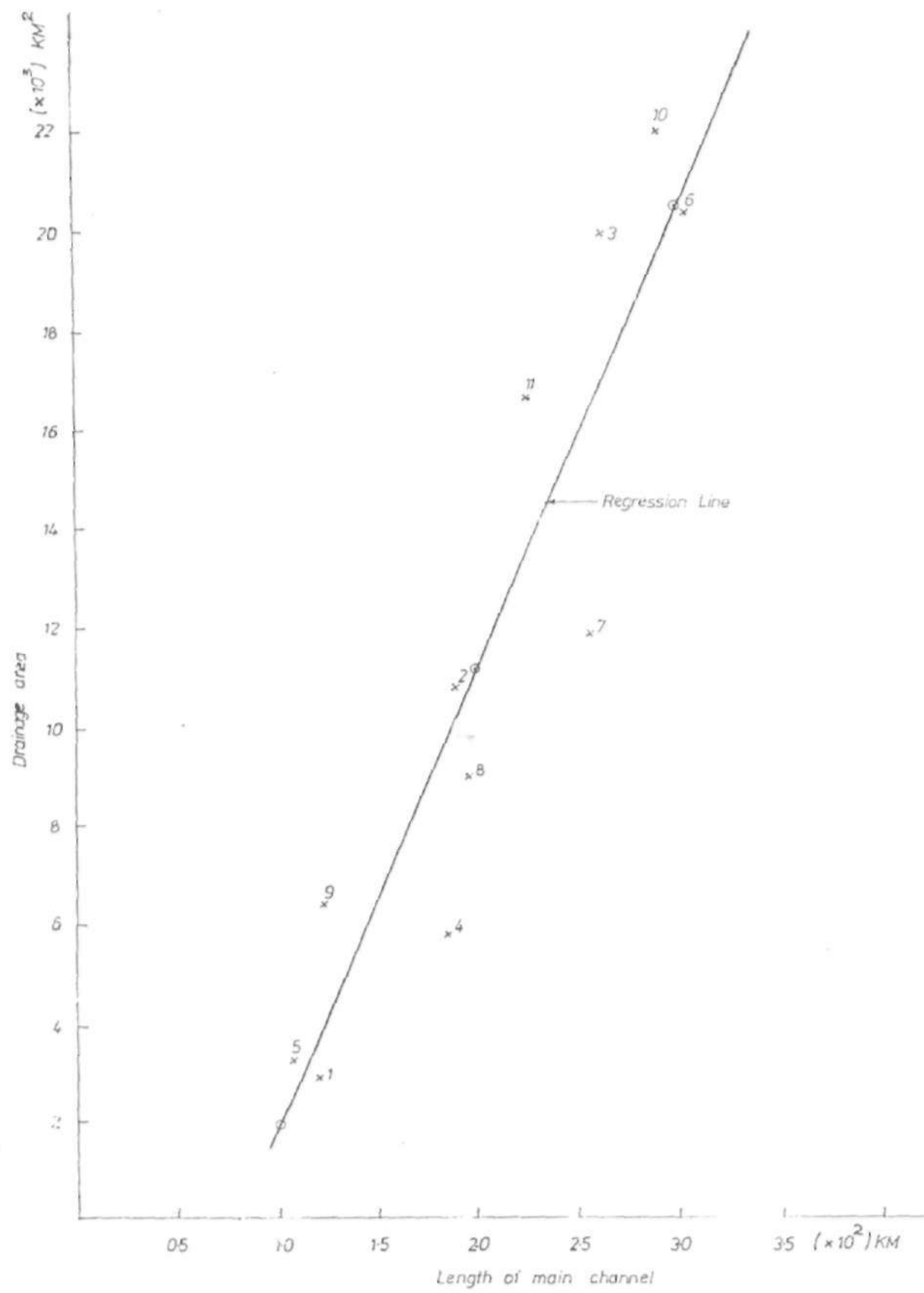


FIG-4.0 CORRELATION BETWEEN DRAINAGE AREA AND LENGTH OF MAIN CHANNEL FOR THE STATIONS

The covariance, $\text{cov} [A, L]$ is given by:

$$\text{cov} [A, L] = \text{cov} [L, A] = \frac{1}{n} \sum_{i=1}^n (A_i - A_m)(L_i - L_m) = \frac{4,526.735.09}{11}$$

$$\therefore \text{cov} [A, L] = \text{cov} [L, A] = 411,521.4.$$

The correlation coefficient between A, and L is given by:

$$r_{A,L} = \frac{\text{cov} [A, L]}{S_A \cdot S_L} = \frac{411,521.4}{7007 \times 69.85} = 0.84$$

Thus since $r_{A,L}$ is high, it means there is a high linear dependence between A and L and hence equations (2.3.85) and (2.3.86) were used to calculate the expected value and variance of the mean discharge respectively. The high linear correlation between A and L is confirmed by Figure 4.0.

4.2.2 Normal Model

4.2.2 (a) Calculation of Expectation of Mean discharge

The expected value of the mean discharge, $E [p_Q | I_R]$ was calculated from equation (2.3.85) using the regression model:

$$Q = 4.41 A_i^{1.89} L_i^{-2.20}$$

as follows:

$$a A_m^b L_m^c = 4.41 (11,767.5)^{1.89} (205.82)^{-2.20} = 1,771.57 \text{ m}^3/\text{s}$$

$$ab(b-1)A_m^{b-2} L_m^c \text{ var} [A] = [4.41 \times 1.89 \times 0.89 (11,767.5)^{-0.11} (205.82)^{-2.20} (7007.6)^2] = 1,056.77 \text{ m}^3/\text{s}$$

$$abc A_m^{b-1} L_i^{c-1} \text{ cov} [A, L] = [4.41 \times 1.89 \times (-2.20) (11,767.8)^{0.89} (205.82)^{-3.2} (411,521.4)] = -1251.60 \text{ m}^3/\text{s}$$

$$ac(c-1) A_m^b L_m^{c-2} \text{ var} [L] = [4.41 \times (-2.2) (-3.2) (11,767.8)^{1.89} (205.82)^{-4.2} (69.85)^2] = 1436.43 \text{ m}^3/\text{s}$$

Substituting these values in equation (2.3.85) gives:

$$E [U_Q] = \left\{ 1,771.57 + \frac{1}{2} [1,056.77 - 2 \times 1251.60 + 1,436.43] \right\} \\ = 1766.57 \text{ m}^3/\text{s}.$$

$$\text{Thus } E [U_Q | I_R] = 1766.57 \text{ m}^3/\text{s}$$

4.2.2 (b) Calculation of variance of the mean annual discharge

The variance of the mean annual discharge, $V [U_Q | I_R]$ was calculated from equation (2.3.86) using regression model

$$Q = 4.41 A_i^{1.89} L_i^{-2.20} \text{ as follows:}$$

$$(ab A_m^{b-1} L_m^{c-2})^2 \cdot \text{var} [A] = \left\{ [4.41 \times 1.89 \times (11,767.5)^{1.89} (205.82)^{-2.20}]^2 \right. \\ \left. (7007.6)^2 \right\} = 3.9757 \times 10^6$$

$$a^2 bc A_m^{2b-1} L_m^{2c-1} \text{cov} [A, L] = [4.41]^2 \times 1.89 \times (-2.20) (11,767.5)^{2.78} \\ (205.82)^{-5.4} (411521.4)] = -2.2173 \times 10^6$$

$$(ac A_m^b L_m^{c-1})^2 \cdot \text{var} [L] = \left\{ [4.41 \times 2.2 \times (11,767.5)^{1.89} (205.82)^{-3.2}]^2 \right. \\ \left. (69.85)^2 \right\} = 1.7495 \times 10^6$$

Substituting these values into equation (2.3.86) gives:

$$V [U_Q | I_R] = (3.9757 - 2 \times 2.2173 + 1.7495) \times 10^6 = 1.2906 \times 10^6 \\ V [U_Q | I_R] = 1.29 \times 10^6 \text{ m}^6/\text{s}^2$$

4.2.2. (c) Calculation of the expectation of variance of the annual discharges

The expectation of the variance of the annual discharges,

$E [\sigma^2 | I_R]$ was calculated from equation (2.3.85) using the regression model: $V = 2.12 A^{3.01} L^{-3.24}$. This was done by replacing Q in equation (2.3.85) by V and substituting $a = 2.12$, $b = 3.01$ and $c = -3.24$.

The result obtained for the expected value is:

$$E[\sigma^2 | I_R] = 1.46 \times 10^5 \text{ m}^6/\text{s}^2.$$

4.2.2. (d) Calculation of the variance of variance of the annual flood discharges:

The variance of the variance of the annual flood discharges, $V[\sigma^2 | I_R]$ was calculated from equation (2.3.86) using the same regression model and making the same substitutions as discussed in item (iii) above. The result obtained for the variance of the annual flood discharges was:

$$V[\sigma^2 | I_R] = 1.63 \times 10^{10} \text{ m}^{12}/\text{s}^4.$$

The result of the previous calculations are summarized in Table (4.2).

Table 4.2: Expectations and Variance of μ and σ^2 for Normal Model.

Parameters	Expectation	Variance
Mean Discharge, μ_Q	1766.57 m ³ /s	$1.29 \times 10^6 \text{ m}^6/\text{s}^2$
Variance of discharge, σ^2	$1.46 \times 10^6 \text{ m}^6/\text{s}^2$	$1.63 \times 10^{10} \text{ m}^{12}/\text{s}^4$

The values shown in Table (4.2) were substituted into equation (2.3.64) and equations (2.3.73) through (2.3.75) to obtain the prior parameters as follows:

$$m' = E[\mu_Q | I_R] = 1766.57 \text{ m}^3/\text{s}$$

$$n' = \frac{E[\sigma_Q^2 | I_R]}{V[\mu_Q | I_R]} = \frac{1.46 \times 10^5}{1.29 \times 10^6} = 0.11 \approx 0$$

$$\begin{aligned} v' &= 4 + \frac{2 E^2[\sigma_Q^2 | I_R]}{V[\sigma_Q^2 | I_R]} = 4 + \frac{2 \times (1.46 \times 10^5)^2}{1.63 \times 10^{10}} \\ &= 4 + 2.62 = 6.62 \approx 7. \end{aligned}$$

$$v' = 7$$

$$w' = \frac{v' - 2}{v'} E[\sigma_Q^2 | I_R] = \frac{7 - 2}{7} \times 1.46 \times 10^5 = 1.04 \times 10^5 \text{ m}^6/\text{s}^2.$$

Observe that the mean annual discharge $m = 1769.6 \text{ m}^3/\text{s}$ obtained from the sample information is close to the mean annual discharge $m' = 1766.57 \text{ m}^3/\text{s}$ obtained from regional information. This shows that, for the normal distribution, the regional information is adequate in predicting the mean annual discharge of the River Taraba at Gassol gaging station.

4.2.3 Log-normal Model Prior Parameters

The prior parameters were obtained from the regression models for the mean and variance of the natural logarithm of the series of annual floods by using the same regional information as for the normal distribution. From chapter 3 the equation obtained from the regression of $\ln Q$ on A_i and L_i was equation (3.6.6).

$$\ln Q_m = 3.04 A_i^{0.266} L_i^{-0.30}$$

Following the same procedure as for Normal distribution and substituting:

$$a = 3.04$$

$$b = 0.266$$

$$c = 0.30$$

into equations (2.3.85) and (2.3.86), the following results were obtained:

$$E[\ln Q | I_R] = 7.2496 \ln (\text{m}^3/\text{s})$$

$$V[\ln Q | I_R] = 0.4618$$

From chapter 3 the regression model for the variance was equation(3.6.7)

$$\ln[V] = 3.7 \times 10^{-9} A^{0.99} L^{1.37}$$

Again following the same procedure as for Normal model and substituting:

$$a = 3.7 \times 10^{-9}$$

$$b = 0.99$$

$$c = 1.37$$

into equations (2.3.85) and (2.3.86), the following results are obtained:

$$E[\sigma^2_{\ln Q} | I_R] = 0.07371 ; \quad V[\sigma^2_{\ln Q} | I_R] = 0.003515$$

The results above are summarized in Table (4.3).

Table 4.3 : Expectations and Variance of μ and σ^2 for Log-Normal Model.

Parameters	Expectation	Variance
Mean Natural Logarithm of Discharge, $\mu_{\ln Q}$	$7.2496 \ln(m^3/s)$	$0.4618 \ln(m^6/s^2)$
Variance of Natural Logarithm of Discharge, $\sigma^2_{\ln Q}$	$0.07371 \ln(m^6/s^2)$	$0.003515 \ln(m^{12}/s^2)$

The values shown in Table (4.3) were substituted into equation (2.3.64) and equations (2.3.73) through (2.3.75) to obtain the following prior parameters:

$$m' = E[\mu_{\ln Q} | I_R] = 7.2496 \ln(m^3/s)$$

$$n' = \frac{E[\sigma^2_{\ln Q} | I_R]}{V[\mu_{\ln Q} | I_R]} = \frac{0.07371}{0.4618} = 0.16 \approx 0$$

$$v' = 4 + \frac{2 E[\sigma^2_{\ln Q} | I_R]}{V[\sigma^2_{\ln Q} | I_R]} = 4 + \frac{2 \times (0.07371)^2}{0.003515}$$

$$= 4 + 3.09 = 7$$

313/13

KASHIJI ARAHIM LIBRARY

$$w' = \frac{v'-2}{v'} \cdot E[\sigma^2_{\ln Q} | I_R] = \frac{7-2}{7} (0.07371) = 0.052$$

$$w' = 0.052 \quad \ln(m^6/s^2)$$

4.2.4 Exceedance Model: Prior Parameters Estimation

4.2.4 (a) For event magnitude distribution, α

The prior parameters of the event magnitude distribution, α , were obtained by using the regression equation (3.6.8) obtained in chapter 3 for the mean exceedance discharge:

$$Z_m = 0.02 A_i^{2.961} L_i^{-3.414}$$

and following the same procedure as for the normal model. The following results were obtained:

$$E[Z_m] = 490.5 \text{ m}^3/\text{s}$$

$$V[Z_m] = 109,340 \text{ m}^6/\text{s}^2$$

The values of the $E[Z_m]$ and $V[Z_m]$ were substituted into equations (2.5.51) and (2.5.52) to give the values of the prior parameters w' and l' respectively as:

$$w' = 1 + \frac{E^2[Z_m]}{V[Z_m]} = 1 + \frac{(490.5)^2}{109,340} = 1 + 2.2$$

$$w' = 3$$

$$l' = w'E[Z_m] = 3 \times 490.5 = 1471.5 \text{ m}^3/\text{s}.$$

4.2.4 (b) Prior Parameters Estimation for the Average Arrival Rate, V .

According to (Wood and Rodriguez-Iturbe, 1975a), the engineer may make some deductive statements about the expected value of V , based on experience.

Table (4.4) shows the flood discharges corresponding to the base-level (Ten-year Return period) for some of the rivers in the Upper Benue River Basin. These base level flood discharges, Q_b , in column (1) were interpolated from the frequency curves Figures 3.4, 3.9, 3.12 and 3.13.

Column (2) shows the number of years of record, T . Column (3) shows the number of discharges, n , exceeding the base level discharge Q_b . Column (4) shows the average arrival rate, v , for the various rivers, calculated from equation (2.5.62) using Maximum Likelihood Estimate Method.

From Table (4.4) one such deductive statement the engineer may simply have made is "based upon my experience in the area, my best estimate of V is 0.1 and there is a 50-50 chance that V could be plus or minus 0.030 of 0.1".

Table 4.4: Calculation of average arrival rate v , for some gauging stations using maximum likelihood method.

River	Gaging Station	Base Level Flood discharge Q_b (m^3/s) (1)	T (years) (2)	n (3)	$V = \frac{n}{T}$ (4)
Taraba	Beli	2,500	31	3	0.10
Donga	Donga	2,120	32	3	0.09
Katsina-Ala	Sevav	3,400	31	3	0.10
Katsina-Ala	Katsina-Ala	2,950	31	3	0.10

Applying equation (2.5.53) by making the following substitutions:

$$E = 0.030$$

$$n = 3$$

The variance of the average arrival rate, V was calculated as:

$$\text{Var}[V] = \frac{1}{3-2} \quad 2 \times (0.03)^2 = 1.80 \times 10^{-3}$$

Applying equation (2.5.54)

$$E[V] = 0.1$$

Equations (2.5.57) and (2.5.58) were then applied to estimate the prior parameters s' and u' respectively.

$$s' = \frac{E[V]}{\text{Var}[V]} = \frac{0.1}{1.80 \times 10^{-3}} = 55.6 \approx 56$$

$$u' = s' E[V] - 1 = 56 \times 0.1 - 1 = 4.6 \approx 5.$$

4.3 POSTERIOR PARAMETERS ESTIMATION

The posterior parameters for the models were obtained from the sample parameters and prior parameters calculated above.

4.3.1 Estimation of Posterior Parameters for Normal Model

Equations (2.3.33) through (2.3.36) derived earlier in chapter two were applied for calculating the posterior parameters of the normal model as follow:

$$n'' = n + n' = 34 + 0 = 34 \text{ years}$$

$$v'' = v' + v + 1 = 7 + 33 + 1 = 41$$

$$\begin{aligned} m'' &= \frac{1}{n''} [nm + n'm'] = \frac{1}{34} [34 \times 1769.6 + 0 \times 1766.57] \\ &= 1769.6 \text{ m}^3/\text{s} \end{aligned}$$

$$\begin{aligned} w'' &= \frac{1}{v''} [nm^2 + vw + n'm'^2 + v'w' - n'm''^2] \\ &= \frac{1}{41} [34(1769.6)^2 + 33 \times 95,160 + 0 \times (1766.57)^2 + 7 \times 1.04 \\ &\quad \times 10^5 - 34 \times (1769.6)^2] \\ &= \frac{1}{41} [33 \times 95,160 + 7 \times 1.04 \times 10^5] = 94,348 \text{ m}^6/\text{s}^2 \end{aligned}$$

4.3.2 Estimation of Posterior Parameters for Log-Normal Model

Equations (2.3.33) through (2.3.36) were also used in calculating the posterior parameters as follow:

$$n'' = n' + n = 0 + 34 = 34 \text{ years}$$

$$v'' = v' + v + 1 = 7 + 33 + 1 = 41$$

$$\begin{aligned} m'' &= \frac{1}{n''} [nm + n'm'] = \frac{1}{34} [34 \times 7.4635 + 0 \times 7.2496] \\ &= 7.4635 \quad \ln(m^3/s) \end{aligned}$$

$$\begin{aligned} w'' &= \frac{1}{v''} [nm^2 + vw + n'm'^2 + v'w' - n''m''^2] \\ &= \frac{1}{41} [34 \times (7.4635)^2 + 33 \times 0.0317 + 0 \times (7.2496)^2 + 7 \times 0.05265 \\ &\quad - 34(7.4635)^2] \\ &= \frac{1}{41} [33 \times 0.0317 + 7 \times 0.05265] = 0.0345. \end{aligned}$$

4.3.3 Estimation of Posterior Parameters for Exceedance Model

The posterior parameters for the event magnitude distribution, α can be evaluated by applying equations (2.5.23) and (2.5.24) as follow:

$$l'' = l' + \sum_{i=1}^n z_i = 1471.5 + 580 = 2051.5 m^3/s.$$

$$w'' = w' + n = 3 + 3 = 6$$

The posterior parameters for V can be evaluated using equations (2.5.39) and (2.5.40):

$$s'' = s' + T = 56 + 34 = 90$$

$$u'' = u' + n = 5 + 3 = 8$$

Tables (4.5), (4.6) and (4.7) show the summary of sample, prior and posterior parameters calculated for the various models in this chapter.

Table 4.5 Sample Parameters

Sample Parameters	Models		
	Normal	log - normal	Exceedance
Mean discharge, (m)	1769.6 (m^3/s)	7.4635($\ln \text{m}^3/\text{s}$)	-
Variance, w	95,160 (m^6/s^2)	0.0317($\ln \text{m}^6/\text{s}^2$)	-
number of records, n	34 (years)	34 (years)	-
degree of freedom, v	33	33	-
For α : n	-	-	3
z	-	-	580 (m^3/s)
For v: T	-	-	34 (years)
n	-	-	3

Table 4.6 Prior Parameters

Prior Parameters	Models		
	Normal	log - Normal	Exceedance
mean discharge, (m^3)	1766.57(m^3/s)	7.2496($\ln \text{m}^3/\text{s}$)	-
Variance (w')	1.04x10 ⁵ (m^6/s^2)	0.0527($\ln \text{m}^6/\text{s}^2$)	-
Number of records (n')	0	0	-
degrees of freedom (v')	7	7	-
For α : w'	-	-	3
l'	-	-	1471.5(m^3/s)
For v : w'	-	-	56
u'	-	-	5

Table 4.7 Posterior Parameters

Posterior Parameters	Models		
	Normal	log - Normal	Exceedance
Mean discharge, (m ³ /s)	1769.6 (m ³ /s)	7.4635 ln(m ³ /s)	-
Variance, (w ⁶)	94,348 (m ⁶ /s ⁶)	0.0345 ln(m ⁶ /s)	-
Number of records, (n ⁰)	34	34	-
degrees of freedom, (v ⁰)	41	41	-
For α : w ⁰	-	-	6
l ⁰	-	-	2051.5
For v : s ⁰	-	-	90
u ⁰	-	-	8

4.4 BAYESIAN METHOD OF ANALYSIS4.4.1 Bayesian Transformed - Normal Model

For the normal model the posterior parameters tabulated in Table (4.7) were substituted into equation (2.3.59) to give:

$$\tilde{f}_Q(q|m^0, \frac{I}{w^0}, v^0) = \frac{1}{B(0.5, 0.5 \times 41)} \left[1 + \frac{(q-1769.6)^2 \cdot 34}{94348 \times 41 \cdot 35} \right]^{-\frac{1}{2}(41+1)} \left[\frac{41 \times 94348 \times 35}{34} \right]^{\frac{1}{2}}$$

$$\begin{aligned} \tilde{f}_Q(q|m^0, \frac{I}{w^0}, v^0) &= \frac{5.01 \times 10^{-4}}{B(0.5, 20.5)} \left[1 + 2.51 \times 10^{-7} (q-1769.6)^2 \right]^{-21} \\ &= 1.29 \times 10^{-3} \left[1 + 2.51 \times 10^{-7} (q-1769.6)^2 \right]^{-21} \dots (4.1) \end{aligned}$$

Bayesian Transformed - Log-Normal Model

Substitution of the posterior parameters for the log-normal model in Table (4.7) into equation (2.4.5) gave:

$$\tilde{f}_{\ln(q)} = \frac{2.096}{q} \left[1 + 0.69 (\ln q - 7.4635)^2 \right]^{-21} \dots (4.2)$$

Substitution of the prior parameters from Table (4.6) into equations (2.3.59) and (2.4.5) for the normal and log-normal models respectively gave:

$$f'(q) = 0 \quad \dots\dots\dots (4.3)$$

$$f'(\ln q) = 0 \quad \dots\dots\dots (4.4)$$

4.4.3 Bayesian Transformed - Exceedance Model

For the exceedance model, the posterior parameter shown in Table (4.7) were substituted into equation (2.5.46) to give:

$$\tilde{G}_z(z) = 0.089t \left[1 + \frac{z}{2051} \right]^{-7} \quad \dots\dots\dots (4.5)$$

Substitution of the prior parameters from Table (4.6) into equation (2.5.46) for the exceedance model gave:

$$G'_z(z) = 0.107t \left[1 + \frac{z}{1471} \right]^{-4} \quad \dots\dots\dots (4.6)$$

4.5 CLASSICAL METHOD OF ANALYSIS

4.5.1 Normal Model

The values of the sample parameter m and w from Table (4.5) were substituted into equation (2.3.91) in case of normal distribution to give:

$$\hat{h} = \frac{1}{w} = \frac{1}{95,160} = 1.05 \times 10^{-5} \text{ s}^2/\text{m}^6$$

$$f_N(q|\hat{\mu}, \hat{h}) = (2\pi)^{-\frac{1}{2}} \exp \left[-\frac{1}{2} \cdot 1.05 \times 10^{-5} (q - 1769.6)^2 \right] \cdot (1.05 \times 10^{-5})^{\frac{1}{2}}$$

$$f_N(q|\hat{\mu}, \hat{h}) = 1.29 \times 10^{-3} \exp \left[-5.26 \times 10^{-6} (q - 1769.6)^2 \right] \dots (4.7)$$

4.5.2 Log-normal Model

Similarly, for the log-normal distribution the sample values m and w from Table (4.5) were substituted into equation (2.4.12) to give:

$$f_{LN}(q|\mu, h) = q^{-1} (2\pi)^{-\frac{1}{2}} \exp \left[-\frac{1}{2}h(\ln q - \hat{\mu})^2 \right] \cdot \hat{h}^{\frac{1}{2}}$$

$$\hat{h} = \frac{1}{w} = \frac{1}{0.0317} = 31.55$$

$$= q^{-1} (2\pi)^{-\frac{1}{2}} \exp \left[-\frac{1}{2} \times 31.55 (\ln q - \hat{\mu})^2 \right] \cdot (31.55)^{\frac{1}{2}}$$

$$f_{LN}(q|\hat{\mu}, \hat{h}) = \frac{2.24}{q} \exp \left[-15.775 (\ln q - 7.4635)^2 \right] \dots \dots \dots (4.8)$$

4.5.3 Exceedance Model

In order to estimate the values of α and v , equations (2.5.60) and (2.5.62) respectively were applied to the historical record of River Taraba at Cassol as follow:

$$\hat{\alpha} = n / \sum_{i=1}^n Z_i = \frac{3}{580} = 5.172 \times 10^{-3}$$

$$\hat{v} = \frac{n}{T} = \frac{3}{34} = 0.088$$

These values of α and v were then substituted into equation (2.5.63) to give:

$$G_z(z) = 1 - F_z(z) = 0.088t \exp \left[-5.172 \times 10^{-3} z \right] \dots \dots \dots (4.9)$$

4.6 Graphical Comparison of Classical, Regional and Bayesian

Graphical comparison of classical, regional and Bayesian methods of flood frequency analysis are shown on Figures (4.1), (4.2) and (4.3) for normal, log-normal and exceedance distributions.

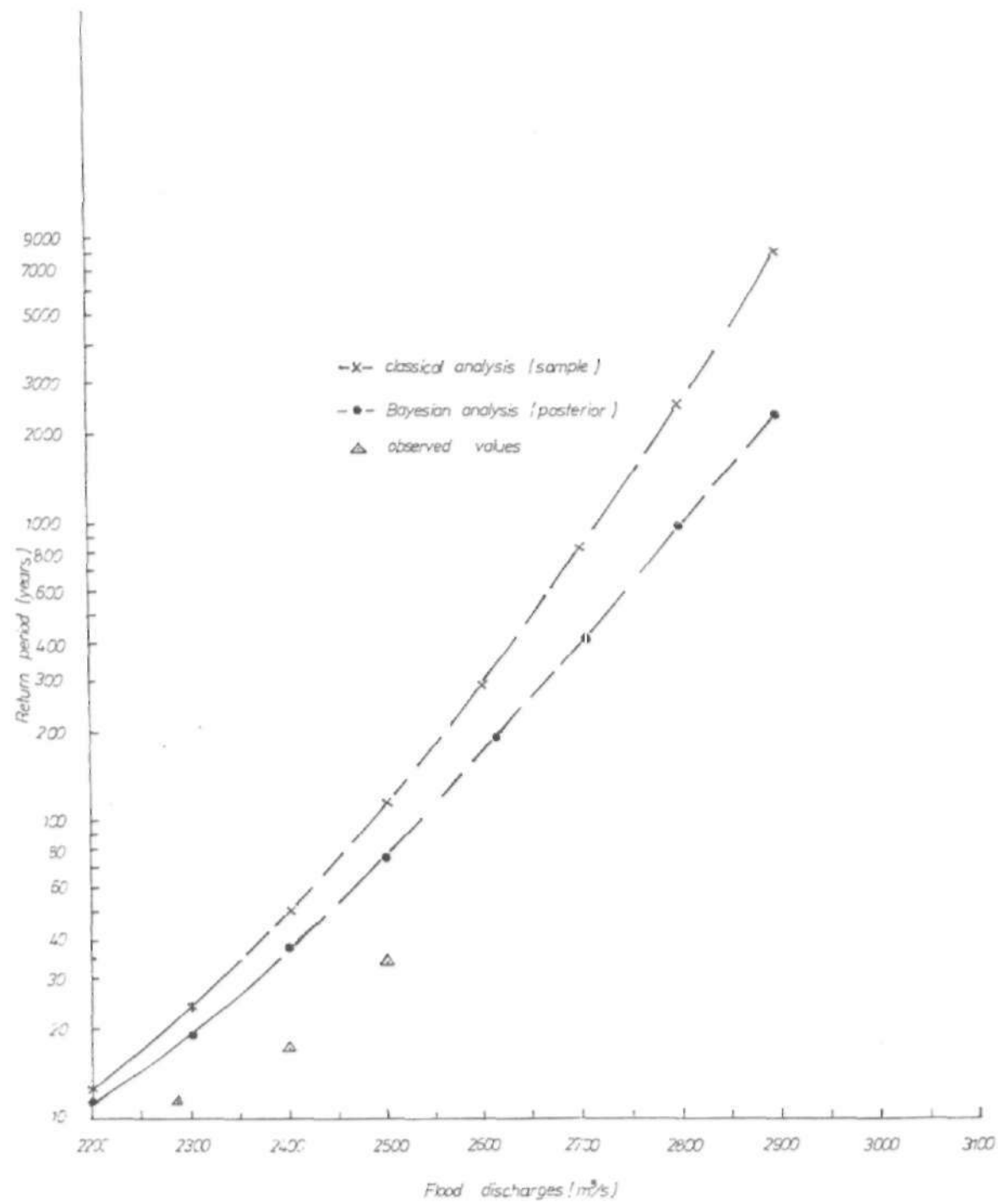


Fig. 4.1 Comparison of classical and Bayesian frequency curves for normal model

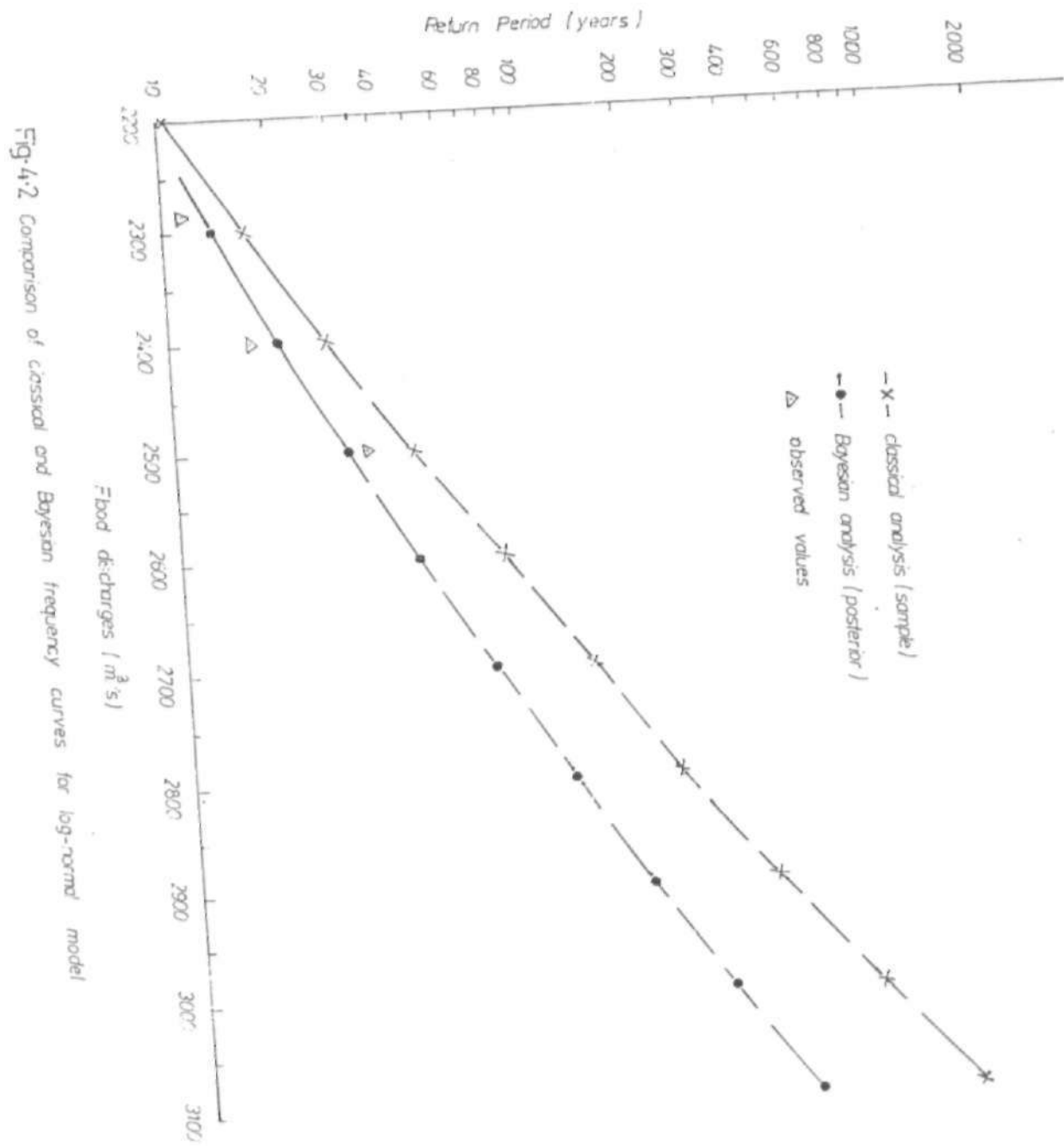


Fig. 4.2 Comparison of classical and Bayesian frequency curves for log-normal model

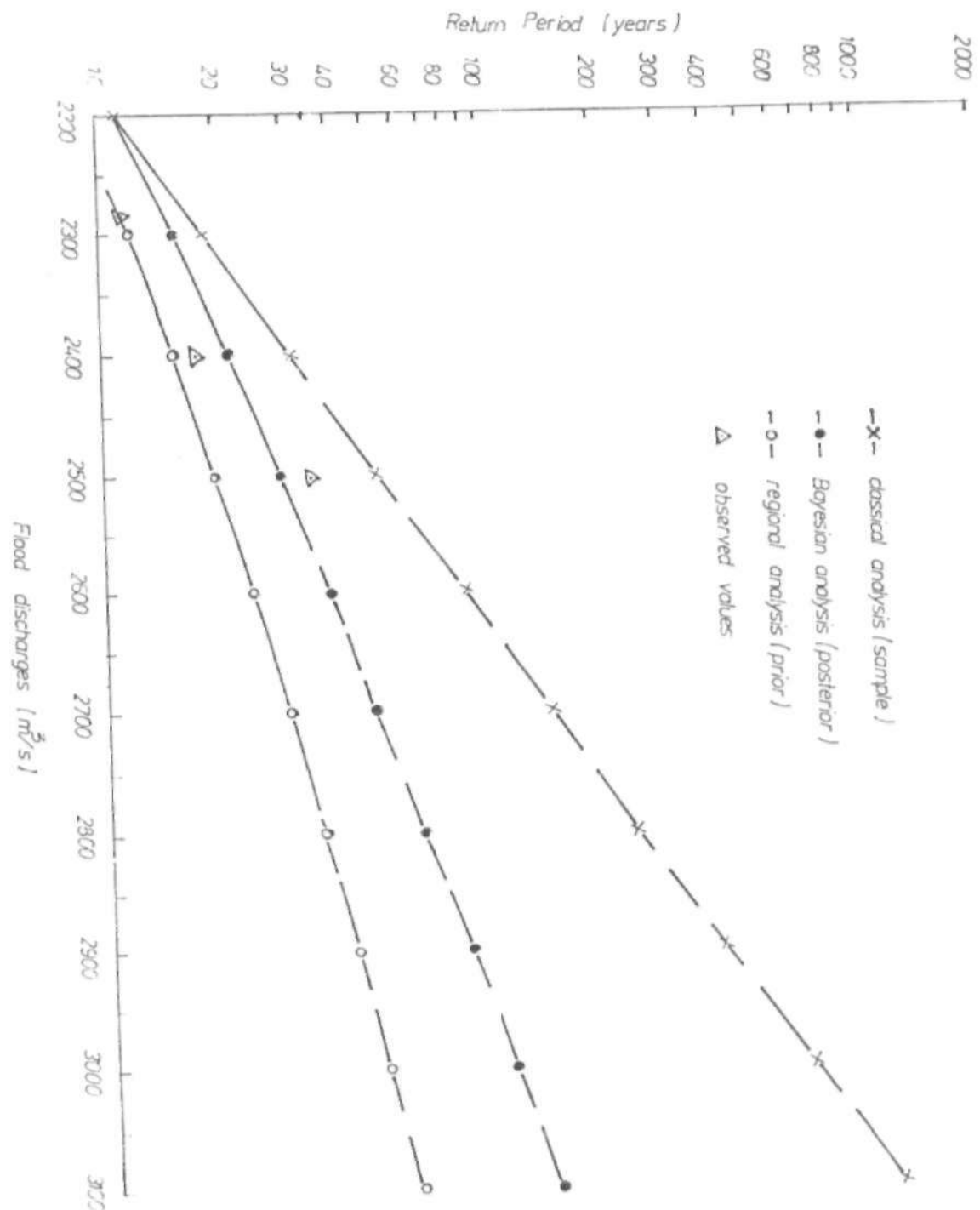


Fig. 4.3 Comparison of classical, Regional and Bayesian frequency curves for exceedance model

MODEL UNCERTAINTY, FORMULATIONS AND APPLICATIONS

5.1

MODEL UNCERTAINTY

In flood frequency analysis hydrologists are often faced with the problem of selecting one statistical model from many competing models. Also, what makes the choice more complex is the fact that many models seem to fit the available data quite well though these models lead to different decisions.

Bayesian procedure can be utilized to select a model from a set of contending models. Thus the use of Bayesian procedure highly reduces the model uncertainty.

Model uncertainty can be considered by defining a composite model of the form: (Wood and Rodriguez-Iturbe, 1975b).

$$\tilde{f}(q|A, \theta) = \theta_1 \cdot f_1(q|A_1) + \dots + \theta_n \cdot f_n(q|A_n) \dots\dots\dots(5.1.1)$$

$$\text{where } A = \bigcup_{i=1}^n A_i$$

The composite model $\tilde{f}(q|A, \theta)$ is conditioned upon a set of unknown model parameters A and an unknown composite model parameters set θ .

The set of probabilistic models $f_1(q|A_1)$, ... and $f_n(q|A_n)$ make up the composite model. These models are conditioned upon a general unknown parameter set A . The parameters $\theta_1, \dots, \theta_n$ take on a value of either 0 or 1; their value is uncertain. If $\theta_i = 1$, then model $f_i(q|A_i)$ is the true model. The constraint:

$$\sum_{i=1}^n \theta_i = 1 \dots\dots\dots(5.1.2)$$

which implies that one and only one model is the true model is imposed.

We shall derive the composite Bayesian model which reduces model uncertainty.

Since we have chosen three models, we shall consider the case where $n = 3$.

The likelihood function for a set of observations Q , where

$Q = \{q_1, q_2, \dots, q_n\}$ (Wood and Rodriguez-Iturbe, 1975b) is just:

$$\begin{aligned} L(A, \theta|Q) &= \theta_1 \prod_{i=1}^n f_1(q_i|A_1) + \theta_2 \prod_{i=1}^n f_2(q_i|A_2) + \theta_3 \prod_{i=1}^n f_3(q_i|A_3) \\ &= \theta_1 L_1(A_1|Q) + \theta_2 L_2(A_2|Q) + \theta_3 L_3(A_3|Q) \dots\dots\dots (5.1.3) \end{aligned}$$

Let us consider now a composite prior distribution. It will be of the form (Wood and Rodriguez-Iturbe, 1975b).

$$\begin{aligned} f'(A, \theta) &= \theta_1 f_1'(A_1|\theta_1 = 1) \cdot P'(\theta_1 = 1) + \theta_2 f_2'(A_2|\theta_2 = 1) \cdot P'(\theta_2 = 1) + \\ &\quad \theta_3 f_3'(A_3|\theta_3 = 1) \cdot P'(\theta_3 = 1) \dots\dots\dots (5.1.4) \end{aligned}$$

where:

$f_i'(A_i|\theta_i = 1)$ — Prior distribution on the parameter set A conditional upon $\theta_i = 1$.

$P'(\theta_i = 1)$ — Prior probability that model i is the 'true' model.

Applying Baye's rule to the posterior distribution of b , conditioned upon the data denoted as $f''(b|\text{data})$ gives:

$$f''(b|\text{data}) = \frac{L(b|\text{data}) \cdot f'(b)}{k} \dots\dots\dots (5.1.5)$$

where:

$L(b|\text{data})$ — likelihood function for b

$f'(b)$ — prior distribution of b

k — normalizing constant

The posterior density function for A, θ is calculated from Baye's rule.

Applying equation (5.1.5)

$$f''(A, \theta) = \frac{L(A, \theta|Q) \cdot f'(A, \theta)}{k^*} \dots\dots\dots(5.1.6)$$

where k^* — normalizing constant given by:

$$k^* = k_1 \cdot P'(\theta_1 = 1) + k_2 \cdot P'(\theta_2 = 1) + k_3 \cdot P'(\theta_3 = 1) \dots\dots(5.1.7)$$

Substituting equations (5.1.3) and (5.1.4) into equation (5.1.6) gives:

$$f''(A, \theta) = \frac{1}{k^*} \left\{ \theta_1 \cdot L_1(A|Q) + \theta_2 \cdot L_2(A_2|Q) + \theta_3 \cdot L_3(A_3|Q) \right\} \cdot$$

$$\left\{ \theta_1 f_1'(A_1|\theta_1=1)P'(\theta_1=1) + \theta_2 \cdot f_2'(A_2|\theta_2=1) \cdot P'(\theta_2=1) + \right.$$

$$\left. \theta_3 \cdot f_3'(A_3|\theta_3=1) \cdot P'(\theta_3=1) \right\} \dots\dots\dots(5.1.8)$$

Expanding equation (5.1.8) and noting that

$$\theta_i L_i(A_i|Q) \cdot f'(A_j|\theta_j=1) \cdot P'(\theta_j=1) = 0$$

where i and j are two different models, gives

$$f''(A, \theta) = \frac{1}{k^*} \left\{ \theta_1 L_1(A_1|Q) \cdot \theta_1 f_1'(A_1|\theta_1=1)P'(\theta_1=1) + \theta_2 L_2(A_2|Q) \cdot \theta_2 f_2'(A_2|\theta_2=1) \cdot \right.$$

$$\left. P'(\theta_2=1) + \theta_3 \cdot L_3(A_3|Q) \cdot \theta_3 f_3'(A_3|\theta_3=1) \cdot P'(\theta_3=1) \right\} \dots\dots\dots(5.1.9)$$

Applying equation (5.1.5)

$$f''(A_i|\theta_i=1) = \frac{L(A_i|Q) \cdot f'(A_i|\theta_i=1)}{k_i}$$

$$L(A_i|Q) \cdot f'(A_i|\theta_i=1) = k_i f''(A_i|\theta_i=1) \dots\dots\dots(5.1.10)$$

Substituting equation (5.1.10) into equation (5.1.9) gives:

$$f''(A, \theta) = \frac{1}{k^*} \left\{ \theta_1^2 k_1 f''(A_1|\theta_1=1) + \theta_2^2 \cdot k_2 f''(A_2|\theta_2=1) \cdot P'(\theta_2=1) + \right.$$

$$\left. \theta_3^2 k_3 f''(A_3|\theta_3=1) \cdot P'(\theta_3=1) \right\} \dots\dots\dots(5.1.11)$$

$$f''(A, \theta) = \theta_1^2 \frac{k_1}{k^*} P'(\theta_1=1) \cdot f''(A_1 | \theta_1=1) + \theta_2^2 \frac{k_2}{k^*} P'(\theta_2=1) \cdot f''(A_2 | \theta_2=1) \\ + \theta_3^2 \frac{k_3}{k^*} P'(\theta_3=1) \cdot f''(A_3 | \theta_3=1) \quad \dots\dots\dots(5.1.12)$$

The posterior model probabilities $P''(\theta_i)$ are given by
(Wood and Rodriguez-Iturbe, 1975b) as:

$$P''(\theta_1=1) = \frac{k_1}{k^*} \cdot P'(\theta_1=1) \quad \dots\dots\dots(5.1.13)$$

$$P''(\theta_2=1) = \frac{k_2}{k^*} \cdot P'(\theta_2=1) \quad \dots\dots\dots(5.1.14)$$

$$P''(\theta_3=1) = \frac{k_3}{k^*} \cdot P'(\theta_3=1) \quad \dots\dots\dots(5.1.15)$$

Substituting equations (5.1.13) through (5.1.15) into equation
(5.1.12) gives:

$$f''(A, \theta) = \theta_1^2 P''(\theta_1=1) \cdot f''(A_1 | \theta_1=1) + \theta_2^2 P''(\theta_2=1) \cdot f''(A_2 | \theta_2=1) \\ + \theta_3^2 P''(\theta_3=1) \cdot f''(A_3 | \theta_3=1) \quad \dots\dots\dots(5.1.16)$$

The composite Bayesian distribution of extreme flood events q
is obtained by applying equation (2.2.4)

$$\tilde{f} = \int_{A, \theta} \tilde{f}(q | A, \theta) \cdot f''(A, \theta) dA d\theta \quad \dots\dots\dots(5.1.17)$$

Substituting equation (5.1.16) into equation (5.1.17)

$$= \int_{A, \theta} [\theta_1 \cdot f_1(q | A_1) + \theta_2 \cdot f_2(q | A_2) + \theta_3 \cdot f_3(q | A_3)] [\theta_1^2 P''(\theta_1=1) \\ f''(A_1 | \theta_1) + \theta_2^2 P''(\theta_2=1) \cdot f''(A_2 | \theta_2=1) + \theta_3^2 P''(\theta_3=1) \cdot f''(A_3 | \theta_3)] dA d\theta \dots\dots(5.1.18)$$

This integration was performed by (Wood et al, 1975b) to
obtain the following:

$$\tilde{f}(q) = P''(\theta_1=1) \cdot \tilde{f}_1(q) + P''(\theta_2=1) \cdot \tilde{f}_2(q) + P''(\theta_3=1) \cdot \tilde{f}_3(q) \dots\dots(5.1.19)$$

(Wood and Rodriguez-Iturbe, 1975b) thus concluded that the composite model
distribution of the models is given by the distributions of the models
weighted by the posterior probability that a particular model is the
true model.

5.2 ANALYTICAL DERIVATION OF THE MARGINAL DENSITY FUNCTION:

5.2.1 Comment

The marginal density function as given by equation (5.1.5), is the probability of observing a set of data and it depends on:

- (i) the probability of model for the stochastic process,
- (ii) the prior probability density function over the parameters of the model,
- (iii) the set of observed data.

The marginal density function will be derived for the three models (Normal, log-normal, and exceedance models).

5.2.2 Marginal Density Function for Normal Model

In chapter 2, it was assumed that the prior on (μ, h) is a natural conjugate prior (that is, it is of the same family as the likelihood function) and is of Normal gamma form (Raiffa and Schlaifer, 1961).

Substituting equation (2.3.21) and equation (2.3.22) into equation (2.3.20) gives:

$$f'(\mu, h | m', w', n', v') = f'_N(\mu | m, hn) \cdot f'_{\gamma_2}(h | w', v') \dots\dots\dots (5.2.1)$$

$$= (2\pi)^{-\frac{1}{2}} \cdot \exp \left[-\frac{1}{2} hn' (m' - \mu)^2 \right] \cdot h^{\frac{1}{2}} (n')^{\frac{1}{2}} \cdot \frac{\exp \left[-\frac{1}{2} v' w' h \right] (\frac{1}{2} v' w' h)^{\frac{1}{2} v' - 1}}{(\frac{1}{2} v' - 1)!} \cdot \frac{1}{\frac{1}{2} v' w' \dots} \dots\dots\dots (5.2.2)$$

where the notations μ, h, m', w', n', v' , are as defined in chapter 2.

$$f'(\mu, h) = (2\pi)^{-\frac{1}{2}} \exp \left[-\frac{1}{2} hn' (m' - \mu)^2 - \frac{1}{2} hv' w' \right] \cdot n'^{\frac{1}{2}} h^{\frac{1}{2}} \cdot h^{\frac{1}{2} v' - 1} \left[\frac{1}{2} v' w' \right]^{\frac{1}{2} v' - 1} \left[\Gamma(\frac{1}{2} v') \right]^{-1} \dots\dots\dots (5.2.3)$$

The marginal likelihood function for the normal model is given by:

$$K_N = \int_{\mu} \int_h L(\mu, h | Q) \cdot f'(\mu, h) d\mu dh \dots\dots\dots (5.2.4)$$

Substituting equations (2.3.15) and (5.2.3) into equation

(5.2.4) gives:

$$K_N = \int_u \int_h \left\{ (2\pi)^{-n/2} \exp \left[-\frac{1}{2} h v w - \frac{1}{2} h n (m - \mu)^2 \right] h^{n/2} \right\} \cdot \left\{ (2\pi)^{-\frac{1}{2}} \exp \left[-\frac{1}{2} h n' (\mu - m')^2 \right. \right. \\ \left. \left. - \frac{1}{2} h v' w' \right] \cdot n'^{\frac{1}{2}} h^{\frac{1}{2}} \cdot h^{\frac{1}{2} v' - 1} \left[\left(\frac{1}{2} v' w' \right)^{\frac{1}{2} v'} \right] \left[\left(\frac{1}{2} v' \right) \right]^{-1} d\mu dh \dots \dots \dots (5.2.5) \right.$$

$$K_N = n'^{\frac{1}{2}} (2\pi)^{-n/2} \cdot (2\pi)^{-\frac{1}{2}} \left(\frac{1}{2} v' w' \right)^{\frac{1}{2} v'} \left[\left(\frac{1}{2} v' \right) \right]^{-1} \int_u \int_h (2\pi)^{-\frac{1}{2}} \cdot \exp \left[-\frac{1}{2} h v w - \frac{1}{2} h n (m - \mu)^2 \right] \cdot \exp \left[-\frac{1}{2} h n' (\mu - m')^2 - \frac{1}{2} h v' w' \right] \cdot h^{\frac{1}{2} v} h^{\frac{1}{2}} \cdot h^{\frac{1}{2}} d\mu dh$$

$$K_N = n'^{\frac{1}{2}} (2\pi)^{-\frac{1}{2}(v+1)} \left[\left(\frac{1}{2} v' w' \right)^{\frac{1}{2} v'} \right] \left[\left(\frac{1}{2} v' \right) \right]^{-1} \int_u \int_h (2\pi)^{-\frac{1}{2}} \left\{ h^{\frac{1}{2}} \exp \left[-\frac{1}{2} h n (m - \mu)^2 \right. \right. \\ \left. \left. - \frac{1}{2} h v w \right] h^{v/2} \right\} \cdot \left\{ h^{\frac{1}{2}} \exp \left[-\frac{1}{2} h n' (m' - \mu)^2 \right] h^{\frac{1}{2} v' - 1} \exp \left[-\frac{1}{2} h v' w' \right] \right\} d\mu dh \dots (5.2.6)$$

From equation (5.2.6), the terms in brackets of the integral have been proved to be identically equal to $f''_N (\mu, h, m'', n'', v'')$ in chapter 2 equation (2.3.30).

Substituting the right side of equation (2.3.27) into the bracketed terms gives:

$$K_N = (n')^{\frac{1}{2}} (2\pi)^{-\frac{1}{2}(v+1)} \left[\left(\frac{1}{2} v' w' \right)^{\frac{1}{2} v'} \right] \left[\left(\frac{1}{2} v' \right) \right]^{-1} \int_u \int_h (2\pi)^{-\frac{1}{2}} h^{\frac{1}{2}} \exp \left[-\frac{1}{2} h n'' (\mu - m'')^2 \right] \cdot h^{\frac{1}{2} v'' - 1} \exp \left[-\frac{1}{2} h v'' w'' \right] d\mu dh \dots \dots \dots (5.2.7)$$

Instead of the equation (5.2.7), (Wood and Rodriguez-Iturbe, 1975b) obtained

$$K_N = (n')^{\frac{1}{2}} (2\pi)^{-\frac{1}{2} v} \left[\left(\frac{1}{2} v' w' \right)^{\frac{1}{2} v'} \right] \left[\left(\frac{1}{2} v' \right) \right]^{-1} \int_u \int_h (2\pi)^{-\frac{1}{2}} h^{\frac{1}{2}} \exp \left[-\frac{1}{2} h n'' (\mu - m'')^2 \right] \cdot h^{\frac{1}{2} v'' - 1} \exp \left[-\frac{1}{2} h v'' w'' \right] d\mu dh$$

which was found to be in error.

The solution of the integral in equation (5.2.7) can be obtained by

integrating between the limits $\mu = 0$ and $\mu = \infty$; $h = 0$ and $h = \infty$

$$\int_0^{\infty} (2\pi)^{-\frac{1}{2}} h^{\frac{1}{2}} e^{-\frac{1}{2} h n'' (\mu - m'')^2} d\mu \cdot \int_0^{\infty} e^{-\frac{1}{2} h v'' w''} \cdot h^{\frac{1}{2} v'' - 1} dh.$$

The integral:

$$\int_0^{\infty} (2\pi)^{-\frac{1}{2}} \cdot h^{\frac{1}{2}} e^{-\frac{1}{2} h n'' (\mu - m'')^2} d\mu = \frac{1}{n''^{\frac{1}{2}}}$$

Since this is a probability density function.

The integral:

$$\int_0^{\infty} e^{-\frac{1}{2} h v'' w''} \cdot h^{\frac{1}{2} v'' - 1} dh$$

can be evaluated by making the following substitutions:

$$\text{Let } y = \frac{1}{2} v'' w'' h$$

$$dy = \frac{1}{2} v'' w'' dh$$

$$\begin{aligned} \int_0^{\infty} e^{-\frac{1}{2} h v'' w''} \cdot h^{\frac{1}{2} v'' - 1} dh &= \int_0^{\infty} e^{-y} \cdot \left[\frac{y}{\frac{1}{2} v'' w''} \right]^{\frac{1}{2} v'' - 1} \cdot \frac{dy}{\frac{1}{2} v'' w''} \\ &= \frac{1}{(\frac{1}{2} v'' w'')^{\frac{1}{2} v''}} \int_0^{\infty} e^{-y} \cdot y^{\frac{1}{2} v'' - 1} dy = \frac{\Gamma(\frac{1}{2} v'')}{(\frac{1}{2} v'' w'')^{\frac{1}{2} v''}} \end{aligned}$$

Thus the integral

$$\begin{aligned} \int_{\mu} \int_h (2\pi)^{-\frac{1}{2}} h^{\frac{1}{2}} \cdot \exp \left[-\frac{1}{2} h n'' (\mu - m'')^2 \right] \cdot h^{\frac{1}{2} v'' - 1} \exp \left[-\frac{1}{2} h v'' w'' \right] d\mu dh \\ = \frac{1}{n''^{\frac{1}{2}}} \cdot \frac{\Gamma(\frac{1}{2} v'')}{(\frac{1}{2} v'' w'')^{\frac{1}{2} v''}} \dots \dots \dots (5.2.8) \end{aligned}$$

Substituting equation (5.2.8) into equation (5.2.7) gives:

$$\begin{aligned} K_N &= (n')^{\frac{1}{2}} (2\pi)^{-\frac{1}{2}(v+1)} \frac{(\frac{1}{2} v' w')^{\frac{1}{2} v'}}{\Gamma(\frac{1}{2} v')} \cdot \frac{1}{(n'')^{\frac{1}{2}}} \cdot \frac{\Gamma(\frac{1}{2} v'')}{(\frac{1}{2} v'' w'')^{\frac{1}{2} v''}} \\ K_N &= \left[\frac{n'}{n''} \right]^{\frac{1}{2}} (2\pi)^{-\frac{1}{2}(v+1)} \frac{\Gamma(\frac{1}{2} v'')}{\Gamma(\frac{1}{2} v')} \cdot \frac{(\frac{1}{2} v' w')^{\frac{1}{2} v'}}{(\frac{1}{2} v'' w'')^{\frac{1}{2} v''}} \dots \dots \dots (5.2.9) \end{aligned}$$

where the parameters n' , n'' , v , v' , v'' , w'' are as defined in chapter 2.

5.2.3 Marginal Density Function For Log-Normal Model

In chapter 2, the probability density function for log-normal distribution was given by:

$$f(q | \mu, h) = \frac{1}{q} (2\pi)^{-\frac{1}{2}} h^{\frac{1}{2}} \exp \left[-\frac{1}{2} h (\ln q - \mu)^2 \right]$$

and the likelihood function for μ and h , given n independent observations of q , is

$$L(\mu, h|Q) = \left[\prod_{i=1}^n q_i \right]^{-1} (2\pi)^{-n/2} \cdot h^{n/2} \cdot \exp \left[-\frac{1}{2} h n \sum (\ln q_i - \mu)^2 \right]$$

Assuming a normal gamma prior for μ and h of the same family as for normal model, then the marginal density function for the log-normal model is found by applying equation (5.2.4)

Thus,

$$K_{LN} = \left[\prod_{i=1}^n q_i \right]^{-1} \int_{\mu} \int_h (2\pi)^{-n/2} \cdot h^{n/2} \cdot \exp \left[-\frac{1}{2} h n \sum (\ln q_i - \mu)^2 \right] \cdot f'(\mu, h) d\mu dh \quad \dots (5.2.10)$$

where :

$$f'(\mu, h) = f'_N(\mu | m, hm) \cdot f'_{\chi^2_2}(h | w', v')$$

The solution of the integral in equation (5.2.10) can be obtained by following the same procedure as for normal model.

Thus the marginal likelihood function K_{LN} is given by:

$$K_{LN} = \left[\prod_{i=1}^n q_i \right]^{-1} \left[\frac{n'}{n''} \right]^{\frac{1}{2}} (2\pi)^{-\frac{1}{2}(v+1)} \frac{\left(\frac{1}{2} v'' \right)}{\left(\frac{1}{2} v' \right)} \cdot \frac{\left(\frac{1}{2} v' w' \right)^{\frac{1}{2} v'}}{\left(\frac{1}{2} v'' w'' \right)^{\frac{1}{2} v''}} \quad \dots (5.2.11)$$

5.2.4 Derivation of Marginal Density Function For Exceedance Model

The probability density function for the exceedance model is given by equation (2.5.9).

That is:

$$f(q|v, \alpha) = v \alpha t \exp [-\alpha(q - q_b)] \quad \text{for } q \geq q_b \quad \dots (5.2.12)$$

$$f(q|v, \alpha) = (1 - v)/q_b \quad \text{for } 0 \leq q \leq q_b \quad \dots (5.2.13)$$

where v is the arrival rate of flood, α is the event magnitude parameter, and q_b is the base level discharge.

Supposing a sample contains $(n+m)$ independent discharges Q of which n is the number of discharges greater than the base level discharge, q_p , and m is the number of discharges less than q_p , then the likelihood function for v and α is derived as follows:

For the exceedance discharges:

$$F(z) = \exp \left[-vt e^{-\alpha z} \right] \quad \text{for } z \geq 0$$

Differentiating the above equation gives the probability density function $f(z)$ given by:

$$f(z) = \frac{dF(z)}{dz} = \frac{d}{dz} (\exp -vt e^{-\alpha z})$$

$$\text{Let } M = \exp \left[-vt e^{-\alpha z} \right]$$

Taking the natural logarithm of both sides

$$\ln[M] = -vt e^{-\alpha z}$$

$$\frac{dM}{dz} \cdot \frac{1}{M} = \alpha vt e^{-\alpha z}$$

$$\frac{dM}{dz} = M \cdot \alpha vt e^{-\alpha z}$$

$$f(z) = \frac{dF}{dz} = \exp \left[-vt e^{-\alpha z} \right] \cdot \alpha vt e^{-\alpha z}$$

$$f(z) = \exp \left[-vt e^{-\alpha z} \right] \cdot \alpha vt e^{-\alpha z}$$

The likelihood function

$$\begin{aligned} L[\alpha, v|q] &= \alpha vt_1 e^{-\alpha z_1} \exp \left[-vt_1 e^{-\alpha z_1} \right] \cdot \alpha vt_2 e^{-\alpha z_2} \exp \left[-vt_2 e^{-\alpha z_2} \right] \\ &\quad \dots \alpha vt_n e^{-\alpha z_n} \exp \left[-vt_n e^{-\alpha z_n} \right] \\ &= \alpha^n v^n \prod_{i=1}^n t_i e^{-\alpha(z_1+z_2+z_3+\dots+z_n)} \cdot e^{-v(t_1 e^{-\alpha z_1} + t_2 e^{-\alpha z_2} + \dots + t_n e^{-\alpha z_n})} \\ &= \alpha^n v^n \prod_{i=1}^n t_i \exp \left[-\left(\alpha \sum_{i=1}^n z_i + v \sum_{i=1}^n t_i \exp -\alpha z_i \right) \right] \end{aligned}$$

In practice $e^{-\alpha z_i} \rightarrow 1$
 $\prod_{i=1}^n t_i \rightarrow 1$

$$L[\alpha, v | (n, t)] = \alpha^n v^n e^{-\alpha \sum_{i=1}^n z_i} \cdot \exp \left[-v \sum_{i=1}^n t_i \right] \dots (5.2.14)$$

For the discharges below the base level discharge q_b

$$f(q | v, \alpha) = \frac{(1-v)}{q_b} \quad \text{for} \quad 0 < q < q_b$$

$$L[v, \alpha | q] = \left[\frac{1-v}{q_b} \right] \cdot \left[\frac{1-v}{q_b} \right] \dots$$

$$L(v, \alpha | q) = \left[\frac{1-v}{q_b} \right]^m$$

The likelihood for the total discharges Q is found from:

$$L(v, \alpha | Q) = L[v, \alpha | q : q < q_b] \cdot L[v, \alpha | q : q \geq q_b] \dots (5.2.15)$$

$$L(v, \alpha | Q) = \frac{(1-v)^m}{q_b^m} \cdot \alpha^n v^n \exp \left[-\alpha \sum_{i=1}^n z_i \right] \cdot \exp \left[-v \sum_{i=1}^n t_i \right] \dots (5.2.16)$$

The marginal likelihood function K_e is defined as:

$$K_e = \int_v \int_{\alpha} L(v, \alpha | Q) \cdot f'(v) \cdot f'(\alpha) \, dv \, d\alpha \dots (5.2.17)$$

The conjugate prior density functions for v and α are of the form: (Wood and Rodriguez-Iturbe, 1975b).

$$f'(v) = \exp(-s'v) \cdot v^{u'-1} \left[s' u' / \Gamma(u') \right] \dots (5.2.18)$$

$$f'(\alpha) = \exp(-l'\alpha) \cdot \alpha^{w'-1} \left[l' w' / \Gamma(w') \right] \dots (5.2.19)$$

Substituting equations (5.2.16), (5.2.18) and (5.2.19) into equation (5.2.17) gives:

$$K_e = \int_v \int_{\alpha} \frac{(1-v)^m}{q_b^m} \cdot \alpha^n v^n \exp \left[-\alpha \sum_{i=1}^n z_i \right] \cdot \exp \left[-v \sum_{i=1}^n t_i \right] \exp(-s'v) \cdot v^{u'-1}$$

$$\left[s' u' / \Gamma(u') \right] \cdot \exp(-l'\alpha) \cdot \alpha^{w'-1} \left[l' w' / \Gamma(w') \right] \, dv \, d\alpha$$

$$= a_b^{-m} \cdot \frac{s'^{u'}}{\Gamma(u')} \int_v \exp \left[- \left(s' + \sum_{i=1}^n t_i \right) v \right] \cdot v^{n+u'-1} (1-v)^m dv \cdot \frac{l'^{w'}}{\Gamma(w')} \\ \int_{\alpha} \exp \left[- \left(l' + \sum_{i=1}^n z_i \right) \alpha \right] \cdot \alpha^{n+w'-1} d\alpha \quad \dots\dots\dots(5.2.20)$$

Substituting the following relations which have been established in chapter 2 into equation (5.2.20)

$$l' + \sum_{i=1}^n z_i = l'' ; \quad u' + n = u'' ; \\ s' + \sum_{i=1}^n t_i = s'' ; \quad w' + n = w''$$

gives:

$$K_E = a_b^{-m} \cdot \frac{s'^{u'}}{\Gamma(u')} \int_v \exp(-s''v) \cdot v^{u''-1} (1-v)^m dv \cdot \frac{l'^{w'}}{\Gamma(w')} \int_{\alpha} \exp(-l''\alpha) \cdot \\ \alpha^{w''-1} d\alpha \quad \dots\dots\dots(5.2.21)$$

When v is much less or equal to 0.1, the expression:

$$(1-v)^m \cong e^{-mv} \quad [v \leq 0.1] \quad \dots\dots\dots(5.2.22)$$

This relation has been developed by the (Wood and Rodriguez-Iturbe, 1975b).

Substituting equation (5.2.22) into equation (5.2.21)

$$K_E = a_b^{-m} \cdot \frac{s'^{u'}}{\Gamma(u')} \int_v \exp \left[-(s'' + m)v \right] \cdot v^{u''-1} dv \cdot \frac{l'^{w'}}{\Gamma(w')} \int_{\alpha} \exp(-l''\alpha) \cdot \\ \alpha^{w''-1} d\alpha \quad \dots\dots\dots(5.2.23)$$

The integral over v can be obtained by making the following substitution:

$$y = (s'' + m)v \\ dy = (s'' + m)dv.$$

and integrating between the limits $v = 0$ and $v = \infty$

$$\begin{aligned}
\int_0^{\infty} \exp [-(s''+m)v] \cdot v^{u''-1} dv &= \int_0^{\infty} \exp (-y) \cdot \left[\frac{y}{s''+m} \right]^{u''-1} \cdot \frac{dy}{(s''+m)} \\
&= \frac{1}{(s''+m)^{u''}} \int_0^{\infty} \exp (-y) \cdot y^{u''-1} dy = \frac{\Gamma(u'')}{(s''+m)^{u''}} \dots\dots\dots(5.2.24)
\end{aligned}$$

The integral over α can be obtained by making substitution similar to that of v and integrating between the limits $\alpha = 0$ and $\alpha = \infty$.

$$\begin{aligned}
y &= l''\alpha \\
dy &= l'' d\alpha
\end{aligned}$$

$$\begin{aligned}
\int_0^{\infty} \exp (-l''\alpha) \cdot \alpha^{w''-1} d\alpha &= \int_0^{\infty} \exp (-y) \cdot \left[\frac{y}{l''} \right]^{w''-1} \cdot \frac{dy}{l''} \\
&= \frac{1}{(l'')^{w''}} \int_0^{\infty} \exp (-y) \cdot y^{w''-1} dy = \frac{\Gamma(w'')}{(l'')^{w''}} \dots\dots\dots(5.2.25)
\end{aligned}$$

Substituting equations (5.2.24) and (5.2.25) into equation (5.2.23) gives:

$$\begin{aligned}
K_E &= q_b^{-m} \cdot \frac{s' u'}{\Gamma(u')} \cdot \frac{\Gamma(u'')}{(s''+m)^{u''}} \cdot \frac{l' w'}{\Gamma(w')} \cdot \frac{\Gamma(w'')}{(l'')^{w''}} \\
K_E &= q_b^{-m} \cdot \frac{s' u'}{(s''+m)^{u''}} \cdot \frac{\Gamma(u'')}{\Gamma(u')} \cdot \frac{l' w'}{l''^{w''}} \cdot \frac{\Gamma(w'')}{\Gamma(w')} \dots\dots(5.2.26)
\end{aligned}$$

5.3

APPLICATION

The analytical derivations for the model uncertainty presented in sections 5.1 and 5.2 were applied to the River Taraba at Gassol in the upper Benue River Basin in Nigeria.

5.3.1 Calculation of Marginal Likelihood Function, K_N , for the Normal Model:

The following parameters have been calculated in chapter 4 as:

$$n' = 0, \quad v = 33, \quad v' = 7, \quad v'' = 41, \quad n'' = 34, \quad w' = 1.04 \times 10^5 \text{ (m}^6/\text{s}^2\text{)}, \text{ and } w'' = 94348 \text{ (m}^6/\text{s}^2\text{)}.$$

Substituting the values of these parameters into equation (5.2.9) gives:

$$K_N = \left[\frac{0}{34} \right]^{\frac{1}{2}} (2\pi)^{-\frac{1}{2}(33+1)} \frac{\Gamma(\frac{1}{2} \times 41)}{\Gamma(\frac{1}{2} \times 7)} \cdot \frac{(\frac{1}{2} \times 7 \times 1.04 \times 10^5)^{\frac{1}{2} \times 7}}{(\frac{1}{2} \times 41 \times 94348)^{\frac{1}{2} \times 41}} = 0$$

$$\text{Thus } K_N = 0$$

5.3.2 Calculation of Marginal Likelihood Function, K_{LN} , for The Log-normal Model:

The following parameters have been calculated in chapter 4 as:

$$n' = 0, \quad v' = 7, \quad v = 33, \quad w' = 0.052, \quad v'' = 41, \quad n'' = 34, \\ \text{and } w'' = 0.0345 \text{ (ln m}^6/\text{s}^2\text{)}$$

Substituting the values of these parameters into equation (5.2.11) gives:

$$K_{LN} = \left[\prod_{i=1}^{34} q_i \right]^{-1} \cdot \left[\frac{0}{34} \right]^{\frac{1}{2}} (2\pi)^{-\frac{1}{2}(33+1)} \frac{\Gamma(\frac{1}{2} \times 41)}{\Gamma(\frac{1}{2} \times 7)} \cdot \frac{(\frac{1}{2} \times 7 \times 0.052)^{\frac{1}{2} \times 7}}{(\frac{1}{2} \times 41 \times 0.0345)^{\frac{1}{2} \times 41}}$$

$$K_{LN} = 0.$$

5.3.3 Calculation of Marginal Likelihood Function, K_E , for The Exceedance Model:

The following parameters have been calculated in chapter 4 as:

$$w' = 3, \quad w'' = 6, \quad l' = 1471.5 \text{ (m}^3/\text{s)}, \quad l'' = 2051 \text{ (m}^3/\text{s)}, \\ n = 3, \quad m = 34 - 3 = 31, \quad \sum z_i = 580 \text{ (m}^3/\text{s)}, \quad s' = 56, \quad u' = 5 \\ s'' = 90, \quad u'' = 8$$

Substituting the values of these parameters into equation

(5.2.25) gives:

$$\begin{aligned}
 K_E &= a_b^{-m} \frac{s^{u'}}{(s'' + m)^{u''}} \cdot \frac{\Gamma(u'')}{\Gamma(u')} \cdot \frac{1^{w'}}{1^{w''}} \cdot \frac{\Gamma(w'')}{\Gamma(w')} \\
 &= (2200)^{-31} \cdot \frac{(56)^5}{(90+31)^8} \cdot \frac{\Gamma(8)}{\Gamma(5)} \cdot \frac{(1471.5)^3}{(2051.5)^6} \cdot \frac{\Gamma(6)}{\Gamma(3)} \\
 &= (2200)^{-31} \cdot \frac{(56)^5}{(121)^8} \cdot \frac{7!}{4!} \cdot \frac{(1471.5)^3}{(2051.5)^6} \cdot \frac{5!}{2!} \\
 &= 2.4260 \times 10^{-42} \cdot x \cdot 10^{-62} \cdot \frac{5.5073 \times 10^8}{4.5950 \times 10^{16}} \cdot \frac{3.18626 \times 10^9}{7.4547 \times 10^{19}} \cdot 210 \times 60 \\
 &= 15659.0 \times 10^{-122} \\
 K_E &= 1.5659 \times 10^{-18}
 \end{aligned}$$

5.3.4 Calculation of Normalizing Constant, K*

The normalizing constant, K* is given by the equation (5.1.7)

$$K^* = K_N P'(\theta_N = 1) + K_{LN} P'(\theta_{LN} = 1) + K_E P'(\theta_E = 1)$$

From the above calculations, $K_N = K_{LN} = 0$, and $K_E = 1.5659 \times 10^{-118}$

Also, since the three models are assumed to be competing model, it is reasonable to assume that their prior probabilities are equal.

Thus we can assume that:

$$P'(\theta_N = 1) = P'(\theta_{LN} = 1) = P'(\theta_E = 1) = 0.33$$

Substituting the values of the prior probabilities and K_N , K_{LN} and K_E into the expression for the normalizing constant:

$$\begin{aligned} K^* &= 0 \times 0.33 + 0.0.33 + 1.5659 \times 10^{-118} \times 0.33 \\ &= 5.16747 \times 10^{-119} \end{aligned}$$

$$\therefore K^* = 5.16747 \times 10^{-119}$$

5.3.5 Calculation Of Posterior Model Probabilities, $P''(\theta_i)$

The posterior model probabilities, $P''(\theta_i)$ were calculated by applying equations (5.1.13), (5.1.14) and (5.1.15)

(a) For Normal Model:

$$\begin{aligned} P''(\theta_N = 1) &= \frac{K_N}{K^*} \cdot P'(\theta_N = 1) \\ &= \frac{0}{5.16747 \times 10^{-119}} \times 0.33 \\ P''(\theta_N = 1) &= 0 \end{aligned}$$

(b) For Log-Normal model:

$$\begin{aligned} P''(\theta_{LN} = 1) &= \frac{K_{LN}}{K^*} \cdot P'(\theta_{LN} = 1) \\ &= \frac{0}{5.16747 \times 10^{-119}} \times 0.33 = 0 \\ P''(\theta_{LN} = 1) &= 0 \end{aligned}$$

(c) For Exceedance model:

$$\begin{aligned} P''(\theta_E = 1) &= \frac{K_E}{K^*} \cdot P'(\theta_E = 1) \\ &= \frac{1.5659 \times 10^{-118}}{5.16747 \times 10^{-119}} \times 0.33 \\ P''(\theta_E = 1) &= 1 \end{aligned}$$

5.3.6 Calculation of the Composite Bayesian Model:

From equation (5.1.19), the composite model is given by:

$$\tilde{f}(q) = P''(\theta_N = 1) \cdot \tilde{f}_N(q) + P''(\theta_{LN} = 1) \cdot \tilde{f}_{LN}(q) + P''(\theta_E = 1) \cdot \tilde{f}_E(q)$$

Substituting the values of the posterior model probabilities calculated above into the above equation

$$\tilde{f}(q) = 0 \cdot \tilde{f}_N(q) + 0 \cdot \tilde{f}_{LN}(q) + 1 \cdot \tilde{f}_E(q)$$

$$\tilde{f}(q) = \tilde{f}_E(q)$$

where $\tilde{f}_E(q)$ is the Exceedance model free from parameter uncertainty given in chapter 4.

That is:

$$\tilde{f}_E(q) = 0.089t \left[1 + \frac{z}{2051.5} \right]^{-7}$$

Thus from the above analysis it can be seen that the exceedance model is the best model that can be used to predict the flood discharges above a base level discharge of 2200 m³/s for the Gassol gaging station records of River Taraba in the Upper Benue River Basin in Nigeria. This exceedance model reduces both parameter uncertainty and model uncertainty.

CHAPTER 6DISCUSSIONS6.1 COMPARISON OF FLOOD DISCHARGES PREDICTED BY CLASSICAL,
BAYESIAN AND REGIONAL METHODS

In the design of a hydraulic structure for example a dam, the life span of the structure dictates the return period to be used for the design. Corresponding to this return period is a design discharge to be chosen. The choice of the exact discharge which will not be exceeded is the central problem of hydrologists.

The studies have been carried out with the aim of forecasting the flood discharges at Gassol gaging station in Upper Benue River Basin for different return periods greater than ten years. This is because most devastating floods have return periods usually greater than ten years. The historical record available at Gassol is continuous and it extends from 1955 to 1988.

To achieve the objective, three different methods of analysis have been tried. These are the Classical, Bayesian and Regional methods of analysis. The classical method has utilised only the historical record of 34 years at the Gassol gaging station. The Regional analysis was based on the historical records of six other gaging stations (from 1955 to 1988) in the same homogeneous area as Gassol. The Bayesian method of analysis utilised both the historical record at Gassol and six Regional records.

Table 6.1 Comparison of Classical, Bayesian and Regional discharges with observed discharges.

Return Period (Years)	Observed Discharges (m^3/s)	Classical Discharges (m^3/s)	Bayesian Discharges (m^3/s)	Regional Discharges (m^3/s)	Differences		
(1)	(2)	(3)	(4)	(5)	(3)-(2)	(4)-(2)	(5)-(2)
11.67	2280	2205	2211	2284	- 75	- 69	+ 4
17.50	2400	2283	2334	2450	- 117	- 66	+ 50
35	2500	2417	2561	2775	- 83	+ 61	+ 275
100		2625	2960	3400	-	-	-
1000		3075	4035	5435	-	-	-

The Table 6.1 shows the comparison of flood discharges predicted by the three methods (using the exceedance model, Figure 4.3) with the observed flood discharges. The observed discharges shown in column (2) were obtained from the historical record of Gassol gaging station with discharges having return periods greater than ten years. The observed discharge record had been cross checked for consistency using the double mass curve method and is therefore, used as the standard. The observed discharges corresponding to return periods of 100 and 1000 years are unknown and are to be predicted using the three methods of analysis.

When the results from the classical method are compared with the observed values, it is found that the discharges predicted by the classical method are consistently lower than the observed values. When the classical method was used to predict discharges corresponding to high return periods, three types of uncertainties might have existed: (i) the uncertainty inherent in the natural generation of discharges due to the randomness of the streamflow (ii) Parameter

uncertainty associated with the estimation of the model parameters from a short record of 34 years and (iii) model uncertainty which is the uncertainty that a particular model may be the true or best model. The classical analysis made no provision for these uncertainties and therefore the method is unreliable when used to predict discharges of return periods much longer than the length of years of record available.

When the results from Bayesian method of analysis are compared with the observed values, the general trend is that, between return periods of 10 and 25 years, the Bayesian method gives lower flood discharges than the observed values whereas for return periods exceeding 25 years, the Bayesian method is found to give higher flood discharge values. For example, for return periods of 11.67 and 17.5 years, the observed discharges at Gassol were $2280 \text{ m}^3/\text{s}$ and $2400 \text{ m}^3/\text{s}$ respectively whereas the Bayesian predicted values were $2210 \text{ m}^3/\text{s}$ and $2335 \text{ m}^3/\text{s}$ respectively. The differences between the observed discharges and those predicted by Bayesian methods are respectively -3% and -2.75% for the above return periods. For the return period of 35 years, the observed discharge at Gassol station was $2500 \text{ m}^3/\text{s}$ while the predicted value from the Bayesian method was $2560 \text{ m}^3/\text{s}$, representing a difference of $+2.4\%$. In the Bayesian method of analysis efforts had been made to reduce the uncertainties. For example, in order to reduce the parameter uncertainty the posterior parameters (made up of the combination of both sample and prior parameters) have been used. To reduce the model uncertainty, a composite model composed of the three distributions was used and this resulted in exceedance model being the best model out of those considered.

Comparison of the results from Regional method with the observed values indicates that the Regional method consistently predicts higher discharges than the observed values. For return periods between 10 and 25 years, the observed values are found to be between the Bayesian values and Regional values but much closer to the regional values. The regional method is found to give a more reliable result than either the Bayesian or classical method within this range of return period. For example, for return periods of 11.67 and 17.5 years, the Regional method predicts discharge of $2285 \text{ m}^3/\text{s}$ and $2450 \text{ m}^2/\text{s}$ respectively as against the observed values of $2280 \text{ m}^3/\text{s}$ and $2400 \text{ m}^3/\text{s}$ respectively. These represent the difference of +0.18% and +2.1% respectively.

For return periods greater than about 25 years, the results predicted by the Regional method was found to be much higher than the observed values. For example, for return period of 35 years, the regional method predicts flood discharge of $2775 \text{ m}^3/\text{s}$ as against the observed value of $2500 \text{ m}^3/\text{s}$. For return periods exceeding 25 years, the observed values are between the classical and Bayesian values, but much closer to the Bayesian values.

6.2 SELECTION OF MODEL THAT FITS THE FLOOD DATA AT GASSOL

In order to select the model that fits the flood data at Gassol three distributions (Normal, log-normal and exceedance) were fitted to the flood data of River Taraba at Gassol gaging station and attempts made to estimate flood discharges with return periods above ten years. For each distribution three methods of analysis were used (Classical, Bayesian and Regional analysis). The frequency curves obtained from the three distributions

are shown in Figures 4.1, 4.2 and 4.3 respectively for normal, log normal and exceedance distributions.

From Figure 4.1, it is found that when the normal model is used the classical and Bayesian methods of analysis predict values of discharges lower than the corresponding observed values for a given return period. For example, for return period of 35 years, the values predicted by the classical and Bayesian methods were $2350\text{m}^3/\text{s}$ and $2390\text{m}^3/\text{s}$ against the actual value of $2500\text{m}^3/\text{s}$. This shows clearly that the normal model is unsuitable for predicting flood discharges above a return period of ten years at Gassol. This ascertainment is confirmed by the facts that the posterior probability of this normal model was found to be zero.

From Figure 4.2, it is observed that the classical method of analysis using log-normal model consistently predicts lower discharges than the observed values. Though, in the case of Bayesian method of analysis using log-normal model, the predicted discharge was found to be close to the observed value at a return period of 35 years, the posterior probability of this model was found to be zero showing that the model is unsuitable.

In the case of the Exceedance model, Figure 4.3, the observed values were found to lie between the values predicted by the regional, Bayesian and classical methods. Also the posterior probability of exceedance model was unity showing that it is the best out of the three models considered.

6.3 SAMPLE, PRIOR AND POSTERIOR MEAN FLOOD DISCHARGES FOR NORMAL MODEL

The sample mean flood discharges for normal model calculated from the historical record of River Taraba at Gassol was $1769.6 \text{ m}^3/\text{s}$. The prior mean discharge obtained from the historical records of eleven other stations for normal model was $1766.6 \text{ m}^3/\text{s}$. The posterior mean flood discharge calculated by combining the sample and prior parameters was found to be $1769.6 \text{ m}^3/\text{s}$. This means the three methods yield close results.

6.4 REGIONAL FREQUENCY CURVES FOR NORMAL AND LOG-NORMAL MODELS

The regional frequency curves for both Normal and Log-normal models were found to be absent in Figures 4.1 and 4.2 respectively. This is because the prior parameter n' (which represents the equivalent length of years of regional record) was found to be negligible. For normal and log-normal models, the values of n' obtained were 0.11 and 0.16 respectively, which according to (Wood and Rodriguez-Iturbe, 1975a) should be regarded as zero.

The value n' was calculated from the ratio of the expectation of the variance of the regional discharges to the variance of the mean regional discharges. That is

$$n' = \frac{E \left[\sigma_Q^2 \mid I_R \right]}{V \left[\mu_Q \mid I_R \right]}$$

CHAPTER 7CONCLUSIONS/RECOMMENDATIONS7.1 CONCLUSIONS

The following conclusions have been drawn from the studies:

1. For the prediction of flood discharges above a return period of ten years of the River Taraba at Gassol gaging station, the exceedance model was found to give the most satisfactory results out of the three models (Normal, log-normal and exceedance) considered.
2. The classical method of flood frequency analysis was found to be unsuitable when used to predict discharges above a return period of ten years. The Regional method of flood frequency analysis gave the most satisfactory and reliable results for River Taraba at Gassol for return periods between 10 and 25 years. For return periods above 25 years, the Bayesian method gave the most satisfactory discharges for River Taraba at Gassol.
3. A single probability distribution may not fit all the flood data at a station. At different portions of the flood data, different probability distributions may be required. Bayesian method of analysis takes care of this through the use of composite probability model which consists of different distributions. This means that the posterior probabilities of these distributions will add up to unity.

4. For the normal distribution, it was found that for the example considered, the sample, prior and posterior mean flood discharges were very close. This shows that for the Normal distribution, the classical, Regional and Bayesian methods yield nearly the same result for the mean discharge in the example considered.

5. The only demerit in the use of the Bayesian method of flood frequency is that it involves vigorous mathematical derivations and sometimes the solutions of the integrals involved are difficult to obtain. However, when these derivations have been carried out, the application is very simple.

7.2 RECOMMENDATIONS

1. It is recommended that the authorities of the Upper Benue River Basin Development should in future use the Regional method of flood frequency analysis to predict discharges having return periods between 10 and 25 years for the Gassol gaging station. For the prediction of flood discharges above a return period of 25 years, the Bayesian method of flood frequency analysis is recommended. This is because the Bayesian method has the advantage of including the parameter uncertainty in the analysis explicitly and by using prior information this uncertainty is greatly reduced.

2. It is recommended that Bayesian analytical derivations similar to the Normal, Log-normal and exceedance distributions be carried out for other common probability distributions to enhance the future use of the Bayesian method of flood frequency analysis.

REFERENCES

Agriculture Livestock and Technical Services (ALTS) Nig. Ltd.,

LIDECO and H. Y. Bologna Italy (1981). Feasibility Studies For the Lower Taraba Basin, Dam Engineering Report, Annex 5 to the General Report, Upper Benue River Basin, Nigeria.

Agriculture Livestock and Technical Services (ALTS) Nig. Ltd.,

LIDECO Italy (1983)., Feasibility Studies for the Lower Taraba Basin, Hydrology Report, Annex 2, Upper Benue River Basin, Nigeria, pp. 68.

Beard, L. R. (1962). Statistical Methods in Hydrology, In: Dune, T. and Leopold, L. B. (Ed.). Water in Environmental Planning, W. H. Freeman and Company, San Francisco pp. 380 - 435.

Benjamin, J. R. and Cornell, C. A. (1970). Probability Statistics, and Decision for Civil Engineers, McGraw-Hill, U. S. A.

Chow, Ven Te (1964). Handbook of Applied Hydrology, McGraw-Hill, New York.

Cunnane, C. and Nash, J. E. (1973). Bayesian Estimation of Frequency of Hydrological Events, Paper Presented at Symposium on the Design of Water Resources Projects with Inadequate Data, Madrid, pp. 48 - 55.

Dalrymple Tate (1960). Flood - Frequency Analysis, U. S. Geological Survey Water Supply Paper 1543 - A pp. 1 - 46.

- Davis, D. R., Duckstein, L. and Fogel M. M. (1976).
Uncertainty in The Return Period of Maximum Hydrologic Events: A Bayesian Approach, Journal of Hydrology
 No. 31 pp 81 - 95.
- Langbein W. B.(1949). Annual Floods and Partial - Duration Flood Series, American Geophysical Union Transaction,
 Vol. 30, pp 879 - 881.
- Odusote, A. O.(1989). Regional Flood - Frequency Analysis of Ogun-Oshun River Basins, M.Sc. Thesis, Department of Water Resources and Environmental Engineering, A.B.U. Zaria, Nigeria.
- Ogbuli, C. M.(1989). Regional Flood Frequency Analysis of Upper Benue River Basin, Uncompleted M.Sc. Thesis, Department of Water Resources and Environmental Engineering, A.B.U. Zaria.
- Raiffa, H. and Schlaifer, R.(1961). Applied Statistical Decision Theory, M. I. T. Press, Cambridge, Mass., U.S.A.
- Shane and Graver (1970). Statistical Decision Theory Techniques for the Revision of Mean Flow Regression Estimates, Water Resources Research, Vol. 6, No. 6. pp 1641 - 1648.
- Todorovic, P. and Zelenhasic, E. (1970). A Stochastic Model for Flood Analysis, Water Resources Research, Vol. 6, No. 6, pp 1641 - 1648.
- Valde's, J. B., Rodriguez-Iturbe, I. and Vicens, G. J. (1977)
Bayesian Generation of Synthetic Stream-Flows 2. Multivariate Case, Water Resources Research, Vol. 13, No. 2, pp 291 - 295.
- Vicens, G. J., Rodriguez-Iturbe I. and Schaake JR., J. C. (1975)
Bayesian Framework for the Use of Regional Information in Hydrology, Water Resources Research, Vol. 11, No. 3, pp 405 - 414.

Vicens, G. J., Rodriguez-Iturbe, I. and Schaake JR., J. C. (1975)
Bayesian Generation of Synthetic Stream-Flows,
 Water Resources Research, Vol. 11, No. 6, pp. 827 - 838.

Viessman, W. JR., Knapp, J., Lewis, G. L. and Harbaugh, T. E. (1977).
Introduction to Hydrology, Harper and Row, Publishers,
 New York. pp. 201 - 228.

Wood, E. F., (1973). Flood Control Design with Limited Hydrologic
 Data - A Comparison of the Classical and Bayesian
 Approaches, Paper Presented at Symposium on the Design of
 Water Resources Projects with Inadequate Data, Madrid.
 pp. 469 - 483.

Wood, E. F., Rodriguez-Iturbe, I. (1975a). Bayesian Inference and
 Decision Making for Extreme Hydrologic Events, Water
 Resources Research, Vol. 11, No. 4, pp. 533 - 542.

Wood, E. F., Rodriguez-Iturbe, I., (1975b). A Bayesian Approach to
 Analyzing Uncertainty Among Flood Frequency Models,
 Water Resources, Vol. 11, No. 6, pp. 839 - 843.

Wood, E.F., Rodriguez-Iturbe, I., and Schaake, J.C., (1974). The
 Methodology of Bayesian inference and decision making
 applied to extreme hydrologic events, Rep. 178,
 Ralph M. Parsons Lab. for Water Resources and Hydrodyn.,
 Mass. Inst. of Technology, Cambridge.

A - 1

FLOOD DATA FOR PEAK DISCHARGE

EXCEEDANCE PROBABILITY ANALYSIS FOR STATION 1

Flood Data for River..... MAYO KAM AT MAYO KAM

Period of Record..... N = 34 Years; Drainage Area..... 2,986 Km²

DATE	ANNUAL FLOOD Q (m ³ /s)	REARRANGED VARIABLE IN DESCENDING ORDER	ORDER NUMBER M	RECURRENCE INTERVAL, $T_r = \frac{N+1}{M}$	PROBABILITY OF EXCEEDANCE $P_{T_r} = \frac{100\%}{T_r}$
(1)	(2)	(3)	(4)	(5)	(6)
1955	(400)	580	1	35.00	2.86
1956	(420)		2		
1957	(410)	504	3	11.67	8.57
1958	(330)		4		
1959	(140)		5		
1960	(390)		6		
1961	(340)		7		
1962	(300)		8		
1963	(500)	410	9	3.88	25.77
1964	(320)		10		
1965	(390)	400	11	3.18	31.45
1966	(350)		12		
1967	(340)		13		
1968	(480)	380	14	2.50	40.00
1969	(540)		15		
1970	(420)		16		
1971	(170)	350	17	2.06	48.54
1972	(320)		18		
1973	(370)		19		
1974	267		20		
1975	254		21		
1976	350		22		
1977	400		23		
1978	410		24		
1979	300	300	25	1.40	71.43
1980	504		26		
1981	580		27		
1982	380	267	28	1.25	80
1983	200	254	29	1.21	82.64
1984	150	200	30	1.17	85.47
1985	(325)		31		
1986	(380)		32		
1987	(290)	150	33	1.06	94.34
1988	(200)		34		

After Ogtuli

APPENDIX A.

A - ?

FLOOD DATA FOR PEAK DISCHARGE

EXCEEDANCE PROBABILITY ANALYSIS FOR STATION 2

Flood Data for River... TARABA AT BELI

Period of Record..... N = 34 Years; Drainage Area..... 10,850 Km²

DATE	ANNUAL FLOOD Q (m ³ /s)	REARRANGED VARIABLE IN DESCENDING ORDER	ORDER NUMBER M	RECURRENCE INTERVAL, $T_r = \frac{N+1}{M}$	PROBABILITY OF EXCEEDANCE $P_{T_r} = \frac{100\%}{T_r}$
(1)	(2)	(3)	(4)	(5)	(6)
1955	2322	2830	1	35.00	2.86
1956	2382	2700	2	17.50	5.71
1957	2180	2550	3	11.67	8.57
1958	1612	2467	4	8.75	11.43
1959	1100	2426	5	7.00	14.29
1960	2426	2382	6	5.83	17.14
1961	2125	2350	7	5.00	20.00
1962	1642	2322	8	4.38	22.86
1963	2830	2220	9	3.88	25.71
1964	2220	2200	10	3.50	28.50
1965	2467	2180	11	3.18	31.43
1966	(1940)	2150	12	2.92	34.29
1967	1860	2125	13	2.69	37.14
1968	2550	2100	14	2.50	40.00
1969	2700	2095	15	2.33	42.86
1970	2000	2010	16	2.19	45.71
1971	1372	2000	17	2.06	48.57
1972	2100	1957	18	1.94	51.43
1973	2150		19		
1974	2010	1897	20	1.73	57.14
1975	1798	1860	21	1.67	60.00
1976	1789	1798	22	1.59	62.86
1977	1737	1789	23	1.52	65.71
1978	1500	1767	24	1.46	68.57
1979	1957	1737	25	1.40	71.43
1980	2350		26		
1981	1767	1642	27	1.30	77.14
1982	1897	1612	28	1.25	80.00
1983	1415	1500	29	1.21	82.86
1984	1398	1415	30	1.17	85.71
1985	2095	1398	31	1.13	88.57
1986	2200	1372	32	1.09	91.43
1987	(1650)		33		
1988	(1340)	1100	34	1.03	97.14

After Ogtuli

A - 3

FLOOD DATA FOR PEAK DISCHARGE

EXCEEDANCE PROBABILITY ANALYSIS FOR STATION 3

Flood Data for River..... TARABA AT TELA

Period of Record..... N = 34 Years; Drainage Area..... 19,979 Km²

DATE	ANNUAL FLOOD Q (m ³ /s)	REARRANGED VARIABLE IN DESCENDING ORDER	ORDER NUMBER M	RECURRENCE INTERVAL, $T_r = \frac{N+1}{M}$	PROBABILITY OF EXCEEDANCE $P_{T_r} = \frac{100\%}{T_r}$
(1)	(2)	(3)	(4)	(5)	(6)
1955	(2550)		1		
1956	(2750)	3200	2	17.50	5.71
1957	(2700)		3		
1958	(2140)		4		
1959	(1040)	3000	5	7.00	14.29
1960	(2480)	2800	6	5.83	17.14
1961	(2200)		7		
1962	(2000)		8		
1963	(3200)		9		
1964	(2100)	2600	10	3.50	28.57
1965	(2450)		11		
1966	(2300)		12		
1967	(2200)		13		
1968	(3050)		14		
1969	(3400)		15		
1970	(2700)		16		
1971	(1200)		17		
1972	(2100)	2200	18	1.94	51.43
1973	(2140)		19		
1974	(2200)		20		
1975	(1600)		21		
1976	(1950)		22		
1977	(2080)		23		
1978	(1466)		24		
1979	2200		25		
1980	2800		26		
1981	3000		27		
1982	3200		28		
1983	1700	1700	29	1.21	82.86
1984	(2600)		30		
1985	(2100)	1466	31	1.13	88.57
1986	(2500)		32		
1977	(1900)		33		
1988	(1350)		34		

After Orbuli

APPENDIX A.

A - 4 FLOOD DATA FOR PEAK DISCHARGE

EXCEEDANCE PROBABILITY ANALYSIS FOR STATION 4

Flood Data for River..... BANTAJI AT SUNTAI

Period of Record..... N = 34 Years; Drainage Area..... 5.815 Km²

DATE	ANNUAL FLOOD Q (m ³ /s)	REARRANGED VARIABLE IN DESCENDING ORDER	ORDER NUMBER M	RECURRENCE INTERVAL, $T_r = \frac{N+1}{M}$	PROBABILITY OF EXCEEDANCE $P_{T_r} = \frac{100}{T_r} \%$
(1)	(2)	(3)	(4)	(5)	(6)
1955	(710)	960	1	35.00	2.86
1956	(750)	950	2	17.5	5.71
1957	(740)	910	3	11.67	8.57
1958	(645)	900	4	8.75	11.43
1959	(450)	856	5	7.00	14.29
1960	(463)		6		
1961	450	800	7	5.00	20.00
1962	750		8		
1963	(850)	750	9	3.88	25.71
1964	900	745	10	3.50	28.51
1965	508		11		
1966	(670)		12		
1967	632		13		
1968	950		14		
1969	910	646	15	2.33	42.86
1970	619		16		
1971	600		17		
1972	630		18		
1973	(640)	632	19	1.84	54.29
1974	552		20		
1975	646	619	21	1.67	60.00
1976	856	605	22	1.59	62.86
1977	378	600	23	1.52	65.71
1978	407		24		
1979	800	562	25	1.40	71.43
1980	745	530	26	1.35	74.29
1981	960	508	27	1.30	77.14
1982	530		28		
1983	420	463	29	1.21	82.86
1984	600		30		
1985	(635)	450	31	1.13	88.57
1986	(700)	420	32	1.09	91.43
1987	(580)	407	33	1.06	93.29
1988	(470)	378	34	1.03	97.14

After Orbuli

APPENDIX A.

A - 5 FLOOD DATA FOR PEAK DISCHARGE

EXCEEDANCE PROBABILITY ANALYSIS FOR STATION - 5

Flood Data for River..... BANTAJI AT SUNTAI BAKUNDI

Period of Record..... N = 34 Years; Drainage Area..... 3,350 Km²

DATE	ANNUAL FLOOD Q (m ³ /s)	REARRANGED VARIABLE IN DESCENDING ORDER	ORDER NUMBER M	RECURRENCE INTERVAL, $T_r = \frac{N+1}{M}$	PROBABILITY OF EXCEEDANCE $P_{T_r} = \frac{100\%}{T_r}$
(1)	(2)	(3)	(4)	(5)	(6)
1955	(740)	900	1	35.00	2.86
1956	(780)		2		
1957	(750)		3		
1958	(650)		4		
1959	(450)	800	5	7.00	14.29
1960	(710)	780	6	5.83	17.14
1961	(660)		7		
1962	(630)		8		
1963	(850)		9		
1964	(640)		10		
1965	(710)		11		
1966	(680)		12		
1967	(660)		13		
1968	(800)		14		
1969	(870)		15		
1970	(750)		16		
1971	(480)		17		
1972	(640)		18		
1973	(650)		19		
1974	(660)		20		
1975	(530)		21		
1976	(620)		22		
1977	(540)		23		
1978	(560)		24		
1979	(780)	620	25	1.40	71.43
1980	800		26		
1981	900	590	27	1.30	77.14
1982	620	560	28	1.25	80.00
1983	(550)		29		
1984	590	540	30	1.17	85.71
1985	(645)		31		
1986	(700)		32		
1987	(610)		33		
1988	(500)		34		

After Ogntuli

A - 6

FLOOD DATA FOR PEAK DISCHARGE

EXCEEDANCE PROBABILITY ANALYSIS FOR STATION - 6

Flood Data for River.....DONGA AT GINTIN DORROWA.....

Period of Record.....N = 34..... Years; Drainage Area.....20,390 Km².....

DATE	ANNUAL FLOOD Q (m ³ /s)	REARRANGED VARIABLE IN DESCENDING ORDER	ORDER NUMBER M	RECURRENCE INTERVAL, $T_r = \frac{N+1}{M}$	PROBABILITY OF EXCEEDANCE $P_{T_r} = \frac{100\%}{T_r}$
(1)	(2)	(3)	(4)	(5)	(6)
1955	(2040)	2500	1	35.00	2.86
1956	(2150)		2		
1957	(2100)		3		
1958	(1810)		4		
1959	(1220)	2170	5	7.00	14.29
1960	(2000)		6		
1961	(1840)		7		
1962	(1740)		8		
1963	(2380)		9		
1964	(1780)		10		
1965	(2000)	2000	11	3.18	31.43
1966	(1880)		12		
1967	(1840)		13		
1968	(2300)		14		
1969	(2500)		15		
1970	(2100)		16		
1971	(1280)		17		
1972	(1760)		18		
1973	(1800)		19		
1974	(1850)	1800	20	1.75	51.14
1975	(1500)	1790	21	1.67	60.00
1976	(1700)		22		
1977	(1760)		23		
1978	(1750)		24		
1979	(1790)		25		
1980	2170	1750	26	1.35	74.29
1981	2500		27		
1982	2000		28		
1983	1550		29		
1984	1500	1550	30	1.17	85.71
1985	(1780)		31		
1986	(2000)		32		
1987	(1660)		33		
1988	(1380)		34		

After Orbuli

A - 7 FLOOD DATA FOR PEAK DISCHARGE

EXCEEDANCE PROBABILITY ANALYSIS FOR STATION - 7

Flood Data for River..... DONGA AT DONGA

Period of Record..... N = 34 Years; Drainage Area..... 11,909 Km²

DATE	ANNUAL FLOOD Q (m ³ /s)	REARRANGED VARIABLE IN DESCENDING ORDER	ORDER NUMBER M	RECURRENCE INTERVAL, $T_r = \frac{N+1}{M}$	PROBABILITY OF EXCEEDANCE $P_{T_r} = \frac{100\%}{T_r}$
(1)	(2)	(3)	(4)	(5)	(6)
1955	1776	2400	1	35.00	2.86
1956	2047	2272	2	17.50	5.71
1957	1800	2150	3	11.67	8.57
1958	1632	2100	4	8.75	11.43
1959	1250	2047	5	7.00	14.29
1960	1890	2010	6	5.83	17.14
1961	1645	1930	7	5.00	20.00
1962	1548	1910	8	4.38	22.86
1963	2272	1890	9	3.88	25.71
1964	1580	1850	10	3.50	28.57
1965	1730	1800	11	3.18	31.43
1966	1850	1776	12	2.92	34.29
1967	(1560)	1750	13	2.69	37.14
1968	1750	1740	14	2.50	40.00
1969	2150	1730	15	2.33	42.86
1970	1680		16		
1971	1350	1700	17	2.06	48.51
1972	1510	1690	18	1.94	51.43
1973	(1720)	1680	19	1.84	54.29
1974	1910	1670	20	1.75	57.14
1975	1475	1650	21	1.67	60.00
1976	1690	1645	22	1.59	62.86
1977	1700	1632	23	1.52	65.71
1978	2110	1615	24	1.46	68.57
1979	1670	1600	25	1.40	71.43
1980	2100	1580	26	1.35	74.29
1981	2400		27		
1982	1740	1548	28	1.25	80.00
1983	1600	1510	29	1.21	82.86
1984	1490	1490	30	1.17	85.71
1985	1650	1475	31	1.13	88.57
1986	1230	1420	32	1.09	91.43
1987	1615	1350	33	1.06	94.29
1988	1420	1250	34	1.03	97.14

After Orbuli

A - B
..... FLOOD DATA FOR PEAK DISCHARGE

EXCEEDANCE PROBABILITY ANALYSIS FOR STATION - 8

DONGA AT MANYA
Flood Data for River.....

Period of Record..... $N = 34$ Years; Drainage Area..... 9.040 Km^2

DATE	ANNUAL FLOOD $Q \text{ (m}^3/\text{s)}$	REARRANGED VARIABLE IN DESCENDING ORDER	ORDER NUMBER M	RECURRENCE INTERVAL, $T_r = \frac{N+1}{M}$	PROBABILITY OF EXCEEDANCE $P_{T_r} = \frac{100}{T_r} \%$
(1)	(2)	(3)	(4)	(5)	(6)
1955	(1050)	1550	1	35.00	2.86
1956	(1130)	1400	2	17.50	5.71
1957	(1080)		3		
1958	(920)	1100	4	8.75	11.43
1959	602		5		
1960	735		6		
1961	726	1050	7	5.00	20.00
1962	982	1040	8	4.38	22.86
1963	970	1030	9	3.88	25.71
1964	858	1020	10	3.50	28.57
1965	1050		11		
1966	1030	1000	12	2.92	34.29
1967	1000	982	13	2.69	37.14
1968	947	970	14	2.50	40.00
1969	1040	965	15	2.33	42.86
1970	835	947	16	2.19	45.71
1971	840	945	17	2.06	48.51
1972	864	940	18	1.94	51.43
1973	850		19		
1974	837		20		
1975	846	864	21	1.67	60.00
1976	945	858	22	1.59	62.86
1977	584	852	23	1.52	65.71
1978	852	850	24	1.46	68.57
1979	965	846	25	1.40	71.43
1980	940		26		
1981	1400	840	27	1.30	77.14
1982	1550	837	28	1.25	80.00
1983	1020	835	29	1.21	82.86
1984	1100	735	30	1.17	85.71
1985	(900)	726	31	1.13	88.57
1986	(1010)		32		
1987	(840)	602	33	1.06	94.29
1988	(660)	584	34	1.03	97.14

After Gtuli

A - 9 FLOOD DATA FOR PEAK DISCHARGE

EXCEEDANCE PROBABILITY ANALYSIS FOR STATION - 9

DONGA AT ABONG

Flood Data for River.....

Period of Record..... N = 34 Years; Drainage Area..... 6,424 Km²

DATE	ANNUAL FLOOD Q (m ³ /s)	REARRANGED VARIABLE IN DESCENDING ORDER	ORDER NUMBER M	RECURRENCE INTERVAL, $T_r = \frac{N+1}{M}$	PROBABILITY OF EXCEEDANCE $P_{T_r} = \frac{100\%}{T_r}$
(1)	(2)	(3)	(4)	(5)	(6)
1955	(2700)		1		
1956	(2850)		2		
1957	(2800)	3100	3	11.67	8.57
1958	(2400)	2900	4	8.75	11.43
1959	1650		5		
1960	2700		6		
1961	2300	2800	7	5.00	20.00
1962	2230		8		
1963	(2500)		9		
1964	(2370)	2700	10	3.50	28.57
1965	2800		11		
1966	2510		12		
1967	2400		13		
1968	3100		14		
1969	(3330)	2510	15	2.33	42.86
1970	(2820)	2500	16	2.26	45.71
1971	(1700)	2450	17	2.06	48.57
1972	(2370)	2400	18	1.94	51.43
1973	2250		19		
1974	2120		20		
1975	2150		21		
1976	2330		22		
1977	2450	2330	23	1.52	65.71
1978	2500	2300	24	1.46	68.57
1979	2280	2280	25	1.40	71.43
1980	2900	2250	26	1.35	74.29
1981	(3500)	2230	27	1.30	77.14
1982	(2620)		28		
1983	(2050)		29		
1984	(2680)	2150	30	1.17	85.71
1985	(2340)	2120	31	1.13	88.57
1986	(2650)		32		
1987	(2230)		33		
1988	(1830)	1650	34	1.03	97.14

After Gumbel

APPENDIX A.

A - 10

FLOOD DATA FOR PEAK DISCHARGE

EXCEEDANCE PROBABILITY ANALYSIS FOR STATION - 10

Flood Data for River.....KATSINA ALA AT SEVAV.....

Period of Record.....N = 34 Years; Drainage Area.....22,000 Km².....

DATE	ANNUAL FLOOD Q (m ³ /s)	REARRANGED VARIABLE IN DESCENDING ORDER	ORDER NUMBER M	RECURRENCE INTERVAL, $P_r = \frac{N+1}{M}$	PROBABILITY OF EXCEEDANCE $P_{T_r} = \frac{100\%}{T_r}$
(1)	(2)	(3)	(4)	(5)	(6)
1955	(2600)	3950	1	35.00	2.86
1956	3708	3708	2	17.50	5.71
1957	2900	3500	3	11.67	8.57
1958	2475	3356	4	8.75	11.43
1959	1820	3250	5	7.00	14.29
1960	2875	3200	6	5.83	17.14
1961	2790	3123	7	5.00	20.00
1962	2200	3080	8	4.38	22.86
1963	3080	3000	9	3.88	25.71
1964	(2450)	2900	10	3.50	28.57
1965	2550	2875	11	3.18	31.43
1966	2615	2810	12	2.92	34.29
1967	2810	2790	13	2.69	37.14
1968	3250	2700	14	2.50	40.00
1969	3200	2615	15	2.33	42.86
1970	(2240)	2585	16	2.19	45.71
1971	2000	2550	17	2.06	48.51
1972	1680	2520	18	1.94	51.43
1973	2254	2475	19	1.84	54.29
1974	2520		20		
1975	2060		21		
1976	2300	2400	22	1.59	62.86
1977	2420	2360	23	1.52	65.71
1978	2360	2350	24	1.46	68.57
1979	2700	2300	25	1.40	71.43
1980	3356	2254	26	1.35	74.29
1981	3950		27		
1982	2585	2200	28	1.25	80.00
1983	2150	2150	29	1.21	82.86
1984	2350	2060	30	1.17	85.71
1985	3123	2000	31	1.13	88.57
1986	3000	1820	32	1.09	91.43
1987	2400	1720	33	1.06	94.29
1988	1720	1680	34	1.03	97.14

After Oct 11

A - 11

FLOOD DATA FOR PEAK DISCHARGE

EXCEEDANCE PROBABILITY ANALYSIS FOR STATION - 11

Flood Data for River..... KATSINA ALA AT KASTINA ALA

Period of Record..... N = 34 Years; Drainage Area..... 16700 Km²

DATE	ANNUAL FLOOD Q (m ³ /s)	REARRANGED VARIABLE IN DESCENDING ORDER	ORDER NUMBER M	RECURRENCE INTERVAL, $T_r = \frac{N+1}{M}$	PROBABILITY OF EXCEEDANCE $P_{T_r} = \frac{100}{T_r} \%$
(1)	(2)	(3)	(4)	(5)	(6)
1955	2670	3450	1	35.00	2.86
1956	2826	3164	2	17.50	5.71
1957	2610	3000	3	11.67	8.57
1958	1910		4		
1959	1635	2826	5	7.00	14.29
1960	2428	2670	6	5.83	17.14
1961	3658	2658	7	5.00	20.00
1962	1809		8		
1963	3000	2610	9	3.88	25.71
1964	2050	2598	10	3.50	28.57
1965	2095	2480	11	3.18	31.43
1966	2329	2465	12	2.92	34.29
1967	2145		13		
1968	(2900)	2428	14	2.50	40.00
1969	3164	2329	15	2.33	42.86
1970	2598	2324	16	2.19	45.71
1971	1820	2246	17	2.06	48.57
1972	2007	2241	18	1.94	51.43
1973	2480	2200	19	1.84	54.29
1974	2122	2150	20	1.75	57.14
1975	1700	2145	21	1.67	60.00
1976	1920	2142	22	1.59	62.86
1977	1950	2122	23	1.52	65.71
1978	2246	2095	24	1.46	68.57
1979	2246	2050	25	1.40	71.43
1980	(2650)	2007	26	1.35	74.29
1981	3450	1950	27	1.30	77.14
1982	2465	1920	28	1.25	80.00
1983	2324	1910	29	1.21	82.86
1984	2241	1820	30	1.17	85.71
1985	2200	1809	31	1.13	88.57
1986	(2450)	1700	32	1.09	91.43
1987	2150	1635	33	1.06	94.29
1988	1550	1550	34	1.03	97.14

After G. buli

A - 12

FLOOD DATA FOR PEAK DISCHARGE

EXCEEDANCE PROBABILITY ANALYSIS FOR STATION - 12

Flood Data for River.....TARABA AT GASSOL.....

Period of Record.....N = 34..... Years; Drainage Area.....20,513 Km².....

DATE	ANNUAL FLOOD Q (m ³ /s)	REARRANGED VARIABLE IN DESCENDING ORDER	ORDER NUMBER M	RECURRANCE INTERVAL, $T_r = \frac{N+1}{M}$	PROBABILITY OF EXCEEDANCE $P_{T_r} = \frac{100}{T_r} \%$
(1)	(2)	(3)	(4)	(5)	(6)
1955	1930	2500	1	35.00	2.86
1956	2038	2400	2	17.50	5.71
1957	2000	2280	3	11.67	8.57
1958	1710	2195	4	8.75	11.43
1959	1120	2047	5	7.00	14.29
1960	1885	2038	6	5.83	17.15
1961	1725	2017	7	5.00	20.00
1962	1635	2000	8	4.83	22.83
1963	2280	1930	9	3.88	25.71
1964	1680	1890	10	3.50	28.57
1965	1880	1885	11	3.18	31.45
1966	1785	1880	12	2.92	34.25
1967	1720	1865	13	2.69	37.17
1968	2195	1785	14	2.50	40.00
1969	2400	1730	15	2.33	42.86
1970	2017	1725	16	2.19	45.66
1971	1188	1720	17	2.06	48.58
1972	1675	1710	18	1.94	51.55
1973	1700	1700	19	1.84	54.29
1974	1730	1690	20	1.75	57.14
1975	1410	1685	21	1.67	60.00
1976	1605	1680	22	1.59	62.89
1977	1670	1675	23	1.52	68.71
1978	1650	1670	24	1.46	68.57
1979	1690	1650	25	1.40	71.43
1980	2047	1635	26	1.35	74.07
1981	2500	1605	27	1.30	77.14
1982	1865	1575	28	1.25	80.00
1983	1460	1552	29	1.21	82.86
1984	1552	1460	30	1.17	85.71
1985	1685	1410	31	1.13	88.57
1986	1890	1275	32	1.09	91.43
1987	1575	1188	33	1.06	94.29
1988	1275	1120	34	1.03	97.09

After cutoff