ESTIMATION OF MAXIMUM FLOODS FOR RIVER BENUE AT NUMAN BASED ON BAYES THEOREM

BY

NGENE, Charles Ikechukwu
B.ENG(CIVIL ENGINEERING), ASUTECH

A THESIS SUBMITTED TO DEPARTMENT OF WATER RESOURCES AND ENVIRONMENTAL ENGINEERING IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE AWARD OF THE DEGREE OF MASTER OF SCIENCE (WATER RESOURCES & ENVIRONMENTAL ENGINEERING)

OF THE

AHMADU BELLO UNIVERSITY,

ZARIA,

NIGERIA.

NOVEMBER, 1983

DECLARATION

I hereby declare that this thesis is a record of my own research work. It has not been presented by me in any previous application for a higher degree.



NGENE, C.I.

NOVEMBER, 1988

DEDICATION

To my mother Mrs R.O. Ngene (Nne-Ezekwe),
My brothers; Douglas, George, and John, my sisters,
and the blessed memories of my late father
Chief J. O. Ngene and God almighty my creator.

ABSTRACT

This work presents a procedure for determining the most probable maximum floods from flood data which were used in determining the peak design floods and in estimation of the risk of failure of hydraulic structures with certain return periods and expected project life by numerical analysis. Flood data for river Benue at Numan for the periods (1955-1980) and (1982-1987) were fitted to the "Two parameter lognormal distribution", the "Gumbel extreme value distribution" and tested. The break in periods was due to the construction of a dam at Lagdo (Cameroun) in 1982.

The prior probabilities for the period (1955-1980) were estimated and then used as a weighted combination to update the first set of data using Bayes' theorem. A second approach by weighted combination of individual two-parameter distributions using the same Bayes' theorem enabled a selection of the best among the two-parameter distributions. The peak design floods for the period (1982-1987) were determined given the return periods for each distribution while the risk of failure of hydraulic structures was estimated as a function of return periods and expected project life.

ACKNOWLEDGEMENT

The author herein expresses his indebted heartfelt gratitude to his Supervisor Dr. M. Sowinski for his thorough guidance, Dr. S. Mustafa for his useful guidance and the following other important people for their respective contributions towards the completion of this work. They are:

Dr. C. A. Okuofu the Postgraduate Co-Ordinator in the Department, Dr. O.J. Mudiare of Agricultural Engineering Department, Dr. D.W. Ebong of Mathematics Department, Prof. J. A. Ogunrombi and the Head of Department Dr. L.I. Odigie for his elderly concern and assistance.

Engr. and Mrs E.C.N. Okafor, Dr. and Mrs T.C. Onuchuku, and Ogbuoji, Paul Chukwuka for their moral support.

Messrs Yinua Yussuf, Asindi Iloba Chris, and Mbah Valentine Chude for their encouragement.

My mother, Mrs R.N. Ngene (Nne-Ezekwe) and also Stella Mma Obuekwe for enduring the long period of my absence.

CERTIFICATION

This Thesis titled "Estimation of Maximum Floods for river Benue at Numan based on Bayes' Theorem" by NGENE, Charles Ikechukwu meets the regulations governing the award of the degree of Master of Science (Water Resources & Environmental Engineering) of Ahmadu Bello University, and it is approved for its contribution to knowledge and literary presentation.

.....M. Sowinski..... Date: 25/11/88

Dr. M. Sowinski
Chairman, Supervisory Committee

..... Date: 25/11/88

Dr. S. Mustafa
Member, Supervisory Committee

..... Date: 28/11/88

Dr. L. I. Odigie
Head of Department

..... Date: 18-1-89

Prof. D. I. Saror
Ag. Dean, Postgraduate School

TABLE OF CONTENTS

Chapter		Page
	Title page	i
	Declaration	ii
	Dedication	iii
	Abstract	iv
	Acknowledgement	v
	Certification	vi
	Table of Contents	vii
	List of Tables	xii
	List of Appendixes	xvi
	List of Figures	xvii
	Notation	xix
1	INTRODUCTION	1
1.1	Flood probability estimation	1
1.2	Formulation of research problems	2
2	DEFINITIONS, THEORY AND LITERATURE REVIEW	4
2.1	Introduction	4
2.2	Definitions	4
2.2.1	Peak design flood	4
2.2.2	Maximum probable flood	4
2.2.3	Initial prior probability	4
2.2.4	Posterior probability	5
2.2.5	Likelihoods	5
2.2.6	Combined probability	5
2.2.7	Simple risk of failure	5
2.3	Bayes' theorem	6
2.4	Literature Review	8
2.4.1	Posterior probability	8
2.4.1.1	Posterior probabilities for two distributions	9
2.4.1.2	Posterior probabilities for one distribution	10
2.4.2	Prior probability	11

Chapter		page
2.4.3	Likelihoods	13
2.4.3.1	Likelihood for single recorded peak flood	13
2.4.3.2	Likelihood for several recorded peak flows	15
2.4.3.3	Likelihood f of exceedence	16
2.4.4	Combined probabilities	17
2.4.5	Peak design floods	19
2.4.6	Simple risk of failure	21
2.4.7	Previous findings on Flood probability estimation	23
3	VERIFICATION OF HYDROLOGICAL DATA AND SELECTION OF TYPES OF DISTRUCTION	25
3.1	Introduction	25
3.2	Verification of hydrological data	25
3.2.1	Introduction	25
3.2.2	Collection, analysis and adoption of data	27
3.2.3	Conclusion	32
3.3	Selection of types of distribution	42
3.3.1	Introduction	42
3.3.2	Analysis of data for the period (1955 - 1980)	43
3.3.2.1	Introduction	43
3.3.2.2	Probability plotting for Gumbel extreme value distribution and Two-parameter lognormal distribution for the period (1955-1980)	43
3.3.2.3	Summary	47
3.3.3	Analysis of data for the period (1982-1987)	48
3.3.3.1	Introduction	48
3.3.3.2	Probability plotting for Gumbel extreme value distribution and Two parameter lognormal distribution for the period (1982-1987)	48

(...)

Chapter		page
3.3.3.3	Mean and Standard deviation values for the variable, Q, for the period (1982-1987)	52
3.3.3.4	Gumbel extreme value distribution	55
3.3.3.4.1	Introduction	55
3.3.3.4.2	Chi-squared goodness-of-fit test statistics for Gumebel extreme value distribution for the period (1982-1987)	56
3.3.3.4.3	K-S goodness-of-fit test statistics for Gumbel extreme value distribution for the period (1982-1987)	61
3.3.3.5	Two parameter lognormal distribution	64
3.3.3.5.1	Introduction	64
3.3.3.5.2	Chi-squared goodness-of-fit test statistics for Two parameter lognormal distribution for the period (1982-1987)	65
3.3.3.5.3	K-S goodness-of-fit test statistics for Two parameter lognormal distribution for the period(1982-1987)	68
3.3.3.6	Conclusions	70
4	COMPOUND LOGNORMAL DISTRIBUTION	71
4.1	Introduction	71
4.2	Assigning prior probabilities to each of the component distributions	73
4.3.1	Deriving the low, probable and high values of mean and standard deviation	75
4.3.2	Initial probability matrix	78
4.4	Updating the data of the distribution	80
4.4.1	Lognormal distribution	81
4.4.1.1	Likelihoods for single recorded peak flood	81
4.4.1.2	Posterior probabilities for singular recorded peak flood	82
4.4.1.3	Discussion of results	83

Chapter		Page
4.4.1.4	Singular likelihoods for several recorded peak flows for the period (1982-1987)	84
4.4.1.5	Combining singular likelihoods for several recorded peak flows for the Period (1982-1987)	86
4.4.1.6	Posterior probabilities for several recorded peak flows for the period (1982-1987)	86
4.4.1.7	Discussion of results	89
4.5	General conclusion	89
5	COMBINING LOGNORMAL AND GUMBEL DISTRIBUTIONS	90
5.1	Introduction	90
5.2	Assumption of prior weights	91
5.3	Updating the combined distributions	93
5.3.1.1	Likelihoods for singular recorded peak flood	94
5.3.1.2	Posterior weights for singular recorded peak flood for the period (1982-1987)	95
5.3.1.3	Discussion of results	98
5.3.2.1	Singular likelihoods for several recorded peak flows for the period (1982-1987)	99
5.3.2.2	Combining singular likelihoods for several recorded peak flows for the period (1982-1987)	101
5.3.2.3	Posterior weights for several recorded peak flows for the period (1982-1987)	103
5.3.2.4	Discussion of results	109
5.3.3	Summary	109
5.4	Combining probabilities using Lognormal and Gumbel distributions	110
5.4.1.1	Combined probabilities for flood discharges for the period (1955-1980)	111

Chapter		page
5.4.1.2	Discussion of results	116
5.4.2.1	Combined probabilities for flood discharges for the period (1982-1987)	117
5.4.2.2	Discussion of results	121
6	DETERMINATION OF PEAK DESIGN FLOODS	123
6.1	Introduction	123
6.2	Lognormal distribution	124
6.2.1	Determining peak design return floods for given return periods	124
6.2.2	Discussion of results	126
6.3	Gumbel distribution	127
6.3.1	Determining the peak design return floods for given return periods	127
6.3.2	Discussion of results	129
6.4	Summary	130
7	ESTIMATION OF RISK OF FAILURE	131
7.1	Introduction	131
7.2	Lognormal distribution	132
7.2.1	Estimating simple risk of failure	132
7.2.2	Discussion of results	134
7.3	Gumbel distribution	136
7.3.1	Estimating simple risk of failure	136
7.3.2	Discussion of results	137
7.4	Summary	138
8	CONCLUSIONS AND RECOMMENDATIONS	139
8.1	Conclusions	139
8.2	Recommendations	141
	REFERENCES	142
	APPENDIXES	144

LIST OF TABLES

Table		Page
3.1	Flood discharge on river Benue at Numan from data obtained from Upper Benue River Basin and Rural Development Authority, Yola.	34
3.2	Flood discharges on river Benue at Numan from data obtained from Federal Department of Water Resources, Mando Road, Kaduna for the period (1955-1980).	36
3.3	Stage-discharge for rating curve number I from Inland Waterways Division, Lokoja branch, for the period (1955-1972) for river Benue at Numan.	38
3.4	Stage-discharge for rating curve number II from Inland Waterways Division, Lokoja branch for the period (1973-1980), for river Benue at Numan.	39
3.5	Stage-discharge for river Benue at Numan divided into sections from Inland Waterways Division, Lokoja branch, for the period (1955-1987).	40

Table		Page
3.6	Probability test for flood discharges for the period (1955-1980).	44
3.7	Probability test for flood discharges for the period (1982-1987).	49
3.8	Mean and Standard deviation values for the period (1982-1987).	55
3.9	Histogram intervals for flood discharges for Gumbel extreme value distribution for the period (1982-1987).	59
3.10	K-S test statistics for Gumbel extreme value distribution for flood discharges for the period (1982-1987).	63
3.11	Histogram intervals for Two parameter lognormal distribution for flood discharges for the period (1982-1987).	66
3.12	K-S test statistics for Two parameter lognormal distribution for flood discharges for the period (1982-1987).	69
4.1	Prior probability for the period (1955-1980).	75
4.2	The low, probable and high values of mean and standard deviation for the variable, Q.	78

Table		page
4.3	Initial probability matrix	79
4.4	Singular likelihoods and posterior probabilities for 1981 only	83
4.5	Singular likelihoods for six year peak flows for the period (1982-1987)	85
4.6	Combined likelihoods and posterior probabilities for several recorded peak flows.	88
5.1	Assumed prior weights for the period (1982-1987).	92
5.2	Singular likelihoods and posterior weights for 1981 only.	96
5.3	Singular likelihoods and posterior weights for 1981 only	98
5.4	Singular likelihoods for six year peak flows for the period (1982-1987).	100
5.5	Combined likelihoods and posterior weights for several recorded peak flows.	106
5.6	Combined likelihoods and posterior weights for several recorded peak flows.	108
5.7	Combined probabilities for flood discharges for the period (1955-1980).	113

Table		Page
5.8	Combined probabilities for flood discharges for the period (1982-1987).	119
6.1	Peak design return floods for given return periods obtained using Lognormal distribution	127
6.2	Peak design return floods for given return periods obtained using Gumbel distribution	128
7.1	Risks of failure given the expected project life.	134
7.2	Risks of failure given the expected project life.	137

LIST OF APPENDIX

APPENDIX		PAGE
I	Construction of the dam at Lagdo in the Cameroons.	144
II	Cumulative Normal Distribution	145
III	Percentile Values ($X^2_{\alpha,\nu}$) for the Chi-square Distribution with ν degrees of freedom (shaded area = α)	146
IV	Critical values for the K-S test statistic	147
V	Values of the standardized Type I extreme-value distribution	148
V	Ordinates Y of the standard Normal Density.	149

LIST OF FIGURES

Figure		Page
2.1	Partition of S and Q into mutually exclusive events	6
2.2	Compound distribution	19
3.1	Location of study area	26
3.2	Stage-discharge rating curve for river Benue at Numan for the period (1955-1980) from Federal Department of Water Resources, Mando Road, Kaduna.	29
3.3	Stage-discharge rating curve for river Benue at Numan for the periods (1955-1972) and (1973-1987) from Inland Waterways Division, Lokoja Branch, for numbers I and II respectively.	30
3.4	Gumbel extreme distribution probability for river Benue at Numan for the period (1955-1980)	45
3.5	Lognormal probability for river Benue at Numan for the period (1955-1980)	46
3.6	Gumbel extreme distribution probability for river Benue at Numan for the period (1982-1987).	50
3.7	Lognormal probability for river Benue at Numan for the period (1982-1987).	51

Figure		Page
5.1	PDF for river Benue at Numan for the period (1955-1980)	114
5.2	PDF for river Benue at Numan for the period (1982-1987).	120

LIST OF NOTATIONS

a_n = function of n

A = actual data for the period (1982-1987)

b_n = function of n

B = basic data for the period (1955-1980)

B_j = additional data to B

C = combined

D_j = critical value of the chi-squared goodness-of-fit test statistic

$D_{2(max)}$ = observed value of the K-S test statistic

$D_{n,\alpha}$ = critical value of the K-S test statistics

E_j = normal expected number of observations

$f(Q|M_Y,\sigma_Y 2)$ = likelihood of mean M_Y and standard deviation σ_Y, given the variable, Q.

$f_u(u)$ = PDF for the standardized variable, u.

$f_W(w)$ = PDF for the reduced variate, w.

$F(Q_i)$ = CDF for the variable, Q_i.

$F^*(Q^{(i)})$ = observed cumulative histogram

$F_u(u)$ = CDF for the standardized variable, u

$F_W(w)$ = CDF for the reduced variate, w

H = High

H_{ei} = elements in a set or set of events

i = order of samples

i_j = index value of mean

j = index of standard deviation value

k = number of categories

k_D = number of years

$l(Q|\Theta)$ = likelihood of joint occurrence of independent events Q given Θ

L = Low

L_{ij} = likelihood of combination i,j

L_{J2} = likelihood that Q_2 would occur if distribution J is correct

L_{22} = likelihood that Q_2 would occur if distribution 2 is correct

$L^c_{G(Q)}$ = combined likelihood for Gumbel distribution

$L_G(Q_k)$ = singular likelihood for Gumbel distribution

$L^c_{LN}(Q_{\cdot})$ = combined likelihood for Lognormal distribution

$L_{LN}(Q_k)$ = singular likelihood for Lognormal distribution

M_Q = mean value for the variable Q

M_Y = natural logarithms of the mean value

n = size of data

n_F = expected project life

N = total number of samples ofthe distribution

N_f = normalization factor

N_{f1} = normalization factor for first approach

N_{f2} = normalization factor for second approach

O_j = number of observed values

P = probable

P_r = probability

P_{J1} = probability that Q_J will be exceeded using distribution 1

P_{j2} = probability that Q_j will be exceeded using distribution 2

$P'_{i,j}$ = posterior probability

$P_{i,j}$ = prior probability

$P^{j}_{i\ B_j,j}$ = posterior probability (weight) of combination, i,j, for section B_j

$P(H_{ei})$ = prior probabilities of the events H_{ei}

$P(H_{ei}|Q)$ = posterior probabilities of the events H_{ei}

$P(Q|H_{ei})$ = likelihood function

$P(Q)_{LN}$ = density function for lognormal distribution

$P_G(Q)$ = PDF for Gumbel distribution

$P_{LN}(Q)$ = PDF for lognormal distribution

$P_c(Q>Q_i)$ = probability of a flood of magnitude Q exceeding Q_i.

P_L = likelihood of flow Q not being exceeded by another flow of magnitude Q_i in n_F- years

$P^n(Q)$ = functional equation

Q = variable

Q_1 = design magnitude

Q_T = peak design return floods for given return periods, T

r = no. of parameters estimated from the data

RY_{n_F} = simple risk of failure involved in hydrologic design for given expected project life, n_F

S = sample space

S_Θ = summation

T = return period

u = standardized variable

u_G = parameter for the reduced variate

γ_Q = coefficient of variation for the variable Q

w = reduced variate

w_1 = probability that distribution 1 is correct

w_2 = probability that distribution 2 is correct

w_G = assigned prior weight to Gumbel distribution

w_{12} = posterior probability (weight) that distribution 1 is correct (given that Q_2 has occurred)

w_{22} = posterior probability (weight) that distribution 2 is correct (given that Q_2 has occurred)

W_{G1}^{1A} = posterior weight for Section A due to Gumbel distribution for first approach.

W_{G2}^{1A} = posterior weight for section B_1 due to Gumbel distribution for second approach.

W_{LN} = assigned weight to lognormal distribution

$W_{LN_1}^{1A}$ = posterior weight to section A due to lognormal distribution for first appraoch

$W_{LN_2}^{1A}$ = posterior weight for section B_1 due to lognormal distribution for second approach

Y = natural logarithm of Q

σ_Q = standard deviation for the variable Q

σ_Y = natural logarithm of the standard deviation value

α = significance level

$\hat{\alpha}_{ex}$ = scale parameter

α_G = parameter for the reduced variate

$\hat{\beta}$ = location parameter

π = product

Σ = summation

X^2 = observed value of chi-squared goodness-of-fit test statistic

$X^2_{\alpha,\nu}$ = critical value of the chi-squared goodness-of-fit test statistic

ν = degrees of freedom

CHAPTER 1

INTRODUCTION

1.1 FLOOD PROBABILITY ESTIMATION

The estimation of maximum floods is very important because it is the basis from which most hydraulic/ hydrological designs are made. A poor knowledge of this will give room to poor designs leading to immediate or eventual failure of such projects. There are procedures for estimating these maximum floods and the methods are as follows:

(i) Traditional approach using frequency analysis;

(ii) Approach based on probability distribution.

In customary approach a frequency analysis is applied to the available data and floods with certain specified average return periods, T-years (like say 1.5, 6, 12 -yrs) are derived. This is applied in situations where adequate data are available.

In some situations where the consequences of failure are very high, the maximum probable flood is used for design. This maximum probable flood is obtained using the second approach. In this approach a probability distribution is plotted. It is a graph of probability against maximum discharges. The peak flood at the highest point of the probability distribution is the maximum

KASHIM IBRAHIM LIBRARY
AHMADU BELLO UNIVERSITY
ZARIA, NIGERIA.

(a) derivation of parameters for component distributions and assumption of prior probabilities;

(b) updating the prior probabilities using Bayes' theorem;

(iv) Combining Lognormal and Gumbel distributions by:

(a) assuming subjective weights for distributions obtained for the period (1982-1987)

(b) updating the weights assigned to distributions computed for the period (1955-1980) by consideration of

one year record (1981 only) and several year record (1982-1987),

using Bayes' theorem;

(c) estimation of the influence of both approaches on the PDF.

(v) determination of peak design floods for given return periods;

(vi) estimation of simple risk of failure for different expected project life and given return periods;

(vii) drawing conclusions and recommending solutions.

by taking percentage points of areas under the normal curve.

2.2.4 POSTERIOR PROBABILITY:

Posterior probability is a weighted combination of a prior probability and a likelihood function which are normalize to make the sum of the posterior probabilities add up to 1.0.

2.2.5 LIKELIHOODS:

Likelihoods area density functions obtained by combination of mean and standard deviation i.e. each is being specified by a mean and standard deviation and when duely followed the distribution is said to be accepted. In order words, they are in form of models.

2.2.6 COMBINED PROBABILITY:

Combined probability is a weighted combination of two probabilities using two different distributions of the same distribution or different distributions. This could be the prior weight or posterior weight being combined by the probability distribution function of the same distribution distributions.

2.2.7 SIMPLE RISK OF FAILURE:

The simple risk of failure is the probability of occurrence of the variable, Q, greater than the design magnitude, Q_1, which is also equal to the chance of nonoccurrence of $Q \leqslant Q_1$ (Yen (1970)).

2.3 BAYES' THEOREM

According to Kalbfleisch (1979(a)) Bayes' theorem is explained using Fig. 2.1 of the set theory and the explanation is as follows:

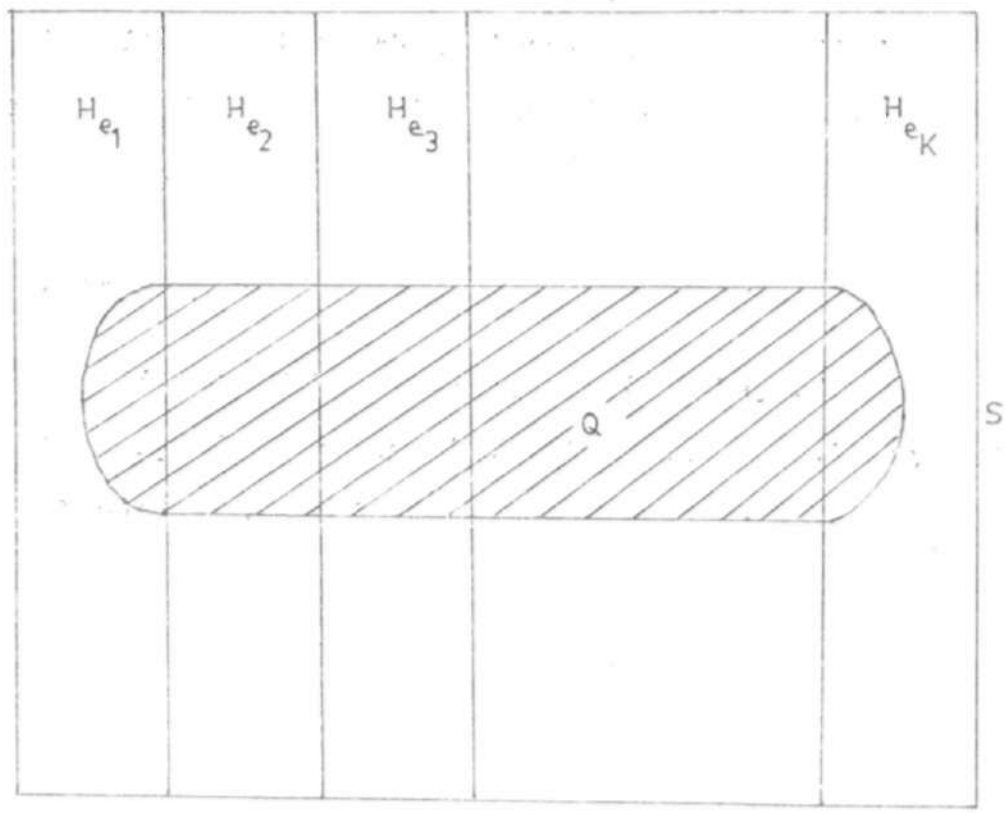


FIG. 2.1 : Partition of S and Q into mutually exclusive events

$$S = H_{e_1} \cup H_{e_2} \cup H_{e_3} \cup \ldots \cup H_{e_k};$$

$$Q = QH_{e_1} \cup QH_{e_2} \cup QH_{e_3} \cup \ldots \cup QH_{e_k}$$

Assume $H_{e_1}, H_{e_2} \ldots, H_{e_k}$ to be a set of equal partition in the sample space and let Q be any other event

as shown in Fig. 2.1. Suppose that the probabilities $P(H_{e_i})$ and the Conditional probability $P(Q|H_{e_i})$ of event Q within each of the sets H_{e_i} of the partition are known. The problem becomes to determine the conditional probability $P(H_{e_i}|Q)$ of the events H_{e_i} when Q is known to have occurred. If

$$P(Q) > 0 \quad \ldots \quad \ldots \quad (2.1)$$

By definition of eqn (2.1), let H_{e_i} and Q be two events defined on the same sample space, with $P(H_{e_i}) > 0$. The conditional probability of H_{e_i} given the observed data, Q is defined by

$$P(H_{e_i}|Q) = \frac{P(Q,H_{e_i})}{P(Q)} \quad \ldots \quad \ldots \quad (2.2)$$

The numerator in eqn (2.2) using eqn (1.2.1) in Box and Tiao (1973) may be expressed as:

$$P(Q,H_{e_i}) = P(Q|H_{e_i})\ P(H_{e_i}) \quad \ldots \quad \ldots \quad (2.3)$$

and substituting eqn (2.3) into eqn (2.2), we get

$$P(H_{e_i}|Q) = \frac{P(Q|H_{e_i})\ P(H_{e_i})}{P(Q)} \quad \ldots \quad \ldots \quad (2.4)$$

According to Kalbfleisch this rule was first given in a paper by the Reverend Thomas Bayes in 1763 and is called Bayes's theorem. He further explained that the

probability P(Q) can be obtained as

$$P(Q) = \sum_{i=1}^{n} P(Q|H_{e_i}) P(H_{e_i}) \quad \ldots \quad (2.5)$$

He said that the probabilities $P(H_{e_i})$ are called the a priori probabilities or prior probabilities of the events H_{e_i}, and that the conditional probabilities $P(H_{e_i}|Q)$ are called the a posterior probabilities or posterior probabilities of the events H_{e_i}. So the conditional probabilities $P(H_{e_i}|Q)$ called posterior probabilities depends on the other conditional probability $P(Q|H_{e_i})$ called the likelihood functions. As such obtaining the posterior probabilities depends or is conditioned on obtaining the likelihood functions.

Bayes's theorem modifies the prior probabilities to incorporate information provides by the occurrence of event Q. From the aforementioned explanations the Bayes's theorem is said to be a direct consequence of the definition of conditional probability. However, use of the Bayes's theorem is sometimes advocated in situations where the prior probabilities $P(H_{e_i})$ cannot be verified empirically.

2.4 LITERATURE REVIEW

2.4.1 POSTERIOR PROBABILITY

The work in posterior probabilities are reviewed in the following two orders:

(i) Posterior probabilities using compound distributions as a weighted combination of individual two-parameter distributions;

(ii) Posterior probabilities using compound distributions as a weighted combination of different parameter of the same distribution.

So the first thing to do is to review works done in (i).

2.4.1.1 POSTERIOR PROBABILITIES FOR TWO DISTRIBUTIONS

According to Russell (1982), if the peak flow in one year is known to be Q_2, then the weights can be updated, from Bayes' theorem which states that:

$$\text{Posterior probability} = \frac{\text{Prior probability x Likelihood}}{\text{Normalization factor}} \quad \ldots \quad \ldots \quad \ldots \quad (2.6(a))$$

or

$$W_{J2} = \frac{W_J \, L_{J2}}{N_F} \quad \ldots \quad \ldots \quad (2.6(b))$$

in which W_{J2} = posterior probability (weight) that distribution J is correct (given that Q_2 has occured),

$W_{J2} = W^J_{LN}$ i. e posterior probability for Lognormal distribution,

L_{J2} = Likelihood that Q_2 would occurs if distribution J is correct (it is given by the height ofthe probability density curve)

$L_{12} = L_{LN}(Q_2)$ i.e. Likelihood for Lognormal distribution,

and $W_{22} = \frac{W_2 L_{22}}{N_F}$ ... ... (2.6(c))

in which W_{22} = posterior probability (weight) that distribution 2 is correct (given that Q_2 has occurred),

$W_{22} = W_G'$ i.e. posterior probability for Gumbel distribution;

L_{22} = Likelihood that Q_2 would occur if distribution 2 is correct (it is given by the height of the probability curve),

$L_{22} = L_G(Q_2)$ i.e. Likelihood for Gumbel distribution;

$$N_F = W_1 L_{12} + W_2 L_{22}$$

$$= W_{LN}' L_{LN}(Q_2) + W_G' L_G(Q_2) \quad \ldots \quad .. \quad (2.7)$$

and is simply a normalization factor to make the sume of the posterior probabilities add up to 1.0.

2.4.1.2 POSTERIOR PROBABILITIES FOR ONE DISTRIBUTION

According to Russell (1982), this is compound distribution that uses each likelihood to update a prior probability and renormalizing to make the new (posterior) probabilities add up to 1.0, i.e.

$$P'_{i_1 j} = \frac{P_{i_1,j}\, L_{i_1,j}}{N_F} \quad \ldots \quad \ldots \quad (2.8)$$

in which $P'_{i_1,j}$ = posterior probability (weight) assigned to combination i_1, j;

i_1 = index value of mean (1-n);

j = index of standard deviation value (1-n);

n = size of sample

$P_{i_1,j}$ = Prior probability of combination i_1,j;

$L_{i_1,j}$ = likelihood of combination of i_1, j, given the data Q,

$= L_{i_1,j}(Q_1)$ i.e. likelihood of the variable Q, given the design magnitude, Q_1, and

N_f = normalization factor $= \sum_{i_1=1}^{n} \sum_{j=1}^{n} P_{i_1,j} L_{i_1,j}$

... ... (2.9)

2.4.2 PRIOR PROBABILITY

Several authors contributed on what prior probability could be and some of their contributions are as follows:

According to Box and Tiao (1973) "a priori distribution is supposed to represent knowledge about

ABMU... LIBRARY
... BELLO UNIVERSITY

parameters before the outcome of a projected experiment is known". On the other hand, Macdonald (1966) has it that the prior probability could be as:

$$P(H_{e_i}) = \frac{H_{e_i}}{\Sigma H_{e_i}} \quad \ldots \quad \ldots \quad (2.10)$$

where H_{e_i} are independent events in a container and this events could be drawn randomly without any replacement.

According to Arnold (1971), there are situations in which investigators know little or intend to continue as if they knew little, about parameters of a model. Hence, there is need for clear-cut or definite rules for selecting a priori distribution that would look like knowing little. However, filling this need has been a problematic aspect of the Bayesian approach to inference used by Bodo and Unny (1976). This led Jeffreys to suggest some rules.

Jeffreys in Arnold (1971) suggested two rules for selecting a priori distribution, which, according to him "... cover the commonest cases". He stated the rules as follows:

(i) "If the parameter may have any value in a finite range, or from $-\infty$ to $+\infty$ its prior probability should be taken as uniformly distributed.

(ii) If it rises in such a way that it may conceivably have any value from 0 to ∞, the prior probability of its logarithm should be taken as uniformly distributed".

This will indeed act as a guide in selecting types of model. It actually applied to the two-parameter models of Lognormal and Gumbel distributions shown in Bodo and Unny (1976). The next thing to do is review works done in area of likelihoods.

2.4.3 LIKELIHOODS

The previous works done in area of likelihoods were carried out in three main ways. The ways are:

(i) Likelihood for single recorded peak flood;

(ii) Likelihood for several recorded peak flows (combining Likelihoods);

(iii) Likelihood of exceedence

The order in which these works were done starting with (i) are as follows:

2.4.3.1 LIKELIHOOD FOR SINGLE RECORDED PEAK FLOOD

According to Bodo and Unny (1976) in their example of 'prior distribution' in which they used two-parameter model forms to represent order of alternatives typically entertained as potential flood models. They are given

by the height oftheprobability density functions as:

For Lognormal model:

$$f(Q|M_Y,\sigma_Y^2) = \frac{1}{Q\sigma_Y(2\pi)^{\frac{1}{2}}} \exp\left[-\tfrac{1}{2}\left(\frac{\ln Q - M_Y}{\sigma_Y}\right)^2\right] \quad \ldots (2.11)$$

$$0 < Q < \infty \quad -\infty < M_Y < \infty \quad 0 < \sigma_Y^2 < \infty$$

in which $f(Q|M_Y, \sigma_Y^2) = L_{i_j,j}(Q)$ = likelihood of mean M_Y. and standard deviation σ_Y, given the variable $Q = L_{LN}(Q)$;

i_j = index of mean value (from 1-n)

j = index of value of standard deviation (from 1-n);

For Gumbel model:

$$f(Q|\alpha_{ex}, \beta) = \frac{1}{\alpha_{ex}} \exp\left\{ - \frac{(Q-\beta)}{\alpha_{ex}} - \exp\left[- \frac{(Q - \beta)}{\alpha_{ex}}\right]\right\} \ldots (2.12)$$

$$-\infty < Q < \infty \qquad 0 < \alpha_{ex} < \infty \qquad -\infty < \beta < \infty$$

In which $f(Q|\alpha_{ex},\beta) = L(Q)$ i.e. likelihood $= L_G(Q)$. They did not explain what the parameters $\hat{\alpha}_{ex}$ and $\hat{\beta}$ really meant. Another researcher Haan (1982), explained that the parameters $\hat{\alpha}_{ex}$ and $\hat{\beta}$ are scale and location parameters with $\hat{\beta}$ being the mode of the distribution. Some other researchers in their own work explained how these parameters of the type 1 extreme value distribution called the Gumbel distribution could be

estimated. They are Lowery and Nash (1970) in Haan (1982) using method of moments, after trying several other methods, gave the estimators to be:

$$\hat{\alpha}_{ex} = \frac{\sigma_Q}{1.283} \quad \ldots \quad \ldots \quad (2.13)$$

and

$$\hat{\beta} = M_Q - 0.45\ \sigma_Q \text{ (maximum)}$$

$$= M_Q + 0.45\ \sigma_Q \text{ (minimum)} \quad \ldots \quad .. \quad (2.14)$$

Bodo and Unny (1976) did tests on the two-parameters models to provide basis for their use in Bayesian probabilities andto become more familiar with the data in use. The tests were ran using Kolmogorov-Smirnov (K-S) test statistics for 46 events because it is a relatively weak test, it was conducted at 90% confidence level, and it accepted all the models. A second test using Chi-squared test, a less generous test was conducted at 99% confidence level using the same 46 events and it as well accepted the two distributions. This is why the two distributions are considered in the sample data obtained for this study.

2.4.3.2 LIKELIHOOD FOR SEVERAL RECORDED PEAK FLOWS

Bodo and Unny (1976) as a method of combining likelihoods from several year records explained that the likelihood function $L(Q|\theta)$ was derived as the probability of the joint occurrence of independent events Q given

θ i.e.

$$L\ (Q|\theta) = \prod_{i=1}^{n} f(Q|\theta) \qquad \ldots \qquad (2.15)$$

According to Russell (1982), "when there are several recorded peak flows, the likelihood of each combination is the product of the likelihoods computed from the individual pieces of data", i.e.

$$L^{c}_{i_1,j}\ (Q) = \Pi L_{i_1,j}\ (Q_{k_D}) \qquad \ldots \qquad \ldots \qquad (2.16)$$

In which $K_D = 1, 2, \ldots, n$

c means combined.

What he showed was that for likelihoods for say three years, their combined likelihoods using eqn (2.16) will be:

$$L^{c}_{11}\ (Q) = L_{11}(Q_1) \times L_{11}(Q_2) \times L_{22}(Q_3) \ldots \quad (2.16(a))$$

$$L^{c}_{22}\ (Q) = L_{22} \times L_{22}\ (Q_2) \times L_{22}\ (Q_3) \quad \ldots \quad (2.16(b))$$

$$L^{c}_{33}\ (Q) = L_{33}\ (Q_1) \times L_{33}(Q_2) \times L_{33}(Q_3) \ldots (2.16(c))$$

2.4.3.3 *LIKELIHOOD OF EXCEEDENCE*

According to Russell (1982) "for information on floods exceeded or not exceeded a given number of times

the likelihood are based on the probability of a flow being exceeded in any one year and for Lognormal distribution it becomes:

$$P(Q \geq Q_1) = 1 - \frac{1}{2\pi \sigma_Y} \int_{-\infty}^{Q} \exp\left[-\frac{1}{2} \frac{(\log Q - M_Y)}{\sigma_Y} \right] dQ \quad \ldots \quad \ldots \quad (2.17)$$

i.e.

$$P(Q \geq Q_1) = 1 - CDF \quad \ldots \quad \ldots \quad (2.18)$$

In which $P(Q > Q_1)$ = Probability that the flow, Q, will be exceeded in any one year given the distribution with mean, M_Y and Standard deviation σ_Y.

He further said that for n_F years, where n_F is the expected project life, that the likelihood of flow Q not being exceeded will be expressed as:

$$P_L = (1 - P(Q \leq Q_1))^{n_F} \quad \ldots \quad .. \quad (2.19)$$

2.4.4 COMBINED PROBABILITIES

Combining probabilities which could be in the form of combining prior weights or posterior weights with its probability density functions of two distributions was expressed by Russell (1982). In explaining this he used an example based on peak flood regime on a river and assumed that this river could be featured with

one or other of the probability distributions shown in Fig. 2.2. An additional assumption was made that either of the distributions is the correct one. Based on these assumptions theprobabilities for the two distributions were combined as follows:

$$P_c\ (Q \leq Q_1) = W_1\ P_{11} + W_2\ P_{12} \ldots \quad \ldots \quad (2.20)$$

In which $P_c\ (Q \leq Q_1)$ = probability of a flood of magnitude Q_1 exceeding or equalling Q;

Q = variable,

Q_1 = design magnitude

W_1 = probability that distribution 1 is correct,

$W_1 = W_{LN} = W'_{LN}$ i.e. this could be equal to the prior or posterior probabilities respectively depending on its application and it has been assumed to be the Lognormal distribution;

W_2 = probability that distribution 2 is correct;

$W_2 = W_G = W'_G$ i.e. this could be equal to prior or posterior probability respectively depending on its application and it has been assumed to be the Gumbel distribution,

P_{11} = probability that Q_1 will be exceeded using distribution 1,

= probability density function for Lognormal distribution, $f_u(u)$;

P_{12} = probability density function for Gumbel distribution, $f_W(w)$

He further stated that W_1 and W_2 may be considered as weights. Their summation is equal to 1.0.

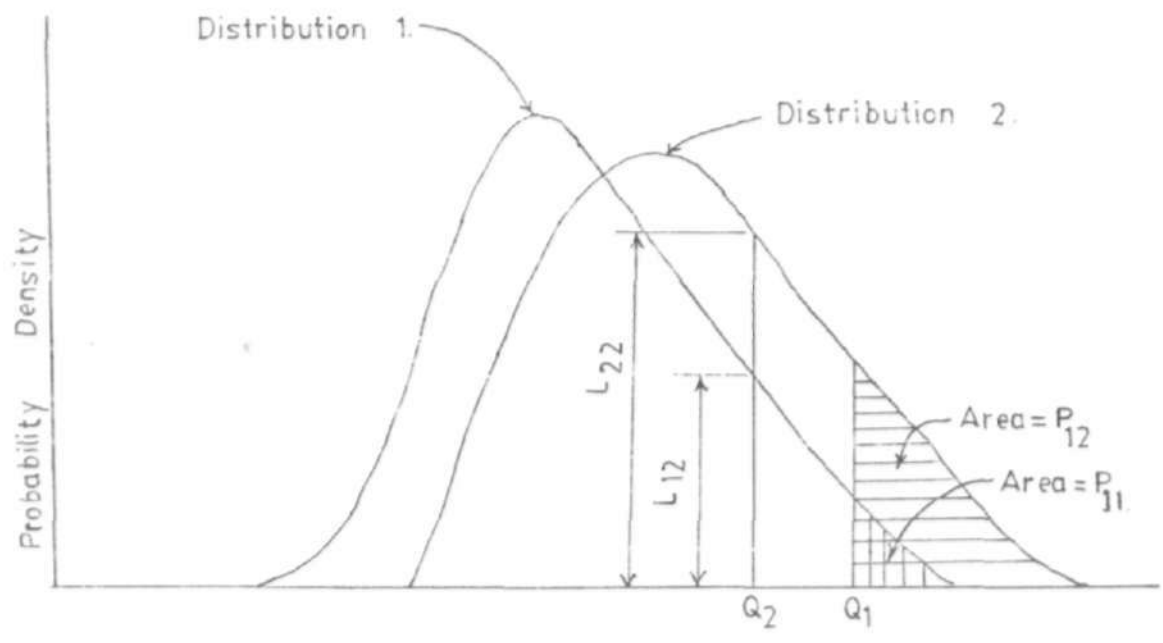


FIG. 2.2: COMPOUND DISTRIBUTION

2.4.5 PEAK DESIGN FLOODS

According to Gumbel (1958) the initial distributions of floods are exponential and that the probability of a flood being less than Q is

$$\Phi\ (Q) = \exp\ -e^{-\alpha_G(Q - U_G)} \qquad \ldots \qquad \ldots\ (2.21)$$

This eqn (2.21) is based on the statistical theory of floods and that if this theory holds he gave that the observed floods will be estimated as

$$Q = U_G + \frac{W}{\alpha_G} \qquad \ldots \qquad \ldots \qquad (2.22)$$

In which Q = observed flood;

U_G = most probable annual flood after differentiating eqn (2.21) twice,

W = reduced variate.

The parameters α_G and U_G are parameters of the estimate and Gumbel (1958) estimated them to be:

$$\hat{U}_G = M_Q - \frac{\bar{W}_N}{\alpha_G} \qquad \ldots \qquad \ldots \qquad (2.23)$$

and

$$\frac{1}{\alpha_G} = \frac{\sigma_Q}{\sigma_N} \qquad \ldots \qquad \ldots \qquad (2.24)$$

where M_Q and σ_Q are the sample mean and standard deviation;

$\bar{W}_N$ and σ_N are the mean and standard deviation of the reduced values obtained from the solutions of

$$\Phi\ (W_M) \quad = \frac{m}{N+1} \quad \ldots \quad \ldots \quad (2.25)$$

Indeed, according to Benjamin and Cornell (1970) the parameters α_G and U_G could be determined more directly using eqns (2.26) and (2.27), respectively without having to go all this steps that might create one mistake or the other, i.e.

$$\alpha_G = \frac{1.282}{\sigma_Y} \quad \ldots \quad \ldots \quad (2.26)$$

$$U_G = M_Y - \frac{0.577}{\alpha_G} \quad \ldots \quad \ldots \quad (2.27)$$

We see that eqn (2.22) did not have the parameters on the right to be in their exponential form and as such will give very low peak flows which could not possibly be the peak design floods. In the body of the text an approach due to formulations from the reduced variate in Benjamin and Cornell will be used. This determination is for the evaluation of peak design floods in Gumbel distributions.

2.4.6 SIMPLE RISK OF FAILURE

Yen (1970) treated the statistical theory that form the basis for theevaluation of the simple risk of failure involved in hydrologic designs by invoking "the theory that the sum of the probabilities of occurrence and nonoccurrence of an event is equal to unity i.e.

$$P\ (Q > Q_1) + P\ (Q \leq Q_1) = 1 \quad \ldots \quad \ldots \quad (2.28)$$

To be able to evaluate the simple risk involved in any hydrologic design he derived a number of steps as follows:

(i) Interprete the probability of the variable, Q, being equal to or greater than the variate, Q_1, in each year as:

$$P\ (Q \geq Q_1) = \frac{1}{T} \quad \ldots \quad \ldots \quad (2.29)$$

where T is the return period in years and he defined it as the average length of time in which statistically Q_1 will be equalled or exceeded once and only once.

(ii) In natural hydrologic phenomena Q is usually a continous variable, as such, the probability of $Q = Q_1$ is zero and

$$P\ (Q \geq Q_1) + P(Q \leq Q_1) = 1.0 \quad \ldots \quad \ldots \quad (2.30)$$

(iii) Following the definition of return period, then the probability of $Q \leq Q_1$ for each year will become:

$$P\ (Q \leq Q_1) = 1 - \frac{1}{T} \quad \ldots \quad \ldots \quad (2.31)$$

An assumption was made from the beginning that the occurrence of the events Q was independent from one another and as such the probability of ocurrence of

$Q \leq Q_1$ for each year for the entire period of n_F- years becomes:

$$P^{n_F}(Q \leq Q_1) = P(Q \leq Q_1) = (1 - \frac{1}{T})^{n_F} \quad \ldots \quad (2.32)$$

Therefore, applying the theory of the sum of probabilities of occurrence and nonoccurrence of an event to eqn (2.32) will give his definition of simple risk of failure which he defined as "the probability of occurrence of the variable, Q, greater than the design magnitude, Q_1, which is also equal to the chance of nonoccurrence of $Q \leq Q_1$, i.e.

$$\text{Risk} = P(Q > Q_1) = 1 - (1 - \frac{1}{T})^{n_F} \quad \ldots \quad (2.33)$$

It is this definition of simple risk of failure of hydrologic designs for hydraulic structures that is used for its estimation.

2.4.6 PREVIOUS FINDINGS ON FLOOD PROBABILITY ESTIMATION

From the review of the work that has been done on the problem for flood probability estimation, it is important to point out that flood probability estimation has received little attention and when it has to be a case in Nigeria little or no attention at all. Researchers have concentrated on flood probability estimation without showing ways of estimating maximum floods based on Bayes' theorem in situations where data are insufficient or where there are an obstruction to

the river or stream flowing capacity thereby affecting its daily water level. It is the realization of this neglect particularly in Nigerian situation where there are little or no records at all that necessitated the investigation for the development of a procedure for estimating maximum floods based on Bayes' theorem.

CHAPTER 3

VERIFICATION OF HYDROLOGICAL DATA AND SELECTION OF TYPES OF DISTRIBUTION

3.1 INTRODUCTION

This chapter has been divided into two broad sub-chapters with a second chapter being divided into two sub-chapters again. The divisions are as follows:

(i) Verification of hydrological data;

(ii) Selection of types of distributions for the two sets of data (1955-1980) and (1982-1987).

3.2 VERIFICATION OF HYDROLOGICAL DATA

3.2.1 INTRODUCTION

In hydrological designs a good design calls for a good data and vice-versa. So the first thing that need to be done before actual work is started is to collect and verify the hydrologic data to be used. These data were to be collected for river Benue at Numan. Numan is a town in Gongola State of Nigeria as shown in Fig. 3.1. Gongola State has its capital city as Yola and Numan is about 100km North-west of Yola. Numan was chosen because (a) the data were mostly available; (b) It is in one of the two major rivers in Nigeria; and (c) It has not had most of the artifical interruption like the construction of a dam etc unlike other sites downstream of river Benue. Fig. 3.1 as to be mentioned in Section

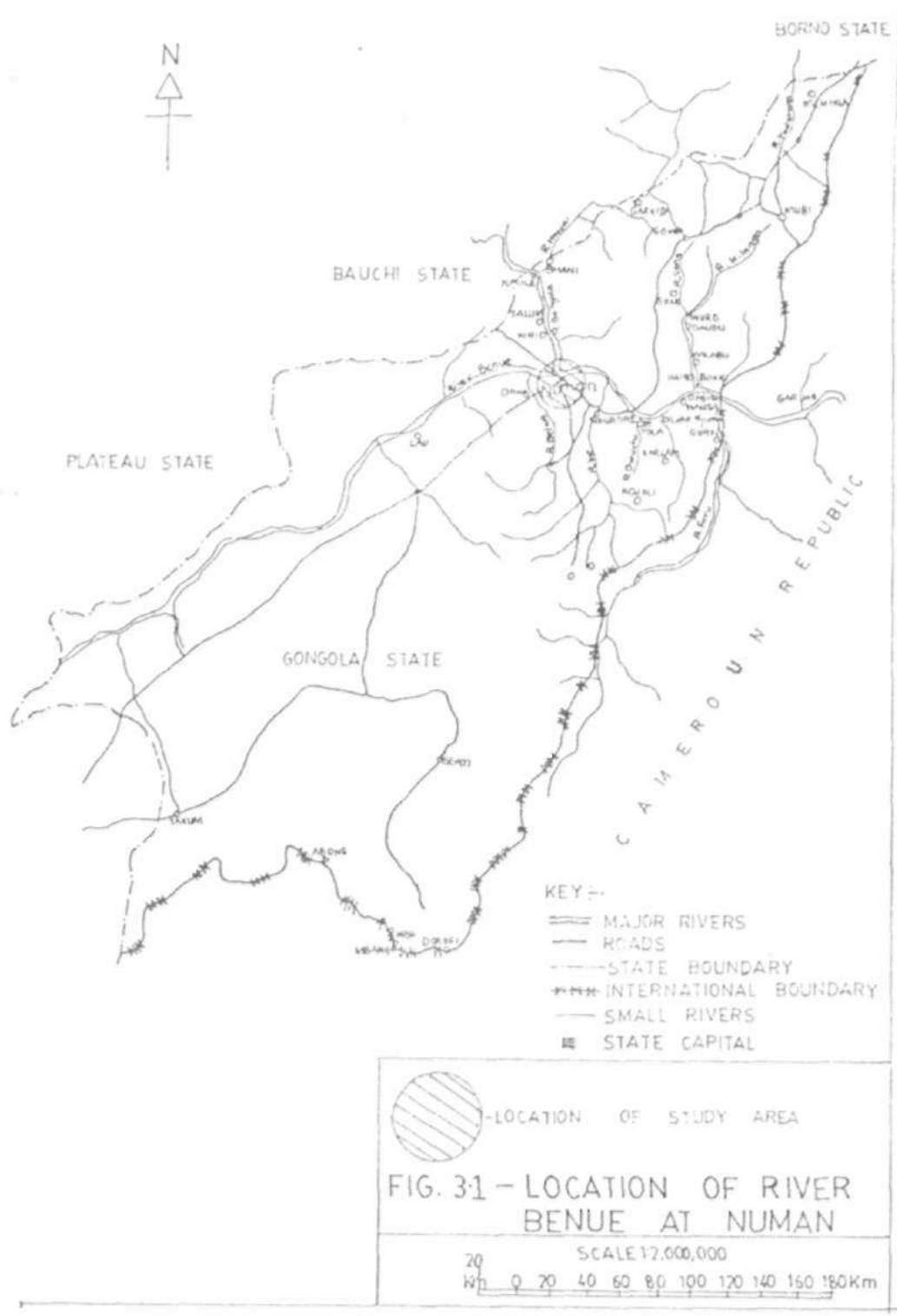


FIG. 3·1 – LOCATION OF RIVER BENUE AT NUMAN

3.2.2 was collected from Upper Benue River Basin and Rural Development Authority, Yola (UBRBRDA).

After the data collection there were disagreements of flood discharge data and rating curves for river Benue at Numan that came from different departments. Certain steps were taken to solve the problem acruing to each hydrologic data, and adopt the most suitable. The steps taken in carrying out these entails visiting various hydrological departments and coming home to analyse these data. These were done as follows:

3.2.2 COLLECTION, ANALYSIS AND ADOPTION OF DATA

The first department visited was UBRBRDA, and the data collected from this department include:

(i) Fig. 3.1 - Showing location of river Benue at Numan;

(ii) Appendix I - Showing the construction of a dam in 1982 at Lagdo in the Camerouns;

(iii) Table 3.1 - Showing initial flood discharges for river Benue at Numan for the period (1955-1980).

Use were made of these informations collected from this department as follows:Fig. 3.1 was only used showing the location of river Benue at Numan as aforementioned in Section 3.2.1. The use of Appendix I is to be

mentioned later in this section after visitation of other departments because it constitute a total obstruction to the river flowing capacity and no design will be possible without its consideration. Table 3.1 containing initial flood discharge for river Benue at Numan was not used because the flood discharges in cumecs for the years 1978, 1979 and 1980 and also for the years 1971 and 1972 were continuously $\geq$5600 and 2604 respectively. This data though collected by UBRBRDA is from Inland Waterways Division, Lokoja Branch (IWD). This in hydrology can not be used for any design and it became necessary that a good source of data, more reliable and off this error be resorted to.

Better results were obtained from the Federal Department of Water Resources, Mando Road, Kaduna, for the period (1955-1980). This data though collected by the IWD was used by "Isu Associates (Nig.) Ltd. (Consulting Engineers and Planners)" through the Federal Department of Water Resources, Mando Road, Kaduna and is shown in Table 3.2. Though it provided the solution to discharges but there was a problem of not showing any daily water level. This created a problem of not being able to correlate their rating curve shown in Fig. 3.2 with their discharges. Secondly, there were no points on this rating curve. This rating curve was to be compared with the ones plotted from data collected from IWD, to find whether they have the same datum or are any place the same.

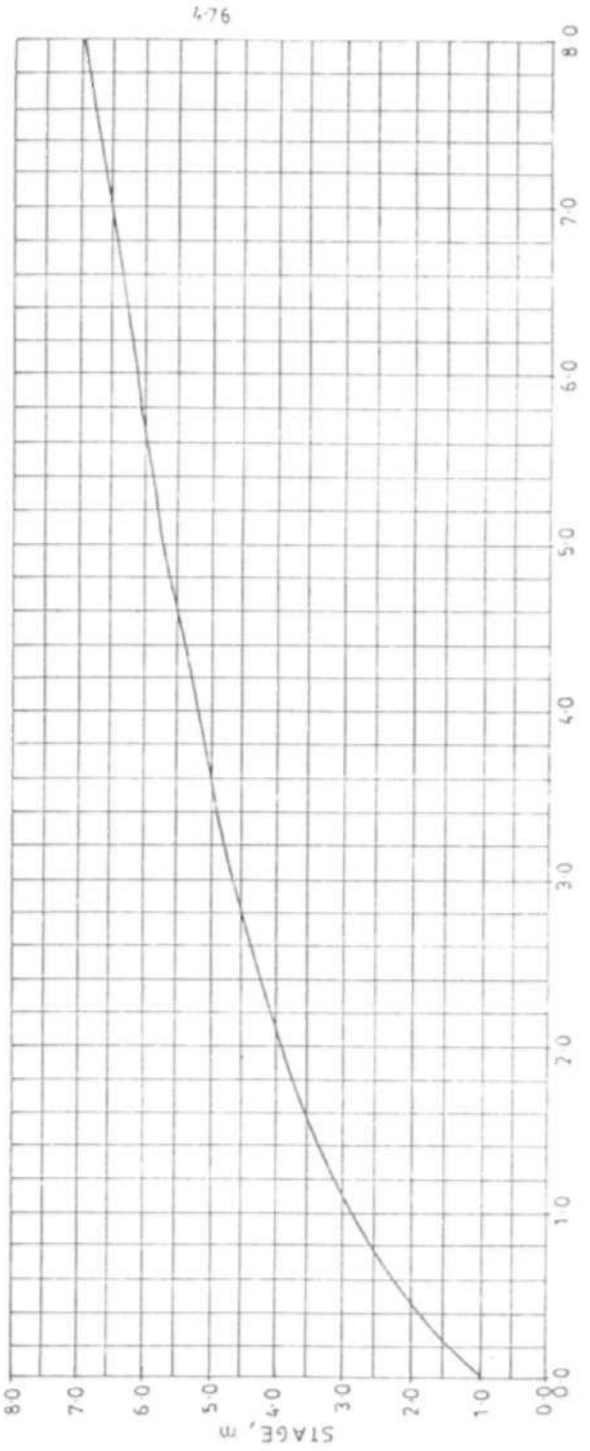
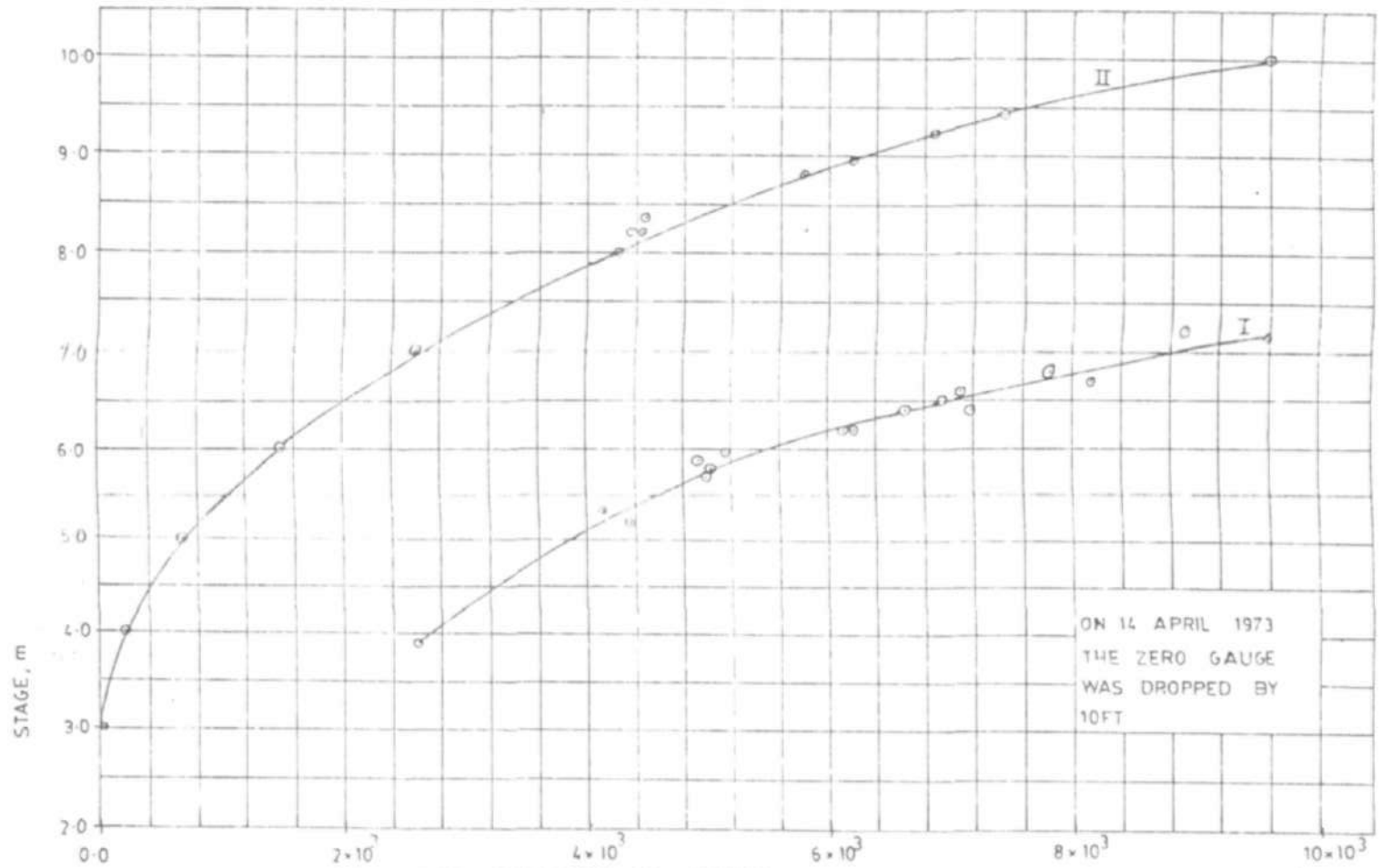


FIG. 3:2— STAGE—DISCHARGE RATING CURVE FOR RIVER AT NUMAN FOR THE PERIOD(1955–1980) FROM FED. DEPT. OF WATER RESOURCES, MANDO ROAD, KADUNA.



PEAK FLOOD DISCHARGES IN CUMECS

FIG. 3·3 — STAGE—DISCHARGE RATING CURVES FOR RIVER BENUE AT NUMAN FOR THE PERIODS (1955—1972) AND (1973—1987) FROM INLAND WATERWAYS DIVISION, LOKOJA BRANCH, FOR NUMBERS I & II RESPECTIVELY.

From information collected from IWD, there were problems because there was only one rating table available and also some of the readings from the available data were discovered to be in feet instead of metres. The main problem with this rating table is that it has for a maximum water level of 8.03m it gives a discharge of 4310 cumecs and as such all readings of water level above 8.03m were given a discharge of greater than 4310 (>4310) cumecs. This cannot be used for any hydrological design and their came up an urgent and dare need to plot the available stage-discharge in that station. This was done using Tables 3.3 and 3.4 for rating curves numbers I and II of Fig. 3.3 respectively following the work of Wilson (1981) who defined rating curve as a graph drawn connecting the water level elevation, or stage of a river channel at a certain-section with the corresponding discharge at that section. From the results of the plotting it was discovered that the IWD, has two rating curves plus the rating table.

The first rating curve, No 1, was for the period (1955-1972) before the zero gauge was dropped by 10ft on 14 April, 1973, to forestall negative readings. Moreover, the information collected within this period were the ones in cubic feet per second. So they were converted into cumecs before being used for plotting as shown in rating curve, number 1 of Fig. 3.3, and also Table 3.3. Table 3.5 containing

stage-discharge for river Benue at Numan for the period (1955-1987) is obtained from this department. The next thing becomes to make use of Appendix I.

From information collected from UBRBRDA, it was discovered that a dam had been constructed on river Benue at Lagdo in the Camerouns in 1982 and became fully operational in 1984 as shown in Appendix I. This dam which information was given in a seminar paper presented at Ilorin, by Normand (1984) is said to have a live storage capacity of 4700 million cubic metres. It has indeed added to the low flood discharge capacity of this river for the period (1982-1987). Based on this, the data was divided into two major sections as shown in Table 3.5. The first section, 'B', implying basic data before impoundment of the rifer is for the period (1955-1980), the additional section to this section, 'B_1', implying an additional data to the basic data before impoundment is for 1981 only, and the second section, 'A', is for the period (1982-1987) being actual data after impoundment of the river upon which the designs for peak flood discharges and risk of failure will be based.

3.2.3 CONCLUSION:

It is observed from Tables 3.1, 3.2, and 3.5 that Table 3.1 did depart much from other ones and that Table 3.2 stopped at 1980 while only Table 3.5 did

reach upto 1987 and was the only one with water level of the river. As a result of critical analysis of data the maximum discharges for the period (1955-1987) were estimated and are presented in Table 3.5.

Also observed from Figs 3.2 and 3.3 was that the rating curve used by Isu Associates (Nig.) Ltd (Consulting Engineers and Planners) was quite different both in datum and every other aspect from that used by Inland Waterways Division. From analysis the rating table fitted into rating curve number II of Fig. 3.3 and was adopted as the most recent and comprehensive.

TABLE 3.1 FLOOD DISCHARGE ON RIVER BENUE AT NUMAN FROM DATA OBTAINED FROM UPPER BENUE RIVER BASIN & RURAL DEVELOPMENT AUTHORITY, YOLA

NO	YEAR	DATE/MONTH	MAXIMUM DISCHARGES IN CUMECS
(1)	(2)	(3)	(4)
1.	1955	14/10	5946
2.	1956	21/9	5580
3.	1957	9/9	5370
4.	1958	28/9	4180
5.	1959	26/9	6957
6.	1960	24/9	7800
7.	1961	15/9	7068
8	1962	15/9	6212
9.	1963	5/9	9000
10.	1964	1/10	4992
11.	1965	10/9	5046
12.	1966	15/9	6135
13.	1967	10/9	4360
14.	1968	16/9	7135
15.	1969	4/9	9904
16.	1970	19/9	9595
17.	1971	10/9	2604
18	1972	10/9	2604
19.	1973	26/9	4412
20.	1974	22/8&16/9	4310
21.	1975	26/9	5200
22	1976	23/8	4500

TABLE 3.1 (contd.)

NO.	YEAR	DATE/MONTH	MAXIMUM DISCHARGES IN CUMECS
23	1977	9&10/9	4450
24	1978	13 & 14/9	>5600
25	1979	26-13/8&1-7/9	>5600
26	1980	28-31/8&1-10/9	>5600

TABLE 3.2 - FLOOD DISCHARGES ON RIVER BENUE AT NUMAN FROM DATA OBTAINED FROM FEDERAL DEPARTMENT OF WATER RESOURCES, KADUNA, FOR THE PERIOD (1955-1980)

NO	YEAR	DATE/MONTH	MAXIMUM DISCHARGES IN CUMECS
(1)	(2)	(3)	(4)
1.	1955	14/10	6657
2.	1956	22/9	5580
3.	1957	9/9	5370
4.	1958	28/9	4176
5.	1959	26/9	6987
6.	1960	24&25/9	7800
7.	1961	15/9	7088
8.	1962	15/9	6212
9.	1963	5/9	9000
10.	1964	1/10	4992
11.	1965	9-11/9	5046
12.	1966	15/9	6135
13.	1967	10/9	4360
14.	1968	15/9	7136
15.	1969	30&31/8	8165
16.	1970	19/9	9595
17.	1971	-	-
18.	1972	23/8	2604
19.	1973	26/9	4412
20.	1974	13/8	4603
21.	1975	12/9	9596
22.	1976	23/8	4500

TABLE 3.2 (contd.)

NO	YEAR	DATE/MONTH	MAXIMUM DISCHARGES IN CUMECS
23.	1977	14/9	6869
24.	1978	6/9	7421
25.	1979	3/9	5781
26.	1980	3/9	6137

- Implies no data for 1971.

TABLE 3.3 STAGE-DISCHARGE FOR RATING CURVE NO. 1
FOR THE PERIOD (1955-1972) FOR RIVER BENUE AT NUMAN FROM IWD, LOKOJA BRANCH

NO	YEAR	DATE/MONTH	WATER LEVELS IN METRES	MAXIMUM DISCHARGES IN CUMECS
(1)	(2)	(3)	(4)	(5)
1.	1955	14/10	6.41	6657
2.	1956	22/9	5.98	5580
3.	1957	9/9	5.89	5370
4.	1958	28/9	5.32	4176
5.	1959	26/9	6.51	6987
6.	1960	24&25/9	6.80	7800
7.	1961	15/9	6.62	7088
8.	1962	15/9	6.24	6212
9.	1963	5/9	7.12	9000
10.	1964	1/10	5.72	4992
11.	1965	9-11/9	5.75	5046
12.	1966	15/9	6.21	6135
13.	1967	10/9	5.19	4360
14.	1968	15/9	6.43	7136
15.	1969	30&31/8	6.76	8165
16.	1970	19/9	7.15	9595
17.	1971	-	-	-
18.	1972	23/8	3.93	2604

- implies no data for 1971.

TABLE 3.4: STAGE-DISCHARGE FOR RATING CURVE NO II FOR THE PERIOD (1973-1980) FOR RIVER BENUE AT NUMAN FROM IWD, LOKOJA BRANCH

NO	YEAR	DATE/MONTH	WATER LEVEL IN METRES	MAXIMUM DISCHARGES IN CUMECS
(1)	(2)	(3)	(4)	(5)
1.		FROM RATING TABLE	3.00	10.00
2.		"	4.00	195.00
3.		"	5.00	679.00
4.		"	6.00	1495.00
5.		"	7.00	2630.00
6.		"	8.00	4250.00
7.	1973	26/9	8.26	4412.00
8.	1974	13/9	8.38	4603.00
9.	1975	12/9	10.09	9596.00
10.	1976	23/9	8.20	4500.00
11.	1977	14/9	9.24	6869.00
12.	1978	6/9	9.43	7421.00
13.	1979	3/9	8.83	5781.00
14.	1980	3/9	8.97	6137.00
15.	1981	30-31/8&12-27/9	8.16	4480.00
16.	1982	12/8	7.95	4160.00
17.	1983	22/9	7.41	3280.00
18.	1984	30/9	6.84	2440.00
19.	1985	11/9	7.75	3840.00
20	1986	27/9	8.02	4240.00
21	1987	4/9	6.82	2420.00

Note: The maximum discharges from No 1 to 14 were used in plotting Fig. 3.2 while No 15 to 21 were read out from Fig. 3.2 using their daily water levels.

TABLE 3.5: STAGE-DISCHARGE FOR RIVER BENUE AT NUMAN DIVIDED INTO SECTIONS FROM INLAND WATERWAYS DIVISION, LOKOJA BRANCH

NO	SECT-ION	YEAR	DATE/MONTH	WATER LEVEL IN METRES	MAXIMUM DISCHARGES IN CUMECS
(1)	(2)	(3)	(4)	(5)	(6)
1.		1955	14/10	6.41	6657
2.		1956	22/9	5.98	5580
3.		1957	9/9	5.89	5370
4.		1958	28/9	5.32	4176
5.		1959	26/9	6.51	6987
6.		1960	24&25/9	6.80	7800
7.		1961	15/9	6.62	7088
8.		1962	15/9	6.24	6212
9.		1963	5/9	7.12	9000
10.		1964	1/10	5.72	4992
11.		1965	9-11/9	5.75	5046
12.	B	1966	15/9	6.21	6135
13.		1967	10/9	5.19	4360
14.		1968	15/9	6.43	7136
15.		1969	30&31/8	6.76	8165
16.		1970	19/9	7.15	9595
17.		1971	-	-	-
18.		1972	10/9	3.93	2604
19.		1973	26/9	8.26	4412
20.		1974	13/9	8.38	4603
21.		1975	12/9	10.09	9596
22.		1976	23/8	8.20	4500
23.		1977	14/9	9.24	6869

TABLE 3.5 (CONTD.)

NO	SEC-TION	YEAR	DATE/MONTH	WATER LEVEL IN METRES	MAXIMUM DISCHARGES IN CUMECS
24		1978	6/9	9.43	7421
25		1979	3/9	8.83	5781
26.		1980	3/9	8.97	6137
27	B_1	1981	30-31/8& 12-27/9	8.16	4480
28.		1982	12/8	7.95	4160
29.		1983	22/9	7.41	3280
30	A	1984	30/9	6.84	2440
31		1985	11/9	7.75	3840
32.		1956	27/9	8.02	4240
33.		1987	4/9	6.82	2420

B implies Basic data, B_1 implies additional data to basic data, A implies Actual data after impoundment, and - implies no data for 1971.

3.3 SELECTION OF TYPES OF DISTRIBUTION

3.3.1 INTRODUCTION

The "Two-parameter lognormal distribution" and the "Gumbel extreme value distribution" were adopted for the study. They are two-parameter types of distribution and they are used following the work of three researchers: Bodo and Unny (1976) recently in their example of "Prior Distributions" recommended the two-parameter types of distribution and gave their reason that "three-parameters forms were not studied because of the difficulties introduced by the third parameter". Russell (1982), most recently in his work recommended and used the two distributions and the reason for his recommendation/usage is that the two-parameter type of distribution has less problems associated with it than the three-parameter type. As a result, analyses were carried out on these two distributions for the two periods (1955-1980) and (1982-1987).

3.3.2 ANALYSIS OF DATA FOR THE PERIOD (1955-1987)

3.3.2 .1 INTRODUCTION:

In this section an initial probability plotting test was to be carried out for the two distributions and if they accept the sample data there will be no need for therigorous goodness-of-fit test. This first probability plotting test is carried out as follows:

3.3.2.2 PROBABILITY PLOTTING FOR GUMBEL EXTREME VALUE DISTRIBUTION AND TWO PARAMETER LOGNORMAL DISTRIBUTION FOR THE PERIOD (1955-1980)

The procedure in carrying out this test is to use eqn (3.1) to obtain the common probability, P_r, shown in Col. (4) of Table 3.6. This probability is multiplied with 100 to get the plotting positions to be used on the abscissa as shown in Col. (5) of Table 3.6. Equation (3.1) is as follows:

$$P_r = \frac{i - \frac{1}{2}}{n} \quad .. \quad .. \quad (3.1)$$

In which n = sample size;

i = order of samples;

P_r = probability.

Ever before this the flood discharges, $Q^{(i)}$, must have been arranged in ascending order with the lowest discharge at the n^{th} position and the highest discharge at the first position, as shown in Col. (2) of Table 3.6. The next thing, is to do the plotting for both the Gumbel

TABLE 3.6 - PROBABILITY TEST FOR FLOOD DISCHARGES FOR THE PERIOD (1955-1980)

Order (i)	$Q^{(i)}$	$Y^{(i)}=\ln Q^{(i)}$	$P_r = \frac{i - \frac{1}{2}}{n}$	$100P_r$
	CUMECS	CUMECS		
(1)	(2)	(3)	(4)	(5)
1.	9596	9.17	0.02	2
2.	9595	9.17	0.06	6
3.	9000	9.10	0.10	10
4.	8165	9.01	0.14	14
5.	7800	8.96	0.18	18
6.	7421	8.91	0.22	22
7.	7136	8.87	0.26	26
8.	7088	8.87	0.30	30
9.	6987	8.85	0.34	34
10.	6869	8.83	0.38	38
11.	6657	8.80	0.42	42
12.	6212	8.73	0.46	46
13.	6137	8.72	0.50	50
14	6135	8.72	0.54	54
15.	5781	8.66	0.58	58
16.	5580	8.63	0.62	62
17.	5370	8.59	0.66	66
18.	5046	8.53	0.70	70
19.	4992	8.52	0.74	74
20	4603	8.43	0.78	78
21	4500	8.41	0.82	82
22.	4412	8.39	0.86	86
23.	4360	8.38	0.90	90
24.	4176	8.34	0.94	94
25.	2604	7.86	0.98	98

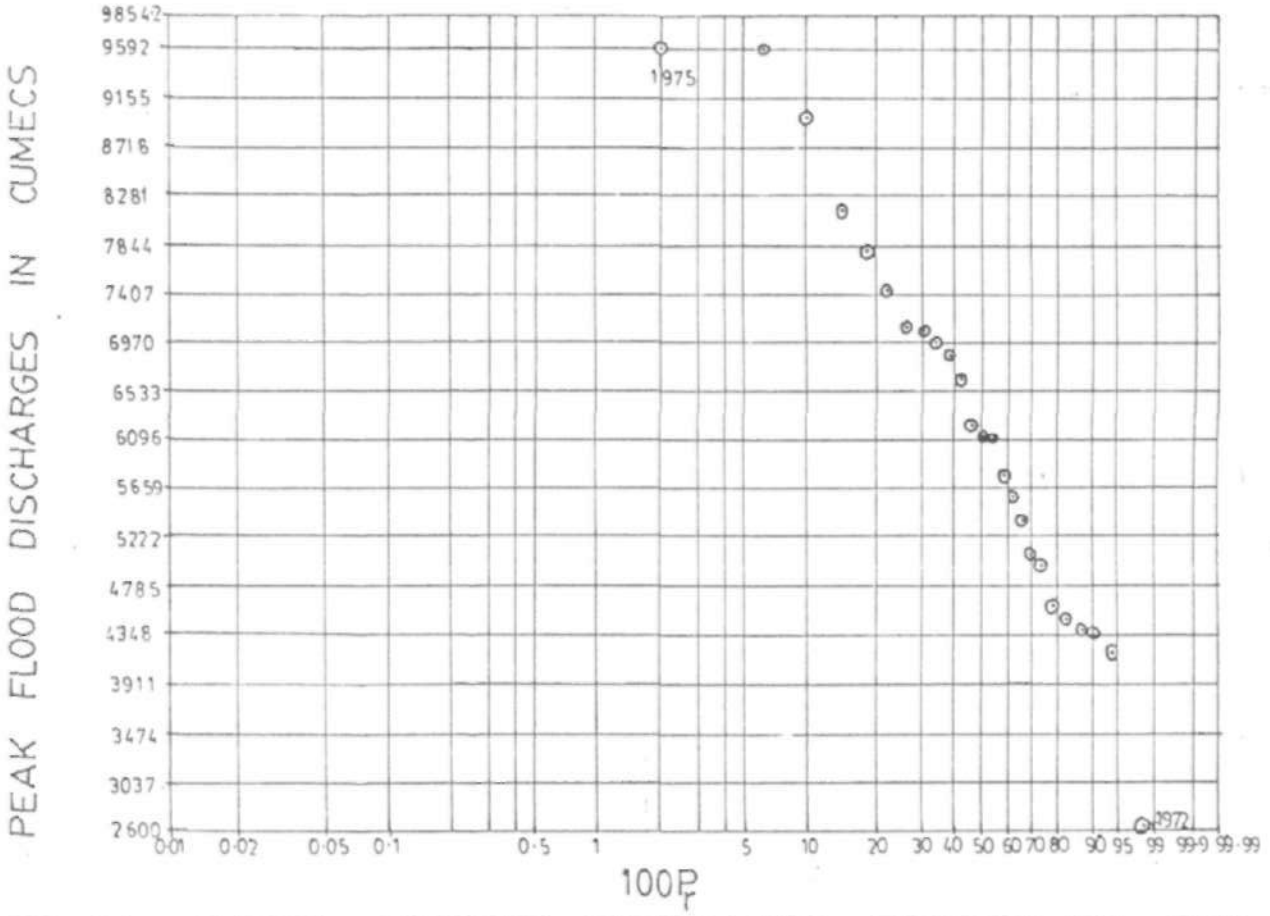


FIG. 3·4 — GUMBEL EXTREME DISTRIBUTION PROBABILITY FOR RIVER BENUE AT NUMAN FOR THE PERIOD (1955–1980)

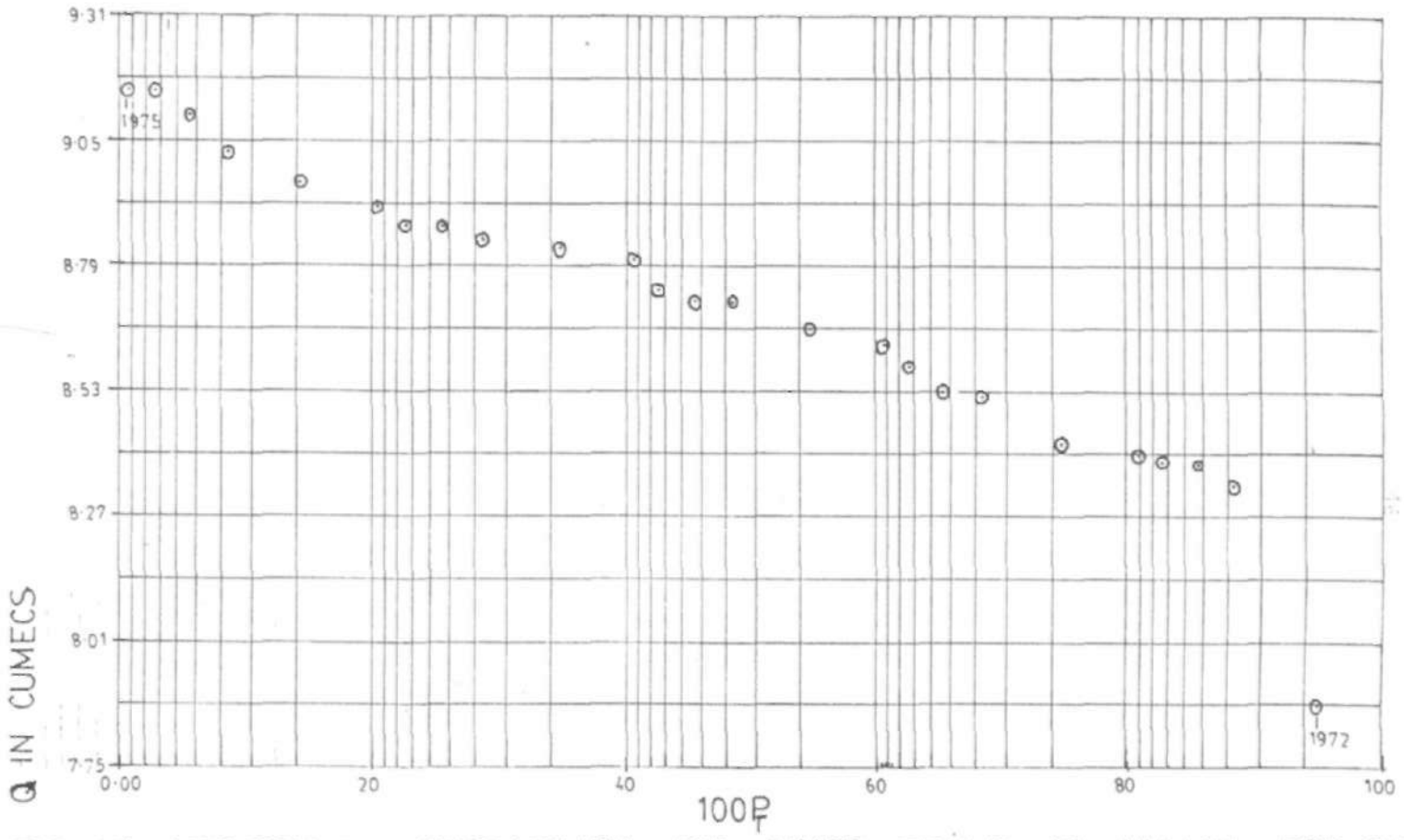


FIG. 3·5 — LOGNORMAL PROBABILITY FOR RIVER BENUE AT NUMAN FOR THE PERIOD (1955 — 1980)

extreme value distribution and the two parameter lognormal distribution using their respective probability papers. It is a plot of discharge, Q, on the ordinate from Col.(2) of Table 3.6 against the probability that has been multiplied by 100 on the abscissa, from Col. (5) of the same Table 3.6.

The plotting for Gumbel distribution is shown in Fig. 3.4 which is on Gumbel extreme distribution probability paper, and the plotting for two parameter lognormal distribution is shown in Fig. 3.5 which is on ordinary lognormal probability paper.

3.3.2.3 SUMMARY

A critical viewing of Fig. 3.4, will show the plotting of the sample data to have a fair straight line slanting from left to right and unbounded too, and as such the sample data is said to follow the extreme value distribution.

The plotting from Fig 3.5 is seen to be a bit in a straight line slanting from left to right like that of the Gumbel and unbounded too at the two extremes similar to the findings of Fisher and Tipett in Kite (1978). Hence, the sample data is concluded to fit the two parameter Lognormal distribution.

3.3.3 ANALYSIS OF DATA FOR THE PERIOD (1982-1987)

3.3.3.1 INTRODUCTION:

In this section like in the first section an initial probability plotting was carried out but if it does not accept the sample data, other tests will then be carried out for the two distributions. How the probability plotting tests for the two distributions were carried out are as follows:

3.3.3.2 PROBABILITY PLOTTING FOR GUMBEL EXTREME VALUE DISTRIBUTION AND TWO PARAMETER LOGNORMAL DISTRIBUTION

The procedure in carrying out this test is to first of all arrange the flood discharges from Col.(6) of Table 3.5 for the period (1982-1987) in ascending order with the lowest discharge at the n^{th} position and the highest discharge at the first position, as shown in Col.(2) of Table 3.7. The next thing is to compute the common probabilities (P_r) for carrying out the test.

Using eqn (3.1) and substituting values accordingly, the common probabilities were obtained as shown in Col. (4) of Table 3.7. This probability is multiplied by 100 to get the plotting positions. to be used on the abscissa as shown in Col. (5) of Table 3.7. The next thing is to do the plotting for both the Gumbel extreme distribution probability and the Two parameter lognormal

distribution probability using their respective probability papers.

As mentioned in section 3.3.2.2 it is a plot of discharge, Q, on the ordinate from Col. (2) of Table 3.7 against the probability, P_r, that has been multiplied by 100 on the abscissa, from Col. (5) of Table 3.7.

TABLE 3.7 : PROBABILITY PLOTTING TEST FOR FLOOD DISCHARGES FOR THE PERIOD (1982-1987)

Order i	$Q^{(i)}$	$Y^{(i)} = \ln Q^{(i)}$	$P_r = \frac{(i - \frac{1}{2})}{n}$	$100P_r$
	CUMECS	CUMECS		
(1)	(2)	(3)	(4)	5
1	4240	8.35	0.083	8.3
2	4160	8.33	0.250	25.0
3	3840	8.25	0.417	41.7
4	3280	8.10	0.583	58.3
5	2440	7.80	0.750	75.0
6	2420	7.79	0.917	91.7

The plotting for Gumbel distribution is shown in Fig. 3.6 which is on Gumbel extreme distribution probability paper and the plotting for two parameter lognormal distribution is shown in Fig. 3.7 which is on Ordinary lognormal probability paper.

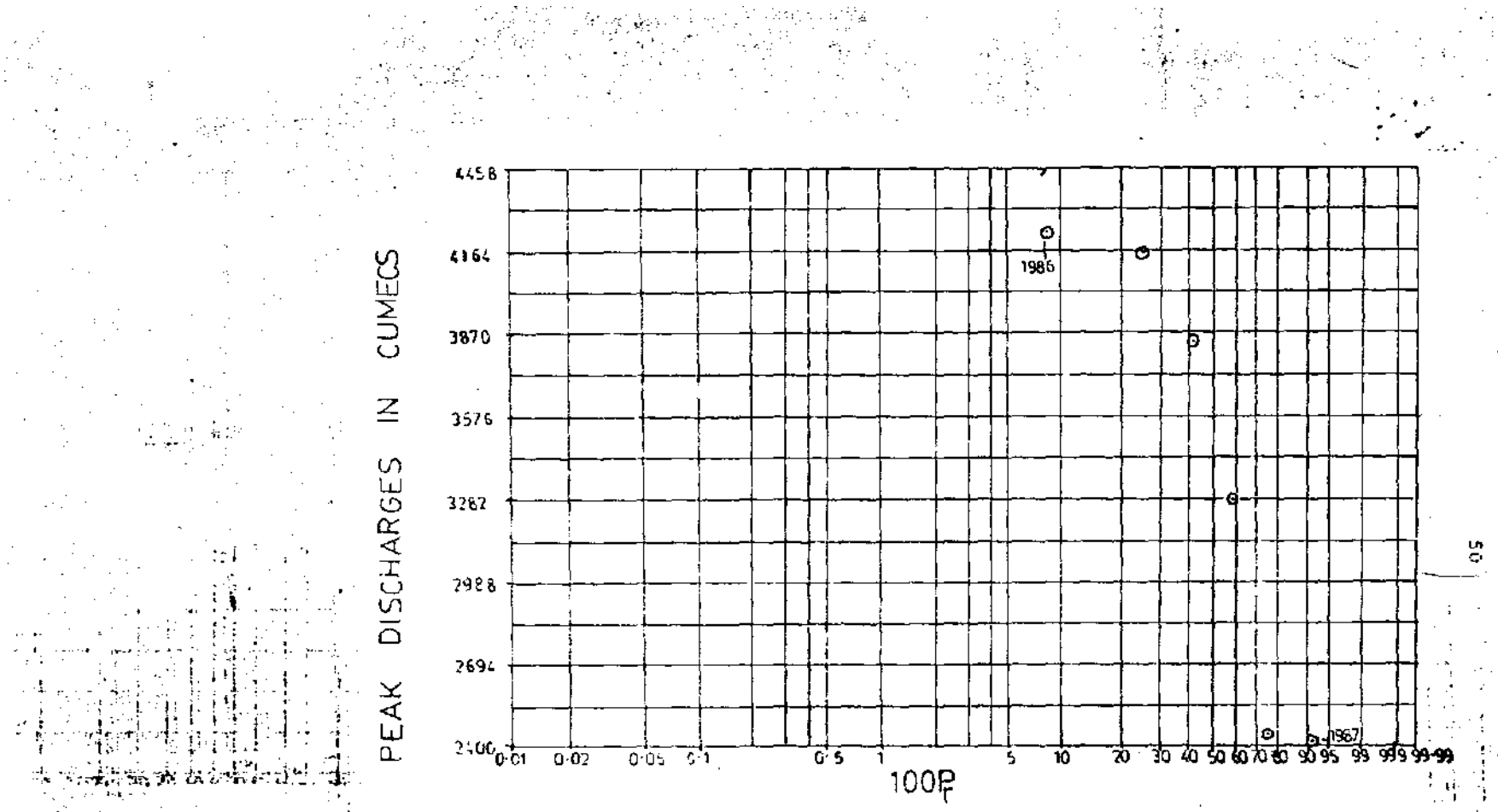


FIG. 3·6 — GUMBEL EXTREME DISTRIBUTION PROBABILITY FOR RIVER BENUE AT NUMAN FOR THE PERIOD(1982-1987)

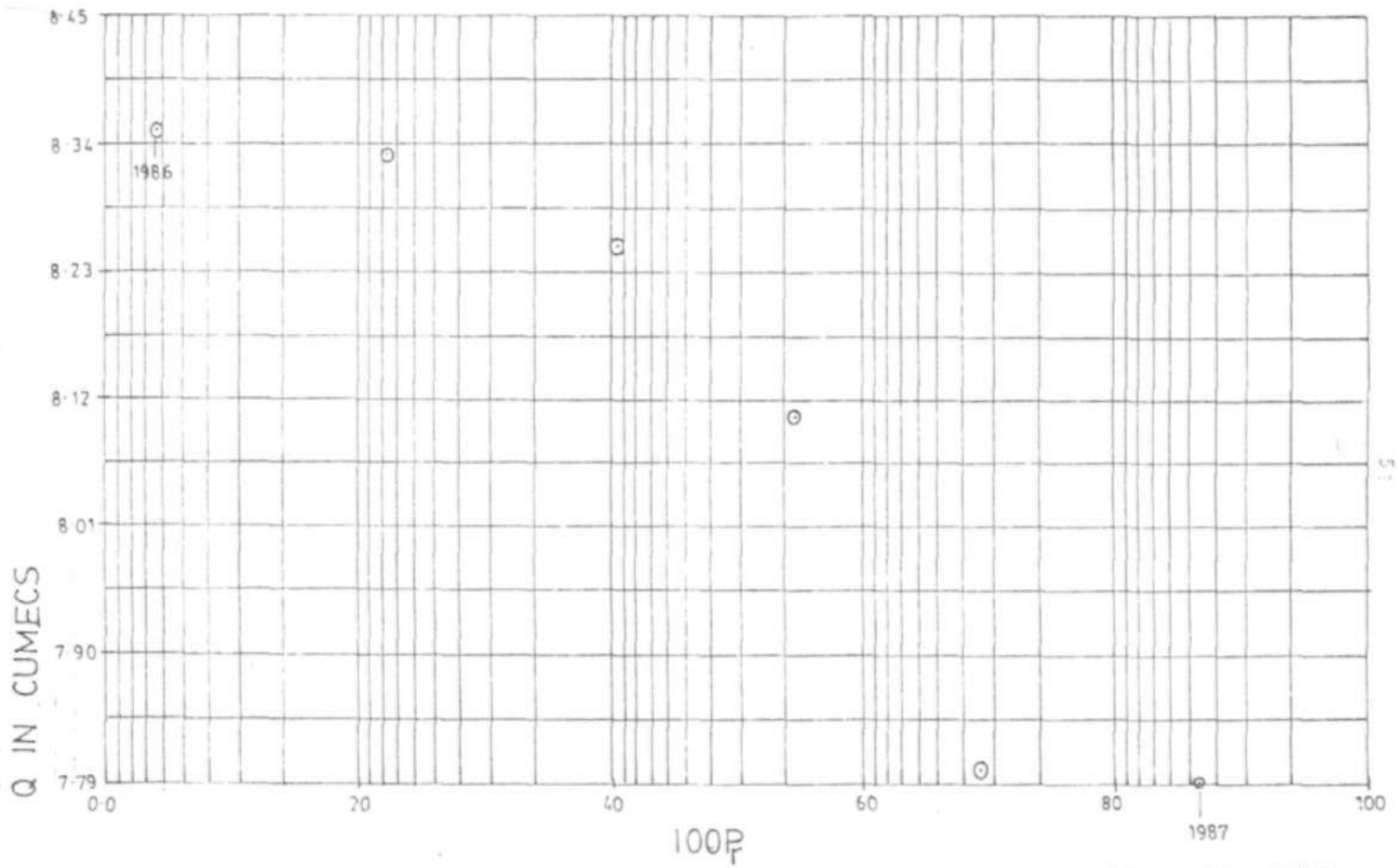


FIG. 3·7 — LOGNORMAL PROBABILITY FOR RIVER BENUE AT NUMAN FOR (1982 — 1987)

Following the work of Fisher and Tipett the two plottings in Figures 3.6 and 3.7 though unbounded at the two extremes but none seems to have a straight line slanting from left to right and as such does not fit into the sample data and immediately calls for the following other tests to be carried out for the two distributions as follows:

i) Chi-squared test statistics for both distributions
ii) Kolmogorov-Smivnov (K-S) test statistics for both distributions.

However, to be able to carry out these other tests the first thing that need to be done is to estimate the mean and standard deviation of these period (1982-1987) as the basic parameters of the tests for the distributions.

3.3.3.3. MEAN AND STANDARD DEVIATION VALUES FOR THE VARIABLE, Q, FOR THE PERIOD (1982-1987)

Employing eqn (3.2) and substituting the flood discharges from Col. (6) of Table 3.5 for the period (1982-1987), the mean value is obtained as shown in Col. (2) of Table 3.8. The formula is as follows:

$$M_Q = \frac{1}{n} \sum_{i=1}^{n} Q_i \quad \ldots \quad \ldots \quad (3.2)$$

In which Q_i = the given sample data

n = size of data

Similarly, the formula for the most used parameter of dispersion called the standard deviation which was used for flood discharges lasting the period (1982-1987) from Col. (6) of Table 3.5. The result of the computation is shown in Col. (3) of Table 3.6 for the estimation of the standard deviation value. The formula is as follows:

$$\sigma_Q = \left(\frac{\sum_{i=1}^{i=n} (Q_i - M_Q)^2}{n} \right)^{\frac{1}{2}} \quad \ldots\ldots \quad (3.3)$$

in which σ_Q = standard deviation for the variable Q_i

After calculating the mean values and the values of the standard deviation for the period (1982-1987) using eqns (3.2) and (3.3) respectively, a next step was taken to derive the natural logarithm of these values.

To obtain the natural logarithm of the mean value, eqn (3.4) was used on the mean value shown in Col. (2) of Table 3.8, the result of which is shown in Col.(4) of Table 3.8 and the formula is as follows:

$$M_Y = \ln M_Q \quad \ldots \quad \ldots \quad (3.4)$$

In which M_Y = natural logarithm of the mean values, M_Q. The procedure towards calculating the natural logarithm of the standard deviation value is to first of all, compute the coefficient of variation. This was done using eqn (3.5) on the values of the mean and standard deviation from Cols (2) and (3) of Table 3.8 respectively, the result of which is shown in Col. (5) of Table 3.8. The formula is as follows:

$$V_Q = \frac{\sigma_Q}{M_Q} \quad \ldots \quad \ldots \quad (3.5)$$

In which V_Q = coefficient of variation for the variable, Q.

After computing the coefficient of variation shown in Col.(5) of Table 3.8, eqn (3.6) was employed from Tung and Larry (1980) to obtain the natural logarithm of the standard deviation value shown in Col. (6) of Table 3.8 and the formula is as follows:

$$\sigma_Y = (\ln (V_Q^2 + 1))^{\frac{1}{2}} \quad .. \quad .. \quad (3.6)$$

In which σ_Y = natural logarithm of the standard deviation value for the variable, Q.

TABLE 3.8 : MEAN AND STANDARD DEVIATION VALUES FOR THE PERIOD (1982-1987)

No	M_Q from eqn (3.2)	σ_Q from eqn (3.3)	M_Y from eqn (3.4)	V_Q from eqn (3.5)	σ_Y from eqn (3.6)
	CUMECS	CUMECS	CUMECS		
(1)	(2)	(3)	(4)	(5)	(6)
1	3397	750	8.131	0.221	0.218

In subsequent sections the actual tests are done starting with the PDF test for Gumbel extreme value distribution.

3.3.3.4 GUMBEL EXTREME VALUE DISTRIBUTION

3.3.3.4.1 INTRODUCTION

To know whether the data available did follow the Type 1 extreme value distribution called the Gumbel extreme value distribution two further tests were carried out as aforementioned and in the following order:

(i) Chi-squared goodness-of-fit test statistics

(ii) K-S test statistics.

3.3.3.4.2 CHI-SQUARED GOODNESS-OF-FIT TEST STATISTICS FOR GUMBEL EXTREME VALUE DISTRIBUTION FOR THE PERIOD (1982-1987)

The Chi-squared goodness-of-fit test statistics, which is related to the histogram deviations from the predicted values, has a traditional formula given as:

$$\chi^2 = \sum_{j=1}^{k} \frac{(O_j - E_j)^2}{E_j} \quad \ldots\ \ldots \quad (3.7)$$

In which χ^2 = the observed value of the Chi-squared goodness-of-fit test statistics;

O_j = number of observed values;

E_j = normal expected number of observations;

k = number of categories considered;

j = order of categories.

This is a test involving hypothesis testing. Hypothesis testing is a conventional procedure for drawing simple conclusions from observed statistical data (Benjamin and Cornell (1970)). In this test the observed value, χ^2, of the Chi-squared test statistics is to be compared with its critical value, D_J, obtained from table at certain significance level, α, and degrees of Freedom, ν. If at a chosen significance level and particular degree of freedom the critical value of the Chi-squared test statistics is greater than its observed value, the sample data is said to fit into the distribution. This hypothesis test that

it fits into the distribution is called the null hypothesis and the test that it does not fit into the distribution is called the alternative hypothesis.

The procedure in the test is to estimate the basic parameters of the observed value of the Chi-squared test statistics in eqn (3.7) starting with the number of observed value.

The process in calculating the "number of observed values" is to first and foremost put the flood discharges into categories or classes based on the limit that $Q_1 \leq Q \leq Q_2$ as shown in Col. (1) of Table 3.9. Where Q_1 is the first interval and Q_2 is the second interval. Based on this limit, the flood discharges from Col. (6) of Table 3.5 for the period (1982-1987) were tallied into these class intervals resulting in the observed values shown in Col. (2) of Table 3.9. The next basic parameters to be computed is the "normal expected number of observations".

To obtain the normal expected number of observations, firstly, reduce all flood discharges standing as class intervals into reduced variate, W, using eqn (3.8) i.e.

$$W = (Y - U_G)\,\alpha_G \quad \ldots \quad \ldots \quad (3.8)$$

In which W = reduced variate;

$$Y = \ln Q \quad \ldots \quad \ldots \quad \ldots \quad (3.9)$$

With the obtained reduced variate, W, go over to Appendix V and get the CDF, $F_W(w)$. Substituting the obtained $F_W(w_i)$ into eqn (3.10) accordingly, the normal expected number of observation for each class interval is found as shown in Col. (3) of Table 3.9. The eqn is as follows:

$$N\left[F_{ex}(Q_{2i}) - F_{ex}(Q_{1i})\right] = N\left[F_W(w_{2i}) - F_W(w_{1i})\right] = E_j \quad \ldots \quad \ldots \quad (3.10)$$

In which N = total number of samples of the distribution;

$F_{ex}(Q_{2i})$ = CDF for the second interval for each class;

$F_{ex}(Q_{1i})$ = CDF for the first interval for each class;

$F_W(w_{2i})$ = CDF for the reduced variate for the second interval for each class;

$F_W(w_{1i})$ = CDF for the reduced variate for the second interval for each class.

The next thing that need to be determined is the "normalized squared deviations".

By substituting the computed values of the number of observed values and the normal expected number of observations for each class interval into eqn (3.11) the normalized squared deviation values are obtained as shown in Col.(4) of Table 3.9. Equation (3.11) is

written as:

$$\text{Normalized squared deviations} = \frac{(O_j - E_j)^2}{E_j} \quad \ldots \quad (3.11)$$

TABLE 3.9 : HISTOGRAM INTERVALS FOR FLOOD DISCHARGES FOR GUMBEL EXTREME VALUE DISTRIBUTION FOR THE PERIOD (1982 - 1987)

Interval for Q	No of Observed values	Model: normal (8.131, 0.218^2) α_G = 5.88, U_G = 8.03 expected no. of observations from eqn (3.12)	Normalized squared deviations
CUMECS			
(1)	(2)	(3)	(4)
2400-3000	2	1.817	0.018
3000-3600	1	2.146	0.612
3600-4200	2	1.070	0.808
4200-4800	1	0.463	0.623
	6	5.496	2.061 = χ^2

The observed value of the Chi-squared test statistics is obtained as a summation of the normalized squared deviations to be 2.061 as shown in Col. (4) of Table 3.9.

Having found theobserved value of the Chi-squared test statistics, the next issue of importance becomes the critical value of the Chi-squared test statistics that would be used in comparing its observed value.

The critical value is obtained from Appendix III as follows:

$$D_1 = \chi^2_{\alpha,\nu} \quad \ldots \quad \ldots \quad (3.12)$$

In which α = significance level;

ν = degrees of freedom;

$$\nu = k - r - 1 \quad \ldots \quad \ldots \quad (3.12(a))$$

k = no. of categories considered;

r = no. of parameters estimated from the data;

D_1 is the critical value of the chi-squared goodness-of-fit test statistics estimated from Appendix III, using the 1 degrees of freedom and 95% significance level of test i.e.

$$D_1 = \chi^2_{0.95,1} = 3.84$$

Since the observed value of the Chi-squared (χ^2)=2.061 is less than the critical value of the Chi-squared $(D_1) = \chi^2_{0.95,1} = 3.84$ accept the null hypothesis and conclude that the data did follow the Gumbel extreme value distribution. This is a one tail test. In the next section the K-S test statistics will be dealt with.

3.3.3.4.3 K-S TEST STATISTICS FOR GUMBEL EXTREME VALUE DISTRIBUTION FOR THE PERIOD (1982-1987)

This is a second quantitative goodness-of-fit test statistics. It concentrates on the deviations between the hypothesized cumulative distribution function, $F_Q(Q^{(i)})$ and the observed cumulative histogram. This test involves hypothesis testing. In this test the observed value, $D_{2(max)}$, is to be compared with the critical value, $D_{n,\alpha}$, of the K-S test statistics obtained from table at certain significance level, α, and given sample size, n. If at a chosen significance level, α, and given sample size, n, the critical value of the K-S test statistics is greater than its observed value the sample data is said to fit into the distribution. This hypothesis test that it fits into the sample data is called the null hypothesis and the test that it does not fit into the distribution is called the alternative hypothesis.

The process in finding the observed value of the K-S test statistics is started by given the equation for the cumulative histogram as follows:

$$F^*(Q^{(i)}) = \frac{i}{n} \quad .. \quad .. \quad .. \quad (3.13)$$

In which $Q^{(i)}$ is the largest observed value in the random variable, in the random sample of size, n. On substitution of values into eqn (3.13) the observed cumulative histograms are obtained as shown in Col. (3) of Table 3.10.

The hypothesized cumulative distribution function which was computed by first, reducing the variables, $Q^{(i)}$, using eqn (3.8) and with the obtained reduced variates, W, go to Appendix V and get the hypothesized cumulative distribution function, $F_W(w^{(i)})$. This given as eqn (3.14) is as follows:

$$F_Q(Q^{(i)}) = F(Q^{(i)}) = F_W(w^{(i)}) \quad \ldots\ldots (3.14)$$

In which $F(Q^{(i)})$ = CDF for the variable, $Q^{(i)}$;

$F_W(w^{(i)})$ = CDF for the reduced variate, $W^{(i)}$

On substitution of values into eqn (3.14) in the order aforementioned, the hypothesized cumulative distribution functions were obtained as shown in Col. (4) of Table 3.10. With this ready, the next thing becomes to calculate the deviations, D_2, between the hypothesized CDF and the observed cumulative histogram.

To compute the observed value ($D_{2(max)}$) of the K-S test statistics eqn (3.15) is employed. This $D_{2(max)}$, is described as the largest of the absolute values of the n differences between the hypothesized CDF from Col. (4) of Table 3.10 and the observed cumulative histogram from Col. (3) of Table 3.10 (sometimes called the empirical CDF), evaluated as the observed values in the sample shown in Col (5) of Table 3.10. This eqn (3.15) is given as follows:

$$D_2 = \max_{i=1}^{n} \left| F^*\left(Q^{(i)}\right) - F_Q\left(Q^{(i)}\right) \right| \quad \ldots \qquad (3.15)$$

The result of the absolute deviations between $F^*(Q^{(i)})$ and $F_Q(Q^{(i)})$ are shown in Col. (5) of Table 3.10.
The largest discrepancy in Col. (5) of Table 3.10 is:

$$\left|F^*(Q^{(i)}) - F_Q(Q^{(i)})\right| = \left|F^*(2440) - F_Q(2440)\right| = 0.312$$

and is the observed value of the K-S test statistics.
The next thing is to obtain the critical value of the K-S test statistics using $D_{n,\alpha}$ in eqn (3.16).

TABLE 3.10: K-S TEST STATISTICS FOR GUMBEL EXTREME VALUE DISTRIBUTION FOR FLOOD DISCHARGES FOR THE PERIOD (1982-1987)

Order (i)	Observed $Q^{(i)}$ CUMECS	$F^*(Q^{(i)})$ from eqn. (3.13)	$F_Q(Q^{(i)}(8.131, 0.218^2)$ α_G=5.88, U_G=8.03 from eqn (3.14)	D_2 from eqn (3.15)
(1)	(2)	(3)	(4)	(5)
1	2420	0.167	$F(2420)=F_W(-1.402)=0.017$	0.150
2	2440	0.333	$F(2440)=F_W(-1.354)=0.021$	0.312*
3	3280	0.500	$F(3280)=F_W(0.386) = 0.507$	0.007
4	3840	0.667	$F(3840)=F_W(1.313)= 0.764$	0.097
5	4160	0.833	$F(4160)=F_W(1.783) =0.845$	0.012
6	4240	1.000	$F(4240)=F_W(1.895) = 0.860$	0.140

Eqn (3.16) is written as:

$$D_{n,\alpha} > D_{2(max)} \quad \ldots \quad .. \quad (3.16)$$

In which n = sample size;

α = significance level;

$D_{n,\alpha}$ = critical value of K-S test statistics;

$D_{2(max)}$ = observed value of K-S test statistics.

For an α level of 5 percent and a sample size of 6, the critical value of the K-S test statistics from Appendix IV is $D_{6,0.05} = 0.521$, greater than $D_{2(max)} = 0.312$. Therefore, the null hypothesis need not be rejected at this significance level. The variable, Q, is said to follow the Gumbel extreme value distribution. This is a one tail test. In the next section tests for two-parameter lognormal distribution will be treated.

3.3.3.5 TWO PARAMETER LOGNORMAL DISTRIBUTION

3.3.3.5.1 INTRODUCTION

In this distribution, two goodness-of-fit test statistics were carried out and they are:

(i) Chi-squared goodness-of-fit test statistics

(ii) K-S goodness-of-fit test statistics.

The steps taken in carrying out these tests starting with the chi-squared goodness-of-fit test statistics are as follows:

3.3.3.5.2 CHI-SQUARED GOODNESS-OF-FIT TEST STATISTICS FOR TWO PARAMETER LOGNORMAL DISTRIBUTION FOR THE PERIOD (1982-1987).

The procedure in carrying out this test was to first of all put the flood discharges from Col.(6) of Table 3.5 for the period (1982-1987) into class intervals as shown in Col.(6) of Table 3.11. The flood discharges from Col. 6 of Table 3.5 for this period (1982-1987) were then tallied into these class intervals, based on the limit that $Q_1 \leqslant Q \leqslant Q_2$. Where Q_1 is the first class interval in each class interval and Q_2 is the second class interval in each class interval. The results of the tallies are the number of observed frequencies (values) shown in Col. (2) of Table 3.11. Having obtained the number of observed values, the next thing becomes to get the "normal expected number of observations".

The normal expected number of observations was obtained by first of all standardizing the flood discharges standing as class intervals from Col. (1) of Table 3.11 using eqn (3.17), i.e.

$$U = \frac{Y - M_Y}{\sigma_Y} \qquad .. \qquad .. \qquad (3.17)$$

In which U = standardized variable.

with the standardized variables, U, go over to Appendix II and get the CDF for the standardized variable, U, written $F_u(U)$. Substituting all the obtained standardized CDF in place of the reduced CDF in eqn (3.10)

accordingly the normal expected number of observations was found as shown in Col. (3) of Table 3.11.

The next basic parameter to be computed towards obtaining the observed value of the chi-squared goodness-of-fit test statistics are the values of the "normalized squared deviations".

TABLE 3.11 ; HISTOGRAM INTERVALS FOR TWO PARAMETER LOGNORMAL DISTRIBUTION FOR FLOOD DISCHARGES FOR THE PERIOD (1982-1987)

Interval for Q	No of observed values	Model: normal $(8.131, 0.218^2)$ expected no of observations from eqn (3.10)	Normalized squared deviation
CUMECS			
(1)	(2)	(3)	(4)
2400-3000	2	1.370	0.290
3000-3600	1	1.932	0.450
3600-4200	2	1.380	0.279
4200-4800	1	0.653	0.184
	6	5.335	$1.203=\chi^2$

The normalized squared deviation values of each category were computed by employing eqn (3.7) and substituting the values of the number of observations from Col.(2) of Table 3.11 with the corresponding normal expected number of observations from Col. (3) of Table

3.11. The resulting normalized squared deviation values are shown in Col. (4) of Table 3.11. The next thing is to calculate the observed value of the Chi-squared goodness-of-fit test statistics.

The observed value of the Chi-squared goodness-of-fit test statistics was computed by employing eqn (3.11)' which is a summation of the obtained normalized squared deviations as shown in Col. (4) of Table 3.11. The next thing is to compute the critical value of the Chi-squared goodness-of-fit test statistics that would be used in comparing the observed value of the Chi-squared tests.

The critical value of the Chi-squared goodness-of-fit test statistics was obtained by employing eqn (3.12(a)) i.e. $\nu = K - r - 1 = 4-2-1 = 1$ and eqn (3.12) i.e. $D_1 = \chi^2_{\nu,\alpha}$. The critical value of the Chi-squared so obtained from Appendix III which is $D_1 = \chi^2_{0.95,1} = 3.84$ is compared with the observed value of the Chi-squared goodness-of-fit test statistics. Since $\chi^2 = 1.203 < D_1 = \chi_{0.95,1} = 3.84$, accept the null hypothesis and conclude that the sample data follow the two parameter lognormal distribution. This is a one tail test.

In the next section the K-S test statistics for two parameter lognormal distribution is carried out as follows:

3.3.3.5.3 K-S GOODNESS-OF-FIT TEST STATISTICS FOR TWO PARAMETER LOGNORMAL DISTRIBUTION FOR THE PERIOD (1982-1987)

The process towards obtaining the observed value of the K-S test statistics is to first of all standardize all the flood discharges from Col. (6) of Table 3.5 for the period (1982-1987) using eqn (3.17) of the standardized variable U. With the values of the standardized variable so obtained, go over to Appendix II and get each hypothesized CDF, $F_u(U^{(i)})$ shown in Col.(4) of Table 3.12. The formulations followed here is the same as that of eqn (3.14), the only thing that in this case the CDF for the reduced variate is now replaced with the standardized hypothesized CDF, $F_u(U^{(i)})$ for the standardized variable, $U^{(i)}$. The next thing is to compute the observed cumulative histogram.

The observed cumulative histogram, $F^*(Q^{(i)})$ was determined by employing eqn (3.13) and substituting values accordingly. The results of the values substituted accordingly into eqn (3.13) are shown in Col. (3) of Table 3.12. The next thing is to calculate the modulus of the deviations between $F^*(Q^{(i)})$ and $F_Q(Q^{(i)})$ using eqn (3.15) is shown in Col. (5) of Table 3.12.

The highest discrepancy in Col. (5) of Table 3.12 is

$$D_{2(max)} = | F^*(Q^{(i)}) - F_Q(Q^{(i)}) | = |F^*(2440) - F_Q(2440)| = 0.269$$

and is called the observed value of the K-S test statistics which is $D_{2(max)} = 0.269$. The next thing is to estimate the critical value of the K-S test statistics using $D_{n,\alpha}$ in eqn (3.16). For an α level of 5 percent and a size of 6, the critical value of the K-S test statistics is $D_{6,0.05} = 0.521 > D_{2(max)} = 0.269$.

TABLE 3.12: K-S TEST STATISTICS FOR TWO PARAMETER LOGNORMAL DISTRIBUTION FOR FLOOD DISCHARGES FOR THE PERIOD (1982-1987)

Order (i)	Observed $Q^{(i)}$	$F^*(Q^{(i)})$ from eqn (3.13)	$F_Q(Q^{(i)})$ (8.131, 0.218^2 from eqn (3.14)	D_2 from eqn (3.15)
	CUMECS			
(1)	(2)	(3)	(4)	(5)
1	2420	0.167	$F(2420)=F_u(-1.557)=1-0.9402$ = 0.060	0.107
2	2440	0.333	$F(2440)=F_u(-1.519)=1.09356$ = 0.064	0.269*
3	3280	0.500	$F(3280)=F_u(-0.162)=1-0.5644$ = 0.436	0.064
4	3840	0.667	$F(3840)=F_u(0.563)$ = 0.713	0.046
5	4160	0.833	$F(4160)=F_u(0.928)$= 0.823	0.010
6	4240	1.000	$F(4240)=F_u(1.015)$ = 0.845	0.155

Therefore, the null hypothesis need not be rejected at this significance level. The variable, Q, is said to follow the two parameters lognormal distribution and as such can be treated as if it were normally distributed. This is a one tail test.

3.3.3.6 CONCLUSIONS

From the results of the two goodness-of-fit test statistics carried out, it would be wise to say that the two parameter lognormal distribution fits more into the sample data than the Gumbel extreme value distribution because of the large deviations between their observed value and their critical value. Secondly, all the observed value of the two parameter lognormal distribution are seen to be lower than that of the Gumbel extreme value distribution, implying, that more accuracy was achieved with the lognormal distribution than with the Gumbel distribution.

CHAPTER 4

COMPOUND LOGNORMAL DISTRIBUTION

4.1 . INTRODUCTION

A compound distribution consists of combination of the same type of distributions but with different parameters. By the different parameters of the distributions it implies, the low, probable and high values of the mean and standard deviation and for an adequate combination of each pair of mean value and value of the standard deviation, the low, probable and high prior probabilities have been established. The low values are the one which are 90% certain to be exceeded; the probable values are the best guess, and the high values are the ones which are 90% certain will not be exceeded. A compound distribution is a product of the prior probability with any particular combination of the mean and standard deviation subject to normalizing to make the sum of all the posterior probabilities add up to unity.

Prior probabilities are initial probabilities assigned corresponding to areas under the normal curve. They are assigned the low, probable and high prior probabilities and these correspond to 9 combinations. The next thing are the parameters of thedistributions.

The different parameters of the distribution are the low, probable and high values of the mean and standard deviation. These three discrete values of the mean and standard deviation will give a total of 9 combinations.

In other words, the compound distribution is made up of 9 component distributions having each of it being specified by the low, probable or high values of the mean and standard deviation and weighted according to its low, probable or high prior probabilities. In this the three distinct values of each of the mean and the standard deviation give a total of 9 combinations.

It is on the basis of this that the posterior probabilities are computed, using Bayes' theorem which is treated in this Chapter. By using Bayes' theorem, the posterior probability of any particular combination of the mean and standard deviation is obtained by multiplying the individual probabilities. This is then normalized to make the sum of all the posterior probabilities add up to 1.0. The result of this gives the most probable maximum flood discharges. The steps taken to arrive at the most probable maximum flood discharges are as follows:

(i) Assigning prior probabilities;

(ii) Assuming different parameters for the component distribution;

(iii) Obtaining the likelihoods

(iv) Obtaining the posterior probabilities.

4.2 ASSIGNING PRIOR PROBABILITIES TO EACH OF THE COMPONENT DISTRIBUTIONS

As a first step towards assigning prior probabilities to low, probable and high durations of the series let us define what this low, probable and high prior probabilities actually mean. The low probabilities are the ones which are 90% sure will be exceeded, the probables are the best guess and the high probabilities are the ones which are 90% sure will not be exceeded. These will give 9 combinations. It has been made so to correspond to the low, probable and high values of the mean and standard deviation. These three distinct values of the mean and the standard deviation will give a total of 9 combinations. Whereas according to Russell (1982) "the probability of any particular combination of mean and standard deviation is obtained by multiplying the individual probabilities and normalizing them to make the sum of all the probabilities add up to 1.0".

So in order to be able to assign prior probabilities to each of the component distributions of low, probable, and high values of the mean and standard deviation, these steps were followed:

(i) Converting the percentage points at 90%, 50% and 90% on a normal curve into decimal points and making them the cumulative distribution function, $F_u(u)$, as shown in Col(2) of Table 4.1.

(ii) Using this CDF computed at the various points go to Appendix II and get the value of u – as shown in Col. (3) of Table 4.1 where u is the standardized normal variate

(iii) With this value of u so obtained we go to Appendix VI to read the probability density function, $f_u(u)$, as shown in Col.(4) of Table 4.1.

(iv) Sum all the computed PDF i.e. $\Sigma f_u(u)$.

(v) Lastly, a ratio of the first probability density function, $f_u(u)$, to the summed PDF, $\Sigma f_u(u)$, gives the low probability, A ratio of the second PDF, to the summed PDF, $\Sigma f_u(u)$, gives the probable probability and a ratio of the third PDF to the sum total of the PDF, gives the high probability as described in eqn (4.1) and the results of the computation are shown in Col. (5) of Table 4.1.

$$\text{Prior probability} = \frac{f_u(u)}{\Sigma f_u(u)} \quad \ldots \quad \ldots (4.1)$$

in which $f_u(u)$ = PDF;
$\Sigma f_u(u)$ = summation of PDF.

TABLE 4.1 – PRIOR PROBABILITY FOR THE PERIOD (1955-1980)

	$F_u(u)$% from normal curve	u from App II	$f_u(u)$ from App VI	Prior probabilities from eqn (4.1)
(1)	(2)	(3)	(4)	(5)
Low (L)	$F_u(u)_L$=90% = 0.9	U_L=1.28	$f_u(u_L)$= 0.176	P_L = 0.234
Probable(P)	$F_u(u)_p$=50% = 0.5	U_p= 0.00	$f_u(U_p)$= 0.399	P_P = 0.532
High (H)	$F_u(u)_H$=90% = 0.9	U_H = 1.28	$f_u(U_H)$=0.176	P_H = 0.234

These prior probabilities are for flood discharges for the period (1955-1980) and as such need to be updated using the different parameters of the distributions. So the next thing now is the derivation of the different parameters of the distributions.

4.3.1 DERIVING THE LOW, PROBABLE AND HIGH VALUES OF MEAN AND STANDARD DEVIATION

There were serious problems here of what to assume to be the low, probable or high values of mean and standard deviation

The Mean values, values of standard deviation, natural logarithm of mean, coefficient of variation and natural logarithm of standard deviation value for flood discharges for the period (1955-1980) from Table 3.5 were determined by employing eqns (3.2), (3.3), (3.4), (3.5) and (3.6) respectively and following the same procedure as in Section (3.3.3.3). The results of the estimates are shown in Table 4.2. The parameters of the distributions for this period (1955-1980) has been assigned the probable value because during this period the river was not interrupted by the construction of a dam or any major hydraulic structure obstructing the natural course of river flow. So as shown in Table 4.2 the values from this period were taken as the probable values.,

The period (1982-1987) witnessed the construction of the dam at Lagdo in the Camerouns from 1982. This dam with a live storage capacity of 4700 million cubic metres in order to be able to damp flood flow has been given maximum release of $1600m^3/s$ for floods with return periods of upto 50 years Normand (1984). The construction of a dam at Lagdo has indeed added much to the low river discharges for this period (1982-1987). Due to this the low values for the mean and standard deviation including their natural logarithms were assigned from this period (1982-1987).

The difference, Δ, between the values of mean and standard deviation for the two periods (1955–1980) and (1982–1987) has been added to the probable values to obtain the high values of mean and standard deviation and subsequently, using them were able to obtain the natural logarithms of the mean values and values of the standard deviation. The steps taken in obtaining this high values of mean and standard deviation values are as follows:

To calculate the natural logarithm for the high value of the standard deviation, the first thing to do is to find the ratio of the value of the standard deviation to the value of the mean called the coefficient of variation using eqn (3.5). On substitution of values of mean and standard deviation into this eqn (3.5) the coefficient of variation for the high value of the standard deviation shown in Col. (5) of Table 4.5 was obtained.

Using the coefficient of variations for the high value of the standard deviation from Col.(5) of Table 4.2 on eqn (3.6), the natural logarithm for the high standard deviation value was obtained as shown in col. (6) of Table 4.2.

Lastly, using the high value of the mean of the variable, Q_i, shown in Col. (3) of Table 4.2 on eqn (3.4), the natural logarithm of the high mean value was computed as shown in Col. (7) of Table 4.2.

TABLE 4.2 – THE LOW(L), PROBABLE (P) AND HIGH (H) VALUES OF MEAN AND STANDARD DEVIATION FOR THE VARIABLE, Q

No.	Period	M_Q from eqn(3.2)	σ_Q from eqn(3.3)	V_Q from eqn(3.5)	σ_Y from eqn(3.6)	M_Y from eqn(3.4)
		m^3/s	m^3/s			m^3/s
(1)	(2)	(3)	(4)	(5)	(6)	(7)
1	1955-1980	6249(P)	1730(P)	0.277(P)	0.272(P)	8.740(P)
2	1982-1987	3397(L)	750(L)	0.221(L)	0.218(L)	8.131(L)
3	Δ+P	9101(H)	2710(H)	0.298(H)	0.292(H)	9.116(H)

Δ = (P) - (L) i.e. (1955-1980) - (1982-1987) for the mean and standard deviation values.

The next thing is to form the initial probability matrix table for the 9-combinations of the 9 component distributions which are as follows:

4.3.2 INITIAL PROBABILITY MATRIX

In this study, the mean and the standard deviation of the logarithms of the annual flood peak values are each specified by three discrete values: The low, probable, and high values. These are assigned initial probabilities of 0.234, 0.532, and 0.234, corresponding to the areas under the normal curve at 90%, 50% and 90%. Similar values were also used by Russell (1982).

The compound distribution is made up of 9 component distributions, each specified by a mean and standard deviation and weighted according to its probability. The three discrete values of each of the mean and the standard deviation give a total of 9 combinations. The probability of any particular combination of mean and standard deviation is found by multiplying the individual probabilities and normalizing to make the sum of all the probabilities add up to 1.0. The initial probability matrix is shown in Table 4.3.

The next thing to do is to update the data of the distribution using the initial probabilities and the parameters of the component distributions. The steps taken in doing this started on how to update the data of the distribution.

TABLE 4.3 - INITIAL PROBABILITY MATRIX

m_i / σ_j	L	P	H	Marginal probability
(1)	(2)	(3)	(4)	(5)
L	0.055	0.124	0.055	0.234
P	0.124	0.283	0.124	0.532
H	0.055	0.124	0.055	0.234
Marginal Probability	0.234	0.532	0.234	

4.4. UPDATING THE DATA OF THE DISTRIBUTION

The data for the period (1955-1980) was used to obtain the type of distribution for the first set of data needed for this "estimation of the most probable maximum floods" as shown in chapter four.

After getting the type of distribution, the need arose to update the data in use to recent date.

The data was updated first by Bayesian analysis using 'B' on 'B_1'. This was exactly to compute the effect of this on the basic data, B, used for types of distribution. During this exercise the result of the posterior probabilities derived for the different parameters of the distributions was checked. It was noted that the probable values have the highest posterior probabilities. These posterior probabilities would then be used as prior probabilities on 'A', to check the same effect (where A is the actual data used for types of distribution, too but after the construction of the dam on river Benue at Lagdo in theCamerouns). If any of the posterior probabilities tends to one, as others tend to zero, that one is immediately assumed to be the probable parameters of this distribution and will be used as the most probable maximum flood in determining the peak design flood given its return period for use in design of any hydraulic structure. In this chapter only the Lognormal distribution has been treated because it accepted the sample data more than the Gumbel distribution and hence, better result

is expected from using it.

4.4.1 LOGNORMAL DISTRIBUTION

In this distribution the calculations were carried out in two perspectives, namely:

(i) for single recorded peak flood;

(ii) for several recorded peak flows.

The likelihoods as a component of the distribution for single recorded peak flood will be treated first.

4.4.1.1 LIKELIHOOD FOR SINGLE RECORDED PEAK FLOOD

The likelihood, indeed, depends on the type of data and for a single recorded peak flood, it is the height of the probability density function. Employing eqn (2.11) i.e.

$$L_{ij}(Q_2) = \frac{1}{Q_2\ \sigma_Y\ (2\pi)^{\frac{1}{2}}}\ \exp\left[-\tfrac{1}{2}\ \frac{(\ln Q_2 - M_Y)^2}{\sigma_Y}\right]$$

and substituting the natural logarithm of the mean value from Col.(4) of Table 3.8 and the natural logarithm of the value of the standard deviation from Col.(6) of Table3 .8 with the flood discharge for 1981 from Col.(6) of Table 3.5, the singular likelihood for lognormal

distribution was obtained as shown in Col. (4) of Table 4.4. Having derived the likelihoods the next thing is to compute the posterior probabilities which are conditioned on likelihoods.

4.4.1.2 POSTERIOR PROBABILITIES FOR SINGULAR RECORDED PEAK FLOW

If the peak flow in one year say 1981 was said to be Q_2, the weights could be updated using Bayes' theorem. Bayes' theorem is the posterior probability and is a product of the prior probability to a weighted combination of mean and standard deviation called the likelihood and normalizing to make the sum of the po-terior probabilities add up to 1.0. They were obtained by employing eqn (2.8) and substituting the prior probabilities from Col. (5) of Table 4.1 and the corresponding likelihoods from Col. (4) of Table 4.3 into it and normalizing with the product of eqns (2.9) to make the sum of the posterior probabilities add up to 1.0. The result of this computation for posterior

TABLE 4.4 – SINGULAR LIKELIHOODS AND POSTERIOR PROBABILITIES FOR 1981, ONLY

M_Y from Table 4.2	σ_Y from Table 4.2	P_{ij} from Table 4.1	L_{ij} (Q_2) from eqn (2.11)	$P_{ij}^{1B_1}$ from eqn(28) (N_F=1.290143x 10^{-4}
(m^3/s				
(1)	(2)	(3)	(4)	(5)
L=8.131	L=0.218	P_{11}=0.234	$L_{11}(Q_2)=1.829\times10^{-4}$	$P_{11}^{1B_1}$ =0.332
P=8.740	P=0.272	P_{22}=0.532	$L_{22}(Q_2)=1.550\times10^{-4}$	$P_{22}^{1B_1}$ = 0.639
H=9.116	H=0.292	P_{33}=0.234	$L_{33}(Q_2)=1.605\times10^{-5}$	$P_{33}^{1B_1}$ = 0.029

Superscript B_1 implies additional data to the basic data, B.

probabilities is shown in Col. (5) of Table 4.4 The next thing is to examine the values of the posterior probabilities obtained.

4.4.1.3. DISCUSSION OF RESULTS

It is observed that the posterior probabilities obtained using the probable parameters of the distributions were the highest followed by the ones obtained using the low parameters of the distribution while the ones computed using the high parameters of the distribution were the lowest. These posterior probabilities will now be used as prior probabilities on the several recorded peak flows for the period (1982-1987) called

section 'A'. By the end of it all, it is expected that some posterior probabilities will tend to zero while one will tend to one. The one that tends to one will become the posterior probability for the probable parameters of the distributions in subsequent use of any hydrologic or hydraulic design. So the next thing is to deal with (ii) of section 4 .4.1 starting with singular likelihoods for several recorded peak flows.

4.4.1.4 SINGULAR LIKELIHOODS FOR SEVERAL RECORDED PEAK FLOWS FOR THE PERIOD (1982-1987)

In this the singular likelihoods for the individual pieces of data was computed for the period (1982-1987). This was done by employing eqn (2.11) and substituting the obtained natural logarithm of the mean value from Col.(4) of Table 3.8 and the natural logarithm of the value of standard deviation from Col.(6) of Table 3.8 with the corresponding flood discharge for each year for the period (1982-1987) from Col. (6) of Table 3.5, the singular likelihoods for lognormal distributions for that period was obtained as shown in Cols (4), (5),(6), (7), (8), and (9) of Table 4.5. In the next section the signular likelihoods from several year recorded peak flows will be combined.

TABLE 4.5 - SINGULAR LIKELIHOODS FOR SIX YEAR PEAK FLOWS FOR THE PERIOD (1982-1987)

M_Y from Table 4.2	σ_Y from Table 4.	P_{ij} from Table 4.1	$L_{ij}(Q)$ from eqn (2.11)					
			Q_3=1982	Q_4=1983	Q_5=1984	Q_6=1985	Q_7=1986	Q_8=1987
M^3/s								
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
L=8.131	L=0.218	P_{11}=0.234	$L_{11}(Q_3)=2.86\times10^{-4}$	$L_{11}(Q_4)=5.506\times10^{-4}$	$L_{11}(Q_5)=2.364\times10^{-4}$	$L_{11}(Q_6)=4.072\times10^{-4}$	$L_{11}(Q_7)=2.578\times10^{-4}$	$L_{11}(Q_8)=2.249\times10^{-4}$
P=8.740	P=0.272	P_{22}=0.532	$L_{22}(Q_3)=1.153\times10^{-4}$	$L_{22}(Q_4)=2.702\times10^{-5}$	$L_{22}(Q_5)=1.528\times10^{-6}$	$L_{22}(Q_6)=7.701\times10^{-5}$	$L_{22}(Q_7)=1.253\times10^{-4}$	$L_{22}(Q_8)=1.387\times10^{-6}$
H=9.116	H=0.292	P_{33}=0.234	$L_{33}(Q_3)=9.039\times10^{-6}$	$L_{33}(Q_4)=9.288\times10^{-7}$	$L_{33}(Q_5)=2.167\times10^{-8}$	$L_{33}(Q_6)=4.523\times10^{-6}$	$L_{33}(Q_7)=1.054\times10^{-5}$	$L_{33}(Q_8)=1.923\times10^{-8}$

4.4.1.5 COMBINING SINGULAR LIKELIHOODS FOR SEVERAL RECORDED PEAK FLOWS FOR THE PERIOD (1982-1987)

For several recorded peak flows, the likelihood of each combination is defined here as the product of the likelihoods computed from the individual pieces of data. To obtain this combined likelihoods employ eqn (2.16) and expanded into eqns (2.16(a)), (2.16(b)), and (2.16(c)) and substitute the singular likelihoods for the six year recorded peak floods (K_D=6) i.e. the period (1982-1987) from Cols. (4), (5), (6), (7), (8), and (9) of Table 4.5 into these eqns (2.16(a)), (2.16(b)), and (2.16(c)) representing the three weighted combination of mean and standard deviation each. The result of this computation is the combined likelihoods shown in Col.(4) of Table 4.6. The next step is to use the combined likelihoods to derive the posterior probabilities.

4.4.1.6 POSTERIOR PROBABILITIES FOR SEVERAL RECORDED PEAK FLOWS FOR THE PERIOD (1982-1987)

This was calculated by employing eqn (2.8) and substituting the combined likelihoods from Col. (4) of Table 4.6 for the likelihoods in the eqn and the posterior probabilities in the eqn and dividing by their normalization factor. Their normalization factor was computed by employing eqn (2.9) which is summation of the product

of the combined likelihoods from Col. (4) of Table 4.6 for the likelihoods in the eqn with the posterior probabilities from Col. (5) of Table 4.4 for the prior probabilities in the eqns. The resulting posterior probabilities from this computations is shown in Col. (5) of Table 4.6. This eqn (2.8) popularly known as Bayes' theorem which could be written here as:

$$\text{Posterior probability for section A} = \frac{\text{Posterior probability for section } B_1 \times \text{Combined likelihood for section A}}{\text{Normalization factor}}$$

$$\text{i.e. } P_{i,j}^{1A} = \frac{P_{i,j}^{1B_1} L_{i,j}^{C}(Q)}{N} \quad \ldots \quad (4.2)$$

in which $P_{i,j}^{1A}$ = posterior probability (weight) for section A assigned to combination i, j;

i = index value of mean (from 1-3);

j = index of standard deviation value (from 1-3)

A = superscript representing section A;

$P_{i,j}^{1B_1}$ = posterior probability of combination i, j, from section B_1;

$L_{i,j}^{C}$ = combined likelihood of data from section A for combination i,j, given the data Q;

N_F = normalization factor = $\Sigma_{i=1}^{3} \Sigma_{j=1}^{3} P_{ij}^{1B_1} L_{ij}^{C}(Q)$

$$= P_{11}^{1B1} \; L_{11}^{c}(Q) + P_{22}^{1B1} \; L_{22}^{c}(Q) + P_{33}^{1B_1} \; L_{33}^{c}(Q)$$

... ... (4.2(a)

The next thing is to discuss the result of the posterior probabilities obtained.

TABLE 4.6 – COMBINED LIKELIHOODS AND POSTERIOR PROBABILITIES FOR SEVERAL RECORDED PEAK FLOWS

M_Y from Table 4.3	σ_Y from Table 4.3	P_{ij}^{1B} from Table 4.4	$L^c_{ij}(Q)$ from eqn (2.16)	P_{ij}^{1A} from (4.2) ($N_F = 2.9179484 \times 10^{-22}$)
m^3/s				
(1)	(2)	(3)	(4)	(5)
L=8.131	L=0.218	P_{11}^{1B1}=0.332	$L^c_{11}(Q)$=8.789x 10^{-22}	P_{11}^{1A}=0.99999986
P=8.740	P=0.272	$P_{22}^{1B}1$ = 0.639	$L^c_{22}(Q)$=6.371x 10^{-29}	P_{22}^{1A} =0.00000014
H=9.116	H=0.292	$P_{33}^{1B}1$ = 0.029	$L^c_{33}(Q)$=1.668x 10^{-37}	P_{33}^{1A} =0.00000000

The superscripts 'C' implies combined and 'A' implies section 'A'.

4.4.3.7 DISCUSSION OF RESULTS

It was seen from the results of the computation that the posterior probability obtained from the high parameters of the distributions was zero at 8 significance level or rather 8 places of decimal and as such was totally rejected for any hydrologic/ hydraulic design. The posterior probability obtained from the probable parameters of the distributions was almost zero at the same 8 places of decimal and as such stand rejected; while the one obtained using the low parameters of the distributions was almost one and as such fully accepted for any hydrologic/hydraulic design.

4.5 GENERAL CONCLUSION

The parameters of the data for the period (1982-1987) having been fully accepted becomes the parameter for the most probable maximum floods for the determination of peak design floods for river Benue at Numan.

CHAPTER 5

COMBINING LOGNORMAL AND GUMBEL DISTRIBUTION

5.1 INTRODUCTION;

Combining distributions is a form of compound distribution brought about by weighted combination of individual two-parameter distributions. In this, the two-parameter types of distribution being used are the "Two parameter lognormal distribution" and the "Gumbel extreme value distribution". In this chapter particular steps were taken in combining the two distributions. The steps taken are in the following orders:

(i) assumption of prior weights:

(ii) obtaining likelihoods for:

1) one year record (1981);

2) several year record (1982-1987);

(iii) obtaining posterior weights for:

1) one year record (1981)

2) several year record (1982-1987);

(iv) Combining PDF using

1) assumed initial weights for the period (1982-1987)

2) obtained posterior weights for the period (1955-1987).

5.2 ASSUMPTION OF PRIOR WEIGHTS

For a proper combination of the component distributions certain basic parameters called the initial weights were assumed. Initial weights are prior weights and to achieve this combination two approaches were assumed and in the order to be undermentioned.

As a first approach initial equivalent weights of 0.5 each was assigned to each of the component distributions being combined such that the sum of the assigned weights add up to unity. This is shown in Col.(2) of Table 5.1 and were to be used as prior weights on Bayes' theorem of this combinations to obtain their modified weights, and also on the PDF of the different distributions during their combinations to access the effect according to Russell (1982).

As a second approach based on the analysis of the accuracy ofthe two goodness-of-fit tests statistic carried out on the two sets of data for the two-parameter types of distributions in which it was discovered that the Two parameter lognormal distribution were more acceptable than the Gumbel extreme value distribution. Based on this subjective initial weights of $W_{LN} = 0.67$ and $W_G = 0.33$ were assigned to the Two parameter lognormal distribution and to the Gumbel extreme value distribution respectively as shown in Col. (3) of Table 5.1. Where

W_{LN} and W_G are the assigned weights to the Two parameter lognormal distribution and the Gumbel extreme value distribution respectively, which its summation must be equal to one.

TABLE 5.1 ASSUMED PRIOR WEIGHTS FOR THE PERIOD (1955-1980)

Order (i)	Initial equivalent weights from Russell(1982)	Initial subjective weights from tests
(1)	(2)	(3)
1	$W_{LN_1} = 0.5$	$W_{LN_2} = 0.67$
2	$W_{G_1} = 0.5$	$W_{G_2} = 0.33$

Subscript 1 implies first approach while subscript 2 implies second approach.

Though there is a need to adhere to the steps mentioned in section 5.1 but a little deviation will not matter provided it is to the clearer understanding of the subsequent sections. Based on this it will be better to show the way in which this combined distributions could be updated.

5.3 UPDATING THE COMBINED DISTRIBUTIONS

In updating this combined distribution the calculations were done in two ways:

i) for single recorded peak flood;

ii) for several recorded peak flows

The data was updated first by Bayesian analysis using 'B' on 'B_1'. This was exactly to compute the effect of this on the basic data, B, used for types of distribution. During this exercise the result of the posterior weights derived for the different distributions was accessed. It would be observed that one of the distributions will consistently have higher posterior weights in both approaches. These posterior weights would then be used as prior weights on 'A' to check the same effect (where A is the actual data used for types of distribution). If the two posterior weights for one distribution are very higher in both approaches, those ones are confirmed to be from a more acceptable distribution than the other ones and should be used for any hydrologic design for hydraulic structures.

With the method of the analysis and its expected results over the next thing is to start with lognormal and Gumbel singular likelihoods as components of the distributions for single recorded posterior weights.

5.3.1.1 LIKELIHOODS FOR SINGULAR RECORDED PEAK FLOOD

The likelihood indeed depends on the type of data. For a single recorded peak flood, it is the height of the probability density function. To compute the likelihood for Lognormal distribution employ eqn (2.11) and substitute the natural logarithms ofthe mean value from Col.(4) of Table3 .8 and values of the standard deviation from Col. (6).of the same Table 3 .8 and the corresponding flood discharge for 1981 from Col.(6) of Table 3.5. The resulting lognormal likelihood ($L_{LN}(Q)$) from the computation is shown in Col. (4) of Table 5.2. To obtain the likelihood for Gumbel distribution, we employ eqn (2.12) but the parameters $\hat{\alpha}_{ex}$and $\hat{\beta}$ of the estimates for the distribution are not known. So the procedure becomes to employ eqn (2.13) for the $\hat{\alpha}_{ex}$ and substitute the standard deviation value from Col. (3) of Table 3.8. The result of thecomputation gives $\hat{\alpha}_{ex}$ = 585. Next, eqn (2.14) was employed for maximum values of $\hat{\beta}$ estimate as the mode of the distribution and on substitution of the mean value from Col. (2) of Table 3.8 and the value of the standard deviation from col.(3) of Table 3.8, the result is $\hat{\beta}$ = 3060. These values of $\hat{\alpha}_{ex}$ and $\hat{\beta}$ were then substituted into eqn (2.12) with the corresponding flood discharge for 1981 from Col.(6) of Table 3.5, to get the Gumbel likelihood ($L_G(Q_2)$) shown in Col.(4) of Table 5.2.

Having found the likelihoods for the two different distributions, the next thing becomes to compute posterior weights which are conditioned on likelihoods.

5.3.1.2 POSTERIOR WEIGHTS FOR SINGULAR RECORDED PEAK FLOW

If the peak flow in one year say 1981 was said to be Q_2, the weight could be updated using Bayes' theorem. The product of this Bayes' theorem is called the posterior weight. So to get this posterior weight in the first approach, firstly, for lognormal distribution employ eqn (2.6(b)) and substitute the lognormal likelihood ($L_{LN}(Q_2)$) from Col.(4) of Table 5.2 and the corresponding initial equivalent weight from Col. 2 of Table 5.1. and normalizing their product to make the posterior weights add up to 1.0 but the summation of the normalization factor is not yet known.

So employing eqn (2.7) and substituting the likelihoods $L_{LN}(Q_2)$ and $L_G(Q_2)$ from Col. (4) of Table 5.2 to their corresponding initial equivalent weights from Col.(2) of Table 5.1, the result is the normalization factor shown in Col. (5) of Table 5.2.

The normalization factor was then substituted into eqn (2.6(b)), the resulting posterior weight for lognormal distribution ($W_{LN1}^{1B_1}$) is shown in Col.(5) of Table 5.2. Where $W_{LN1}^{1B_1}$ implies posterior weight for the additional data B_1 to the basic data, B, for Lognormal

distribution.

For the posterior weight in the same first approach but for Gumbel distribution ($W_{G_1}^{1B_1}$) employing eqn (2.6(c)) and substituting the $L_G(Q_2)$ from Col. (4) of Table 5.2 and the corresponding initial equivalent weight from Col. (2) of Table 5.1 and normalizing with the normalization factor obtained for $W_{LN_1}^{1B_1}$ above. The resulting posterior weight from this computation for Gumbel distribution ($W_{G_1}^{1B_1}$) is shown in Col. (5) of Table 5.2. The next thing is to delve into the second approach.

TABLE 5.2 SINGULAR LIKELIHOODS AND POSTERIOR WEIGHTS FOR 1981, ONLY $\left[\hat{\alpha}_{ex} = 585, \hat{\beta} = 3060 \right]$

m_i from Table 3.8	σ_j from Table 3.8	W_{ij} from Russell (1982)	$L_{LN}(Q_2)$ and $L_G(Q_2)$ from eqns (2.11) & (2.12) resp.	$W_{ij}^{1B_1}$ from eqn(2.6(a)) (N_1 = 0.0001736)
m^3/s				
(1)	(2)	(3)	(4)	(5)
M_Y=8.131	σ_Y=0.218	W_{LN_1} = 0.5	$L_{LN}(Q_2)$=1.829x10^{-4}	$W_{LN_1}^{1B_1}$ = 0.570
M_Q=3397	σ_Q=750	W_{G_1} = 0.5	$L_G(Q_2)$=1.381x10^{-4}	$W_{G_1}^{1B_1}$ = 0.430

Superscript B_1 implies additional data to the basic data, B, used to get types of distribution before the river was impounded. Subscripts 1 means for first approach.

In this second approach $W_{LN_2}^{1B_i}$ is obtained as a product of the assigned subjective initial weight from Col.(3) of Table 5.1 to the lognormal likelihood from Col. (4) of Table 5.2 divided by their normalization factor but this normalization factor for this second approach is not yet known. The normalization factor was derived by employing eqn (2.7) which is a summation of the product of the likelihoods for lognormal and Gumbel distributions from Col.(4) of Table 5.2 to their corresponding initial subjective weights from Col. (3) of Table 5.1. The resulting normalization factor shown in Col. (5) of Table 5.3 was substituted into eqn (2.6 (b)) to obtain the $W_{LN_2}^{1B_i}$ shown in Col. (5) of Table 5.3.

For the Gumbel distribution in this second approach, the posterior weight was computed by employing eqn (2.6(c)) and substituting the Gumbel likelihood ($L_G(Q_2)$) shown in Col. (4) of Table 5.2 and the corresponding initial subjective weight (W_{G2}) shown in Col. (3) of Table 5.1 and normalizing with the normalization factor for $W_{LN_2}^{1B_i}$ shown in Col. (5) of Table 5.3. The resulting Gumbel posterior weight ($W_{G_2}^{1B_i}$) is shown in Col. (5) of Table 5.3.

TABLE 5.3 – SINGULAR LIKELIHOODS AND POSTERIOR WEIGHTS FOR 1981, ONLY $\left[\hat{\alpha}_{ex} = 585,\ \hat{\beta} = 3060\right]$

M_i from Table 3.8	σ_j from Table 3.8	W_{ij_2} from tests	$L_{LN}(Q_2)$ and $L_G(Q_2)$ from eqns (2.11) and (2.12) resp.	$W^{1B_i}_{ij_2}$ from eqn(2.6(a)) (N=1.6726 x 10^{-4}
m^3/s				
(1)	(2)	(3)	(4)	(5)
M_Y=8.131	σ_Y=0.218	W_{LN_2}=0.67	$L_{LN}(Q_2)$=1.829x10^{-4}	W^{1B}_{LN}=0.729
M_Q=3397	σ_Q= 750	W_{G2}=0.33	$L_G(Q_2)$=1.381x10^{-4}	W^{1b}_{G}=0.271

Subscript 2 implies for second approach except for Q_2

5.3.1.3 DISCUSSION OF RESULTS

It was noticed that the posterior weights obtained using the lognormal distribution were consistently higher than the ones obtained using the Gumbel distribution in both approaches and these results were according to expectations.

The next thing is to work for the confirmatory tests which are using the posterior weights for the Gumbel and lognormal distributions as prior weights on the several recorded peak flows for the period (1982-1987) called Section 'A'.

By the end of it all, it is expected that the posterior weights from a particular distribution will tend to one while the ones from alternative distribution will tend to zero implying that the former distribution was more acceptable than the latter for any flood probability estimation for river Benue at Numan.

So the next step is to deal with (ii) of section 5.3. starting with singular likelihoods for several recorded peak flows.

5.3.2.1 SINGULAR LIKELIHOODS FOR SEVERAL RECORDED PEAK FLOWS FOR THE PERIOD (1982-1987)

The procedure in this was to compute the singular likelihoods for the individual class of data for the period (1982-1987). This was done by employing eqns (2.11) and (2.12) for lognormal and Gumbel distributions respectively. This was done for K_D=6yrs i.e. for the six year period, for each distribution and as follows:

For lognormal distribution, on substitution of the logarithms of the mean values from Col.(4) of Table 3..8 and values of the standard deviation from Col.(6) of Table 3.8 and the corresponding flood discharge for each year for the period (1982-1987) from Col.(6) of Table 3.5 into eqn (2.11) the lognormal singular

TABLE 5.4 - SINGULAR LIKELIHOODS FOR SIX YEAR PEAK FLOWS FOR THE PERIOD (1982-1987)

$[\hat{Q}_{ex} = 585;\ \hat{\theta} = 3060]$

M_i from Table 4.8	σ_j from Table 4.8	W_{ij} from Russell (1982)	$L_{LN}(Q_k)$ and $L_G(Q_k)$ from eqns (2.13) and (2.14) respectively					
			Q_3=1982	Q_4=1983	Q_5= 1984	Q_6= 1985	Q_7= 1986	Q_8=1987
M^3/s	M^3/s							
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
M_Y=8.131	σ_Y=0.218	W_{LN}=0.5	$L_{LN}(Q_3)$= 2.86×10^{-4}	$L_{LN}(Q_4)$= 5.506×10^{-4}	$L_{LN}(Q_5)$= 2.364×10^{-4}	$L_{LN}(Q_6)$= 4.072×10^{-4}	$L_{LN}(Q_7)$ 2.578×10^{-4}	$L_{LN}(Q_8)$= 2.249×10^{-4}
M_Q=3397	σ_Q=750	W_G=0.5	$L_G(Q_3)$= 2.239×10^{-4}	$L_G(Q_4)$= 5.907×10^{-4}	$L_G(Q_5)$= 2.753×10^{-4}	$L_G(Q_6)$= 3.462×10^{-4}	$L_G(Q_7)$ = 1.991×10^{-4}	$L_G(Q_8)$= $2.577\times10.^{-4}$

likelihoods ($L_{LN}(Q_k)$) shown in Cols. (4), (5), (6), (7) (8) and (9) were computed.

For Gumbel distribution, on substition of the $\hat{\alpha}_{ex}$ value from section 5.3.1 and value of $\hat{\beta}$ from the same section 5.3.1 for their corresponding flood discharge for each year for the period (1982-1987) from Col. (6) of Table 3.5 into eqn (2.12) the Gumbel singular likelihoods ($L_G(Q_k)$) shown in Cols.(4), (5), (6), (7), (8), and (9) were derived.

With the singular likelihoods for the six years period for the two different distributions, the next thing becomes to compute the combined likelihoods as a necessary factor for the computatiom of posterior weights.

5.3.2.2 COMBINING SINGULAR LIKELIHOODS FOR SEVERAL RECORDED PEAK FLOWS FOR THE PERIOD (1982-1987)

The definition that "for several recorded peak flows, the likelihood of each combination is the product of the likelihoods computed from the individual pieces of data "is adopted here. So the combined likelihoods for the two different distributions are found by employing eqn (2.16) and expanding it into eqns (5.1) and (5.2), thus:

$$L^{c}_{LN}(Q) = L_{LN}(Q_3) \times L_{LN}(Q_4) \times L_{LN}(Q_5) \times L_{LN}(Q_6) \times L_{LN}(Q_7) \times L_{LN}(Q_8). \quad \ldots \quad \ldots \quad \ldots \quad (5.1)$$

in which $L^{c}_{LN}(Q)$ = combined likelihood for lognormal distribution;

c implies combined;

$L_{LN}(Q_{k_D})$ = singular likelihood for lognormal distribution;

k_D = number of years (1,2,3,4,5,6)

= 6 yrs

$$L^{c}_{G}(Q) = L_{G}(Q_3) \times L_{G}(Q_4) \times L_{G}(Q_5) \times L_{G}(Q_6) \times L_{G}(Q_7) \times L_{G}(Q_8) \quad \ldots \quad \ldots \quad (5.2)$$

in which $L^{c}_{G}(Q)$ = combined likelihood for Gumbel distribution;

$L_{G}(Q_k)$ = singular likelihood for Gumbel distribution

For lognormal combined likelihood ($L^{c}_{LN}(Q)$), on substitution of the singular lognormal likelihoods ($L_{LN}(Q_k)$) from Cols. (4), (5),(6), (7), (8) and (9) into eqn (5.1) the $L^{c}_{LN}(Q)$ was obtained as shown in Col. (4) of Table 5.5.

For Gumbel combined likelihood, ($L^c_G(Q)$), on substitution ofthe singular Gumbel likelihoods ($L_G(Q_k)$), from Cols.(4),(5),(6),(7),(8), and (9) into eqn (5.2) the $L^c_G(Q)$ was obtained as shown in Col.(4) of Table 5.5. The next thing to do is to use this combined likelihoods to derive posterior weights for the two different distributions forboth approaches.

5.3.2.3 POSTERIOR WEIGHTS FOR SEVERAL RECORDED PEAK FLOWS FOR THE PERIOD (1982-1987)

The proceedure here for computing the posterior weights has to be done in two approaches. In the first approach the posterior weights for the singular recorded peak flows will be used as prior weights while those for the same singular recorded peak flows but for the second approach will be used as prior weights for Section A in the second approach. The order in which these computations were carried out in their respective approaches are as follows:

In the first approach eqn (2.6(b)) was employed for lognormal distribution and rewritten as eqn (5.3) for the purposes of this computation, i.e.

$$W_{LN_1}^{1A} = \frac{W_{LN_1}^{1B_1} \; L^c{}_{LN}(Q)}{N_{f_1}} \quad \ldots \quad \ldots \quad (5.3)$$

in which $W_{LN_1}^{1A}$ = posterior weight to section A due to lognormal distribution for first approach;

$W_{LN}^{1B_1}$ = posterior weight for Section B_1 due to lognormal distribution for first approach;

$L^c{}_{LN}(Q)$ = combined likelihood for lognormal distribution for Section A,

N_{f_1} = normalization factor for first approach:

$$= W_{LN_1}^{1B_1} \; L^c{}_{LN}(Q) + W_G^{1B_1} \; L_G^c(Q) \quad \ldots \quad (5.4)$$

in which $W_G^{1B_1}$ = posterior weight for section B_1 due to Gumbel distribution for first approach;

$L^c{}_G(Q)$ = combined likelihood for Gumbel distribution for section A;

subscript 1 implies first approach.

In this eqn (5.3) every other factor on the right has been known except the normalization factor. So to calculate the normalization factor eqn (2.7) was employed and redefined here as eqn (5.4) for the purposes of this computation. On substitution of the combined likelihood

for lognormal and Gumbel distributions respectively from Col.(4) of Table 5.5 to their corresponding posterior weights for B_1 from Col.(5) of Table 5.2, into eqn (6.4) the normalization factor was computed as shown in Col(5) of Table 5.5. So substituting the posterior weight for B_1 from col.(5) of Table 5.2 and the combined likelihood from Col.(4) of Table 5.5 and normalizing with the normalization factor shown in Col. (5) of Table 5.5 into eqn (5.3), the lognormal posterior weight shown in Col(5) of Table 5.5 was obtained. The next thing is to get the posterior weight due to Gumbel distribution.

The posterior weight due to Gumbel distribution in this first approach was obtained by substituting the $W_{G_1}^{1B_1}$ from Col.(5) of Table 5.2 and the $L^c{}_G(Q)$ from Col.(4) of Table 5.5 and normalizing with the normalization factor from Col.(5) of Table 5.5 into eqn (2.6(c)) rewritten here as eqn (5.5) for the purpoes of this computation. The resulting posterior weight due to Gumbel distribution was found as shown in Col.(5) of Table 5.6.

$$\text{i.e} \quad W_{G_1}^{1A} = \frac{W_{G_1}^{1B_j} \; L^c{}_G(Q)}{N_{f_i}} \qquad \ldots \qquad \ldots \qquad (5.5)$$

in which $W_{G_1}^{1A}$ = posterior weight to section A due to Gumbel distribution for first approach.

The next thing is to delve into the second approach for obtaining posterior weights for this section.

TABLE 5.5 – COMBINED LIKELIHOODS AND POSTERIOR WEIGHTS FOR SEVERAL RECORDED PEAK FLOWS

$\left[\hat{\alpha}_{ex} = 585,\ \hat{\beta} = 3060\right]$

M_i from Table 3.8	σ_j from Table 3.8	$W^{1B_1}_{ij_1}$ from Table 5.2	$L^c_{LN}(Q)$ & $L^o_G(Q)$ from eqns (5.1) and (5.2) resp.	$W^{1A}_{ij_1}$ from eqn(2.6(a)) ($N_{f1}=7.79097 \times 10^{-22}$
m^3/s	m^3/s			
(1)	(2)	(3)	(4)	(5)
$M_Y=8.131$	$\sigma_Y=0.218$	$W^{1B_1}_{LN_1}=0.570$	$L^c_{LN}(Q)=8.789\times10^{-22}$	$W^{1A}_{LN_1}=0.643$
$M_Q=3397$	$\sigma_Q=750$	$W^{1B_1}_{G1}=0.430$	$L^c_G(Q)=6.468\times10^{-22}$	$W^{1A}_{G1}=0.357$

Superscripts 'C' implies combined and 'A' implies section A

In this second approach eqn (2.6(b)) was employed rewritten here as eqn (5.6) for posterior weight due to lognormal distribution ($W^{1A}_{LN_2}$), i.e.

$$W^{1A}_{LN_2} = \frac{W^{1B_1}_{LN_2}\ L^c_{LN}(Q)}{N_2} \quad \ldots \quad \ldots \quad (5.6)$$

in which $W_{LN_2}^{1A}$ = posterior weight due to lognormal distribution for section A for second approach;

$W_{LN_2}^{1B_1}$ = posterior weight for section B_1 due to lognormal distribution for second approach;

N_{f2} = normalization factor for second approach:

$$= W_{LN_2}^{1B_1} \; L^c{}_{LN}(Q) + W_{G_2}^{1B_1} L^c{}_G(Q) \quad \ldots \quad (5.7)$$

in which $W_{G_2}^{1B_1}$ = posterior weight for section B_1 due to Gumbel distribution for second approach;

subscript 2 implies second approach.

In this eqn (5.6) every other factor is known except the normalization factor. So employ eqn (5.7) and substituting the $W_{LN2}^{1B_1}$ and $W_{G_2}^{1B_1}$ from Col.(5) of Table 5.3 to their corresponding combined likelihoods from Col.(4) of Table 5.5 and normalizing with the normalization factor from Col.(5) of Table 5.6 into it, the $W_{LN_2}^{1A}$ was obtained as shown in Col.(5) of Table 5.6.

In this second approach eqn (2.6(c)) was employed rewritten here as eqn (5.8) for posterior weight for section A due to Gumbel distribution ($W_{G_2}^{1A}$), i.e.

$$W_{G_2}^{1A} = \frac{W_{G_2}^{1B_i} \; L_G^c(Q)}{N_2} \qquad \ldots \qquad \ldots \qquad (5.8)$$

in which $W_{G_2}^{1A}$ = posterior weight for section A due to Gumbel distribution for second approach;

On substitution of the $W_{G_2}^{1B_i}$ from Col.(5) of Table 5.3 and the $L_G^c(Q)$ from Col.(4) of Table 5.5 and normalizing with the normalization factor from Col.(5) of Table 5.6, the $W_{G_2}^{1A}$ was found as shown in Col.(5) of Table 5.6.

TABLE 5.6 – COMBINED LIKELIHOODS AND POSTERIOR WEIGHTS FOR SEVERAL RECORDED PEAK FLOWS

$\left[\hat{\alpha}_{ex} = 585, \; \hat{\beta} = 3060\right]$

M_i from Table 3.8	σ_j from Table 3.8	W^{1B_i} from Table 5.3	$L^c_{LN}(Q)$ & $L^c_G(Q)$ from eqns (5.1) & (5.2) resp.	$W^{1A}_{ij_2}$ from eqn (2.6(a)) (N_{f2} = 8.160009 x 10^{-22})
M^3/s	M^3/s			
(1)	(2)	(3)	(4)	(5)
M_Y=8.131	σ_Y=0.218	$W^{1B_i}_{LN_2}$=0.729	$L^c_{LN}(Q)=8.789\times10^{-22}$	$W^{1A}_{LN_2}$ = 0.785
M_Q=3397	σ_Q=750	$W^{1B_i}_{G_2}$=0.271	$L^c_G(Q)=6.468\times10^{-22}$	$W^{1A}_{G_2}$ 0.215

5.3.2.4 DISCUSSION OF RESULTS

It was seen from the results of the computation from both approaches that the posterior weight due to lognormal distribution was consistently and persistently increasing while that due to Gumbel distribution was consistently and persistently decreasing. As such the posterior weights due to Lognormal distribution are said to be tending towards one while posterior weights due to Gumbel distribution were tending towards zero. Thus the sample data is said to follow the Lognormal distribution more than the Gumbel distribution. In otherwords more accuracy was reached with the Lognormal distribution using the sample data than the Gumbel distribution.

5.3.3 SUMMARY

The Lognormal distribution has more influence on the posterior weights than the Gumbel distribution. Hence, the Lognormal distribution has been recommended for all flood probability estimation for river Benue at Numan. The next thing to do is to combine probabilities using Lognormal and Gumbel distributions.

5.4 COMBINING PROBABILITIES USING LOGNORMAL AND GUMBEL DISTRIBUTIONS

In this section the probabilities were combined using lognormal and Gumbel distributions. This was done for the periods (1955-1980) and (1982-1987). Their PDF were obtained and applied to the eqn of the combining probabilities but this time using two different distributions. On application of their PDF on the combined probabilities certain combined probabilities are expected to be very low while other combined probabilities will be a bit high.

The flood discharges for the period having their combined probabilities tending to one will be said not to be accepting the distributions and as such could not be used as probable maximum floods for any hydrologic/hydraulic design, the alternative flood discharges for the period having their combined probabilities tending to zero will be said to be accepting the distributions and as such could be used as the most probable maximum flood for any hydrologic/hydraulic design on river Benue at Numan.

The way in which these combined probabilities using the two different distributions were carried out are in two orders. They include:

(i) combined probabilities for the flood discharges for the period (1955-1980), and;

(ii) combined probabilities for the flood discharges for the period (1982-1987).

5.4.1.1 COMBINED PROBABILITIES FOR FLOOD DISCHARGES FOR THE PERIOD (1955-1980)

To obtain this combined probabilities the following steps were taken:

The probability density function for the lognormal distribution written $P_{LN}(Q)$ ($f_u(u_i)$) was obtained by first standardizing thevariate, Q_i, using eqn (3.19). On substitution of the flood discharges from Col.(1) of Table 5.7 and the obtained natural logarithms of the mean values from Col.(4) of Table 4.2 and the values of the standard deviation from Col.(6) of the same Table 4.2 into this equation (3.19), the standardized variables, U_i were obtained. With the obtained standardized variable, U, then go over to Appendix VI and obtain the probability density functions (PDF) shown in Col.(3) of Table 5.7. So the next thing is to get the PDF for the Gumbel distribution so as to give room for possible combination of their probabilities.

The PDF for the Gumbel distribution was computed by first reducing the given variable, Q_i, using eqn (3.7). The α_G and U_G value for the period (1955-1980) were derived by employing eqns (2.26) and (2.27) respectively and substituting natural logarithm of mean and standard deviation values from Table 4.2 into them. The values of α_G = 4.71 and U_G = 8.62 were obtained and substituted into eqn (3.7) accordingly with their respective assigned flood discharges from Col.(1) of Table 5.7, their reduced variates (W_i) were computed.

With the W_i, then go over to Appendix V and get the PDF for the respective flood discharges for the Gumbel distribution shown in Col. (5) of Table 5.7.

This PDF's was used to plot a graph of P(Q) against Q in cumecs as shown in Fig. 5.1 and it was seen that the graph did follow the Gumbel distribution because it is cone shaped. The next thing is to assign modified weights to the combined probabilities and to do this eqns (2.20) modified here into eqn (5.9) and (5.10), for the purposes of this computation was employed.

The modified weights obtained for the period (1982-1987) are the constant weights assigned to eqns (5.9) and (5.10). In which 0.643 and 0.357 of eqn (5.9) are the modified weights from col. (5) of Table 5.5 while 0.785 and 0.215 of eqn (5.10) from Col. (5)of Table 5.6 are the subjective modified weights for the same period (1982 - 1987). The equations for the

TABLE 5.7 - COMBINED PROBABILITIES FOR FLOOD DISCHARGES FOR THE PERIOD (1955-1980)

Q_i assigned	U from eqn(3.19) ($8.74, 0.272^2$)	$P_{LN_{APP.VI}}(Q)$ from	w from eqn(3.8) ($\mu=4.71$, $\mu_G=8.62$	$P_G(Q)$ from App. V	$P_{comb}(Q \geq Q_i)$ from eqn(5.9)	$P_{comb}(Q \geq Q_i)$ from eqn (5.10)
M^3/s						
(1)	(2)	(3)	(4)	(5)	(6)	(7)
1000	u_1=1.729	$f_u(U_{LN_1})$=0.0894	w_1=2.78	$f_w(w_1)$=0.0583	$P_c(Q_1)_I$=0.0783	$P_c(Q_i)_{II}$=0.0827
9000	u_2=1.342	$f_u(U_{LN_2})$=0.1622	w_2=2.28	$f_w(w_2)$=0.0924	$P_c(Q_2)_I$=0.1373	$P_c(Q_2)_{II}$=0.1472
8000	u_3=0.909	$f_u(U_{LN_3})$=0.2683	w_3=1.73	$f_w(w_3)$=0.1486	$P_c(Q_3)_I$=0.2256	$P_c(Q_3)_{II}$=0.2426
7000	u_4=0.418	$f_u(U_{LN_4})$=0.3656	w_4=1.10	$f_w(w_4)$=0.2386	$P_c(Q_4)_I$=0.3203	$P_c(Q_4)_{II}$=0.3383
6000	u_5=0.149	$f_u(U_{LN_5})$=0.6054	w_5=0.37	$f_w(w_5)$=0.3460	$P_c(Q_5)_I$=0.5128	$P_c(Q_5)_{II}$=0.5496
5000	u_6=-0.819	$f_u(U_{LN_6})$=0.7148	w_6=-0.48	$f_w(w_6)$=0.3208	$P_c(Q_6)_I$=0.5741	$P_c(Q_6)_{II}$=0.6301
4425	u_7=-1.268	$f_u(U_{LN_7})$=0.8214	w_7^1=-1.06	$f_w(w_7^1)$=0.1610	$P_c(Q_7)_I$=0.5856	$P_c(Q_7^1)_{II}$=0.6794
4000	u_7=-1.640	$f_u(U_{LN_7})$=0.8960	w_7=-1.535	$f_w(w_7)$=0.0449	$P_c(Q_7)_I$= 0.5922	$P_c(Q_6)_{II}$=0.7130
3000	u_8=-2.700	$f_u(U_{LN_8})$=0.9877	w_8=-2.890	$f_w(w_8)$=0.000009	$P_c(Q_8)_I$=0.6352	$P_c(Q_7)_{II}$=0.7753
2000	u_9=-4.189	$f_u(U_{LN_9})$=0.9999	w_9=-4.8	-	$P_c(Q_9)_I$=0.6429	$P_c(Q_8)_{II}$=0.7849
1000	u_{10}=-6.736	-	-	-	-	-

Subscript 'C' implies combined.
prime implies for graphical purposes only.

$W^{IA}_{LN_1}$=0.643 $W^{IA}_{LN_2}$=0.785
$W^{IA}_{G_1}$= 0.357 $W^{IA}_{G_2}$ = 0.215

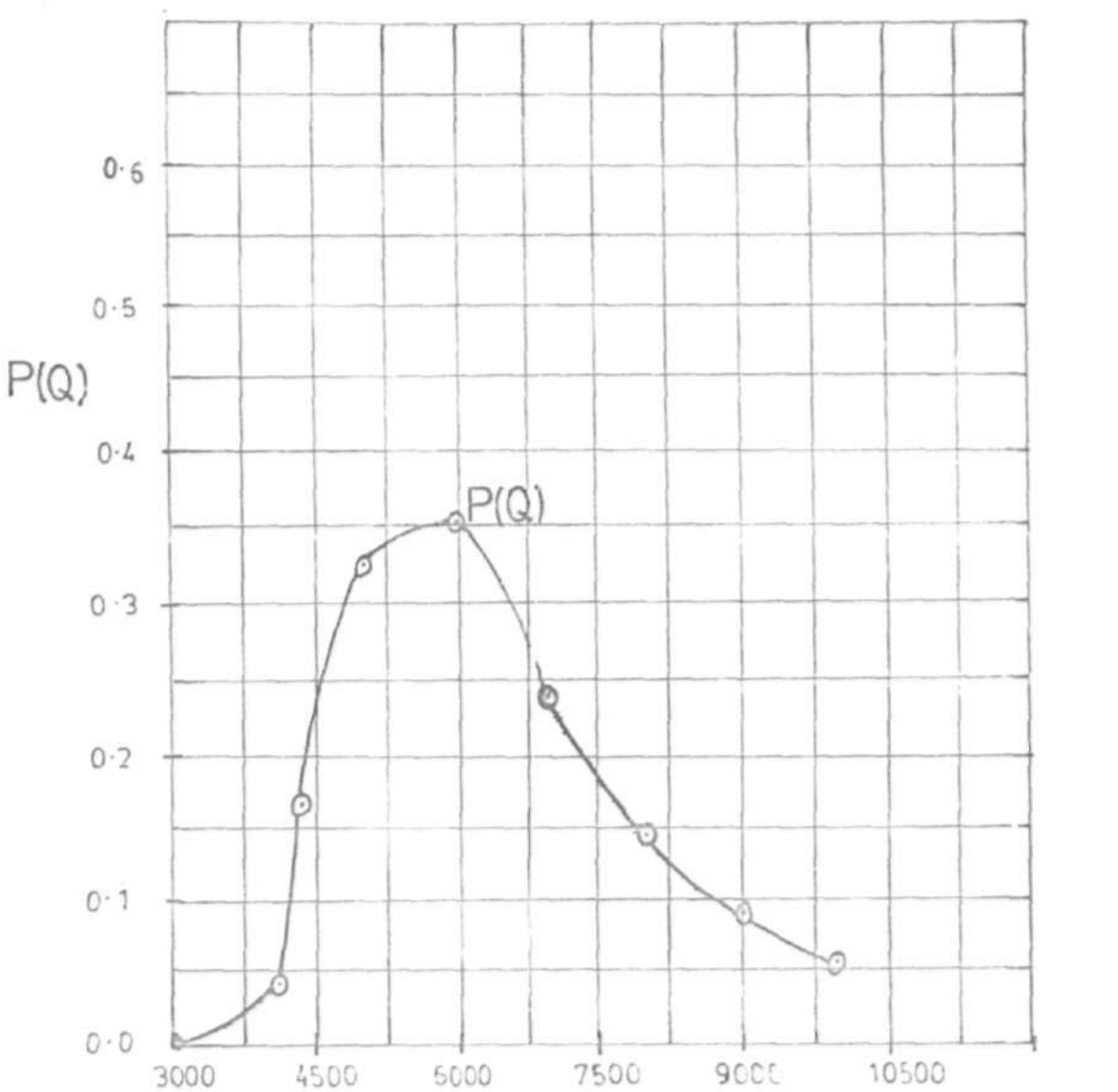


FIG. 5·1— PDF FOR RIVER BENUE AT NUMAN FOR THE PERIOD (1955 –1980)

combined probabilities using two different distributions are as follows:

$$P_{comb}\ (Q \geq Q_i)_I = W^{1A}_{LN_1}\ P_{LN}(Q) + W^{1A}_{G1}\ P_G(Q)$$

$$= 0.643\ P_{LN}(Q_i) + 0.357\ P_G(Q_i) \quad \ldots \quad (5.9)$$

in which $P_{comb}\ (Q \geq Q_i)_I$ = combined probabilities for first approach for the two different distributions;

$W^{1A}_{LN_1}$ = posterior weight due to lognormal distribution for first approach for section A;

$P_{LN}(Q)$ = PDF for lognormal distribution;

$W^{1A}_{G_1}$ = posterior weight due to Gumbel distribution for first approach for section A;

$P_G(Q)$ = PDF for Gumbel distribution

$$P_{comb}(Q \geq Q_i)_{II} = W^{1A}_{LN_2}\ P_{LN}\ (Q) + W^{1A}_{G_2}\ P_G(Q)$$

$$= 0.785\ P_{LN}(Q) + 0.215\ P_G(Q) \quad \ldots \quad (5.10)$$

in which $P_{comb}\ (Q \geq Q_i)_{II}$ = combined probabilities for second approach for the two different distributions;

Q = variable

Q_i = design magnitude or variate

$W^{1A}_{LN_2}$ = posterior weight due to lognormal distribution for second approach for Section A;

$W_{G_2}^{1A}$ = posterior weight due to Gumbel distribution for second approach for Section A.

Using the PDF's shown in Cols (3) and (5) of Table 5.7 on eqns (5.9) and (5.10) the result of the combined probabilities shown in cols (6) and (7) of Table 5.7 were obtained.

5.4.1.2 DISCUSSION OF RESULTS

Since the mean value for this period (1955 -1980) was 6249 cumecs the flood discharges that should attract much attention should be those between 4000 cumecs to 8000 cumecs and above. It was discovered that the assigned weights has more influences on lower flood discharges than higher flood discharges based on the differences between their combined probabilities, i.e. discrepancies between types I and II $P_c(Q \geq Q_i)$.

From Fig 5.1, it was seen that the figure is cone shaped and as such said to follow the Gumbel distribution. Here type I implies combined probabilities for the posterior weights obtained for section A using posterior weights obtained for section B_1 from initial equivalent weights for 1st approach and type II indicates combined probabilities obtained for

section A using posterior weights obtained for section B_1 from initial subjective weights for second approach. The next thing is to examine what is expected at the end of it all.

It is expected that at the end of it all, the two results for the two different periods will be compared to know which one happens to be the most probable maximum floods for practical purposes involving hydrologic / hydraulic designs.

5.4.2.1 COMBINED PROBABILITIES FOR THE FLOOD DISCHARGES FOR THE PERIOD (1982-1987)

To obtain the combined probabilities the following steps were taken:

The probability density function for the lognormal distribution written $P_{LN}(Q)$ $(f_u(u_i))$ was computed by first standardizing the variate, Q_i, using eqn (3.19). On substitution of the natural logarithms of the mean value from Col.(4) of Table 4.2 and value of the standard deviation from Col.(6) of the same Table 4.2 into this eqn (3.19) with their respective assigned flood discharge from Col.(1) of Table 5.8, their standardized values were found as shown in Col.(2) of Table 5.8.

With the obtained standardized variable, $U_{(i)}$, then go over to Appendix VI and get the PDF ($F_u(U_{(i)}$) shown in Col. (3) of Table 5.8.

The PDF for the Gumbel distribution was computed by first reducing the given variable, Q_i, using eqn. (3.8). The α_G and U_G values were determined by employing eqns (2.26) and (2.27) respectively and substituting the natural logarithms of the standard deviation value and the value for the period (1982-1987) from Cols (6) and (4) respectively of Table 3.8. The α_G=5.88S/m^3 and U_G = 8.03m^3/S were derived. With the estimated values of $_G$ and U_G substituted into eqn (3.8) with their respective assigned flood f discharges from Col.(1) of Table 5.8 their reduced variates were obtained.

With the reduced variates, $W_{(i)}$, then go over to Appendix V and get the PDF for the Gumbel distribution shown in Col. (5) of Table 5.8. A graph of the obtained PDF (P(Q)) was plotted against assigned flood discharges ($Q_{(i)}$) in cumecs as shown in Fig. 5.2 and it was seen to follow the Gumbel distribution.

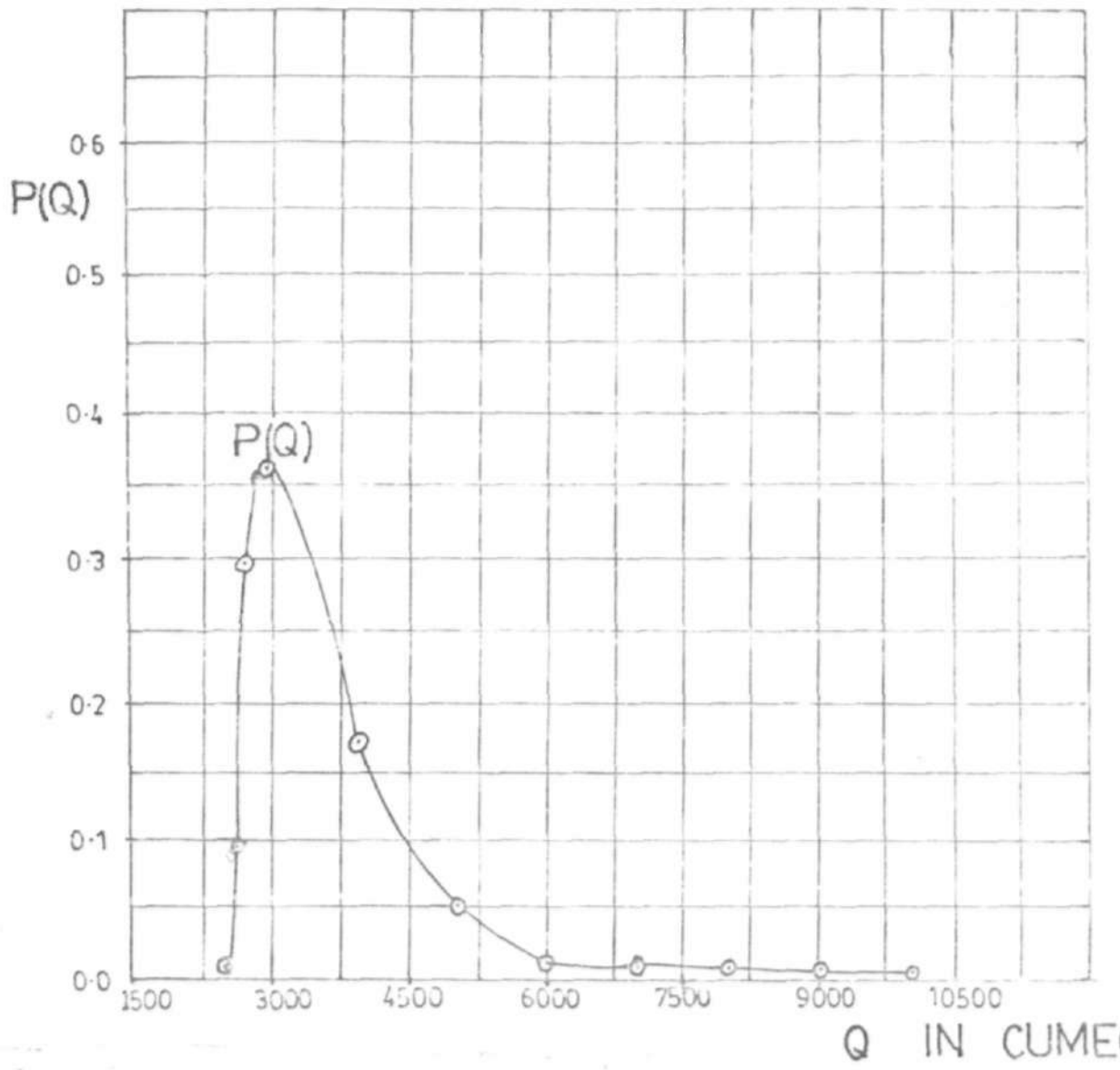
The weights assigned to eqns (5.11) and (5.12) are the initial equivalent weight of 0.5 for each distribution from Col.(2) of Table 5.1 and the initial subjective weights assigned from the result of the analysis of the two goodness-of-fit test statistics carried out for each distribution from Col. (3) of Table 5.1. Note that the summation of the weights assigned to each combination must be equal to one. The equations for the combined probabilities using the two different a distributions for the period(1982-1987)

TABLE 5.8 COMBINED PROBABILITIES FOR FLOOD DISCHARGES FOR THE PERIOD (1982-1987)

Q_i assigned	u from eqn(3.19) (8.131, 0.218^2)	$P_{LN}(Q)$ from App. VI	w from eqn(3.8') (α=5.88, u_G=8.03)	$P_G(Q)$ from App. V	$P_{comb}(Q \geq Q_i)_I$ from eqn(5.11)	$P_{comb}(Q \geq Q_i)_{II}$ from eqn(5.12)
m^3/s						
(1)	(2)	(3)	(4)	(5)	(6)	(7)
1000	u_1=4.94	$f_u(u_1)$=0.000008	w_1=6.94	$f_w(w_1)$=0.00097	$P_c(Q_1)_I$ =0.0005	$P_c(Q_i)_I$=0.0003
9000	u_2=4.47	$f_u(u_2)$=0.000072	w_2=6.32	$f_w(w_2)$=0.0018	$P_c(Q_2)_I$=0.0009	$P_c(Q_2)_{II}$=0.0006
8000	u_3=3.93	$f_u(u_3)$=0.000237	w_3=5.63	$f_w(w_3)$=0.0036	$P_c(Q_3)_I$=0.0019	$P_c(Q_3)_{II}$=0.0013
7000	u_4=3.31	$f_u(u_4)$=0.0022	w_4=4.84	$f_w(w_4)$=0.0079	$P_c(Q_4)_I$=0.0051	$P_c(Q_4)_{II}$=0.0041
6000	u_5=2.61	$f_u(u_5)$=0.0146	w_5=3.94	$f_w(w_5)$=0.0191	$P_c(Q_5)_I$=0.0169	$P_c(Q_5)_{II}$=0.0161
5000	u_6=1.77	$f_u(u_6)$=0.0833	w_6=2.86	$f_w(w_6)$=0.0541	$P_c(Q_6)_I$=0.0687	$P_c(Q_6)_{II}$=0.0737
4000	u_7=0.748	$f_u(u_7)$=0.3016	w_7=1.55	$f_w(w_7)$=0.1718	$P_c(Q_7)_I$=0.2367	$P_c(Q_7)_{II}$=0.2588
3000	u_8=0.57	$f_u(u_8)$=0.6609	w_8=-0.14	$f_w(w_8)$=0.3640	$P_c(Q_8)_I$=0.5125	$P_c(Q_8)_{II}$=0.5629
2750	u_8'=-0.97	$f_u(u_8')$=0.7508	w_8')=0.65	$f_w(w_8')$=0.2821	$P_c(Q_8')_I$=0.5164	$P_c(Q_8')_{II}$=0.5811
2625	u_9'=-1.18	$f_u(u_9')$=0.8011	w_9'=-0.924	$f_w(w_9')$=0.2028	$P_c(Q_9')_I$=0.5020	$P_c(Q_9')_{II}$=0.5876
2000	u_9=-2.43	$f_u(u_9)$=0.9774	w_9=-2.52	$f_w(w_9)$=0.00006	$P_c(Q_9)$=0.4887	$P_c(Q_9)$ =0.6549
1000	u_{10}=-5.61	-	w_{10}=-6.60	-	-	-

The subscript C implies combine
prime implies for graphical purposes only.

W_{LN_1} = 0.5 W_{LN_2} = 0.67
W_{G_1} = 0.5 W_{G_2} = 0.33



- FIG.5·2-PDF FOR RIVER BENUE AT NUMAN FOR THE PERIOD (1982 —1987)

are as follows:

$$P_{comb}\ (Q \geq Q_i)_I = W_{LN_1}\ P_{LN}(Q) + W_{G_1}\ P_G(Q)$$

$$= 0.5P_{LN}(Q) + 0.5\ P_G(Q) \quad \ldots \quad (5.11)$$

in which W_{LN_1} and W_{G_1} are the initial equivalent weights for lognormal and Gumbel distributions respectively.

$$P_{comb}\ (Q \geq Q_i)_{II} = W_{LN_2}\ P_{LN}(Q) + W_{G2}P_G(Q)$$

$$= 0.67\ P_{LN}(Q) + 0.33\ P_G(Q) \quad \ldots \quad (5.12)$$

in which W_{LN_2} and W_{G_2} are the initial subjective weights for lognormal and Gumbel distributions respectively.

Using the PDF's shown in Cols (3) and (5) of Table 5.8 on eqns (5.11) and (5.12) which were modified from eqn(2.20), the result of the combined probabilities shown in Cols (6) and (7) of Table 5.8 were obtained.

5.4.2.2. DISCUSSION OF RESULTS

Since the mean value is 3397 cumecs, the flood discharges that should form bulk of the concentrations are those from 2000 cumecs to 5000 cumecs. It was discovered that the differential weights have negligible effect on the combined probabilities. So no matter how far the weights are adjusted, i.e., from initial equivalent weight to initial subjective weight the effect will be highly negligible. Though with low flows the

pronounced but with high flood discharges by comparing the discrepancies between the results of the type I and II combined probabilities, the effect becomes minimal, even negligible for all practical purposes.

Furthermore, comparing the results of the combined probabilities for the two periods (1955-1980) and (1982-1987) from their differential Tables of 5.7 and 5.8 respectively, it was discovered that the former have higher combined probabilities than the latter and as such are said to be less correct and even rejected. From 5.2 it is seen to be cone shaped and as such said to follow the Gumbel distribution.

CHAPTER 6

DETERMINATION OF PEAK DESIGN FLOOD

6.1 INTRODUCTION

Peak design floods are the maximum floods which any hydraulic structure is designed to sustain within a particular return period, T-yrs above which a structure might fail. Determination of these peak design floods for various return periods is important to avoid the collapse of any designed hydraulic structures and invariably avoiding large monetary damage, loss of life and disruption of community services. In hydrology, certain maximum floods must have existed from historical records. Based on these and other hydrological factors return periods, T-years, are chosen for most probable maximum floods of some particular periods. Conventionally, high return periods, T-years are chosen like say where the consequences of failure is disastrous like in the construction of a dam costing a lot of money its collapse will not only cause loss in terms of monetary value but also loss of life and disruption of community services by inundation of agricultural lands, homes and serious soil leaching. In this, return periods of about 200 years are often chosen to avoid such failure. But in situations where this is less, like in the construction of culverts, bridges, dykes etc much smaller return periods, T-yrs, in the range of 2-200years are chosen.

This return period, T-years, is important in the determination of peak design floods as aforementioned and its application, is used in the determination of the CDF for each distribution.

Based on this determined CDF and the basic parameters of the distributions i.e. the mean values and values of the standard deviation, which must have been estimated from the adopted probable maximum floods during its tests, the peak design floods for river Benue at Numan will be determined, for each of the distributions.

The anlysis on the steps taken for the determination of the peak design floods for river Benue at Numan for the period (1982-1987) for the two parameter types of distribution are as follows:

6.2 LOGNORMAL DISTRIBUTION

6.2.1 DETERMINING THE PEAK DESIGN RETURN FLOODS FOR GIVEN RETURN PERIODS

The procedure in determining the peak design return flood is to evolve an equation that will involve the return period only and have CDF as the subject of the formula. To do this, the values of the right hand sides

of eqns (2.18) and (2.29) are equated to each other and rearranged to have the CDF as the subject of the formula as follows:

$$CDF = 1 - \frac{1}{T} \quad \ldots \quad \ldots \quad (6.1)$$

in which T = return period in years.

For lognormal distribution, on substitution of the given return periods, T-years, shown in Col.(3) of Table 6.1 into eqn (6.1), the CDF ($F_u(u_i)$) was obtained as shown in Col.(4) of Table 6.1. Using these $F_u(u_i)$ go to Appendix II and get the standardized variables, U_i, shown in Col.(5) of Table 6.1.
With these obtained standardized variable for lognormal distribution, the next thing becomes to determine the peak design return floods for river Benue at Numan for the periods (1982-1987).

To do this, eqn (3.17) was employed and rearranged to give eqn (6.2) which is the equation for the determination of the peak design return floods, i.e.

$$Q_T = e^{(u\sigma_Y + M_Y)} \quad \ldots \quad \ldots \quad (6.2)$$

in which Q_T = peak design return floods for lognormal distribution for given return period, T-years.

On substitution of the natural logarithms of the mean values from Col.(4) of Table 3.8 and the values of standard deviation from Col.(6) of Table 3.8 to each standardized variable from Col.(5) of Table 6.1 into eqn (6.2) the peak design return floods were determined as shown in Col.(6) of Table 6.1.

6.2.2 DISCUSSION OF RESULTS

It is seen that the determined design floods did not vary so much, i.e. the difference between a peak design flood for a particular return period did not differ with up to 1000 m^3/sec with that between the next determined peak design flood for a different return period.

It is expected that by the end of it all, the determined peak design return flood using the lognormal distribution will be compared with those determined using Gumbel distribution.

So the next thing is to determine the peak design return floods using the Gumbel's distribution

TABLE 6.1 PEAK DESIGN RETURN FLOODS FOR GIVEN RETURN PERIODS OBTAINED USING LOGNORMAL DISTRIBUTIONS

M_Y from Table 3.8	σ_Y from Table 3.8	T given	$F_u(u)$ from eqn (6.1)	U from App II	Q from eqn(6.2)
M^3/s		Years			M^3/s
(1)	(2)	(3)	(4)	(5)	(6)
		200	0.995	2.575	$Q_{200}=5957$
8.131	0.218	500	0.998	2.880	$Q_{500}=6367$
		1000	0.999	3.080	$Q_{1000}=6650$

6.3 GUMBEL DISTRIBUTION

6.3.1 DETERMINING THE PEAK DESIGN RETURN FLOODS FOR GIVEN RETURN PERIODS

The procedure towards obtaining the peak design return floods for Gumbel distribution is to substitute the given return periods, T-years, shown in Col. (3) of Table 6.1 and get the CDFs ($F_W(w_i)$) as shown in Col. (4) of Table 6.2. Using thisobtained $F_W(w_i)$ go to Appendix V and obtain the

reduced variates, w_i's , shown in Col.(5) of Table 6.2 for each given return period. With these reduce variates, w_i's, for Gumbel distribution, the next thing becomes to evolve an equation fordetermining the peak design return floods for river Benue at Numan for the periods (1982-1987).

Since eqn (2.22) from Gumbel (1958) did not define what the parameter Q really meant, i.e., whether it is in its logarithm form or otherwise, a more comprehensive equation need to be evolved that will make the determination of the peak design return floods a reality for Gumbel distribution.

TABLE 6.2 - PEAK DESIGN RETURN FLOODS FOR GIVEN RETURN PERIODS OBTAINED USING GUMBEL DISTRIBUTION

M_Y from Table 3.8	σ_Y from table 3.8	T given	$F_W(w)$ from Eqn (6.1)	w from App. V	α_G from Table 5.8	u_G from table 5.8	Q from eqn(6.3)
m^3/s		Years			S/m^3	m^3/s	m^3/s
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
		200	0.995	5.30			$Q_{200} = 7566$
8.131	0.218	500	0.998	6.22	5.88	8.03	$Q_{500} = 8847$
		1000	0.999	6.91			$Q_{1000} = 9948$

Based on this eqn (3.8) was employed and rearranged to give eqn (6.3) used in determining the peak design return floods for Gumbel distribution.

$$Q_T = e^{(\frac{w}{\alpha_G} + u_G)} \quad \ldots \quad (6.3)$$

in which Q_T = peak design return floods for Gumbel distribution given the return periods, T-years,

All other parameters remain as previously defined in section 3.3.4.4.2.

On substitution of the α_G value from Table 3.9 and value of u_G from the same Table 3.9 to each reduced variate, w, from Col.(5) of Table 6.2, into eqn (6.3) the peak design return floods were determined as shown in Col.(8) of Table 6.2.

6.3.2 DISCUSSION OF RESULTS

It was noticed that the difference between a particular determined peak design return flood for a particular given return period and the next determined peak design return flood for its different given return period is greater than $1000\ m^3/s$ and as such could not be used for any hydrologic/hydraulic design for river Benue at Numan, quite unlike those determined

using lognormal distribution. Hence, the peak design return floods determined using the lognormal distribution are said to be more accurate than those determined using the Gumbel distribution.

6.4 SUMMARY

The peak design return floods determined using the lognormal distribution are recommended for design of hydraulic structures on river Benue at Numan.

CHAPTER 7

ESTIMATION OF RISK OF FAILURE

7.1 INTRODUCTION:

The simple risk of failure has been defined by Yen (1970) as the probability of occurrence of the variable, Q, greater than the design magnitude, Q_1, which is also equal to the chance of nonoccurrence of $Q \leq Q_1$. This risk of failure is very important in all hydrological/hydraulic designs in that it is a particular factor which any attempt to keep below its specification will lead to immediate or eventual failure of the hydraulic structure in question. Certain factors are considered in the estimation of this risk of failure.

The factors include;

(i) Serviceability;

(ii) expected project life, n-years;

(iii) design return periods, T-years.

The steps taken in the analysis of this risk of failure for any hydraulic project on river Benue at Numan are as follows:

7.2 LOGNORMAL DISTRIBUTION

7.2.1 ESTIMATING SIMPLE RISK OF FAILURE

The procedure towards estimating the risk of failure for structures using lognormal distribution is to estimate first and foremost the likelihood of flow, Q, not being exceeded by another flow of magnitude Q_i in n years. This is obtained by employing eqn.(2.18) and written here as:

$$P_L = \left[1 - p\ (Q \leqslant Q_i) \right]^{n_F} = \left[F_u(u) \right]^{n_F} \quad \ldots \quad (7.1)$$

in which P_L = likelihood of flow Q not being exceeded by another flow of magnitude Q_i in n-years;

n_F = expected project life in years,

$F_u(u)$ = CDF for standardized variable, U,

Q = Variable;

Q_i = Variate.

On substitution of the $F_u(u)$ from Col.(4) of Table 6.1 for a given return period and the given expected project life, n-years, from Col.(3) of Table 7.1 into eqn (7.1), the likelihood of flow Q not being exceeded by another flow of magnitude Q_i in n-years will be estimated and the results of the estimates is shown

in Col. (4) of Table 7.1. The next thing is to estimate the simple risk of failure for hydraulic structures on river Benue at Numan.

The simple risk of failure which by definition "is the probability of occurrence of the variable, Q, greater than the design magnitude, Q_1, which is also equal to the chance of nonoccurrence of $Q \leq Q_1$" i.e.

$$\text{Risk (RY)}_{n_F} = P(Q > Q_1)_{n_F} = 1 - \left(1 - \frac{1}{T}\right)^{n}{}_{F} \ldots \quad (7.2)$$

In which $P(Q > Q_1)_{n_F}$ is the simple risk of failure involved in hydrologic designs for a given expected project life, n-years.

On substitution of the likelihood of exceedence from Col.(4) of Table 7.1 into eqn (7.2), the simple risk of failure involved in hydrologic designs for hydraulic structures was estimated as shown in Col.(5) of Table 7.1.

TABLE 7.1 - RISKS OF FAILURE GIVEN THE EXPECTED PROJECT LIFE

T from Table 6.1	$F_u(u)$ from Table 6.1	n_F = given	P_L from eqn (7.1)	RY_{nF} from eqn (7.2)
Years		Years		
(1)	(2)	(3)	(4)	(5)
200	0.995	50	0.7783	RY_{50}=0.2217
		100	0.6058	RY_{100}=0.3942
		200	0.3670	RY_{200}=0.6330
500	0.998	50	0.9047	RY_{50}=0.0953
		100	0.8186	RY_{100}=0.1814
		200	0.6701	RY_{200}=0.3299
1000	0.999	50	0.3512	RY_{50}=0.0488
		100	0.9048	RY_{100}=0.0952
		200	0.8186	RY_{200}=0.1814

7.2.2 DISCUSSIONS OF RESULTS

It was discovered that the higher the return period, the lower the risk of failure but the higher the expected project the higher the risk of failure. So for high expected

project life, with high return period, the serviceability should be frequent to avoid silts from accumulating at the bottom of the structure and in case of any cracks on the materials of the structure it should be straightened out immediately. For low return period and high expected project life the serviceability should be more frequent so that the peak design floods will not be exceeded due to poor maintenance of silting.

It is expected that by the end of it all, the risk levels estimated using the two different two-parameter types of distributions will be compared. The next thing is to estimate the risk levels using the Gumbel distribution.

7.3 GUMBEL DISTRIBUTION

7.3.1 ESTIMATING SIMPLE RISK OF FAILURE

The procedure towards estimating the simple risk of failure involved in hydrologic designs for given expected project life, n-years for hydraulic structures on river Benue at Numan for Gumbel distribution was to employ eqn (2.19) and is equal to eqn (2.32), written here as

TABLE 7.2 RISKS OF FAILURE GIVEN THE EXPECTED PROJECT LIFE

T from Table 6.1	$F_W(w)$ from Table 6.2	n_F from Table 7.1	P_L from eqn (7.3)	RY_{n_F} from eqn (7.2)
Years		Years		
(1)	(2)	(3)	(4)	(5)
200	0.995	50	0.7783	RY_{50}= 0.2217
		100	0.6058	RY_{100}=0.3942
		200	0.3670	RY_{200}=0.6330
500	0.998	50	0.9047	RY_{50} = 0.0953
		100	0.8186	RY_{100}=0.1814
		200	0.6701	RY_{200} = 0.3299
1000	0.999	50	0.9512	RY_{50}=0.0488
		100	0.9048	RY_{100}=0.0952
		200	0.8186	RY_{200}=0.1814

7.3.2 DISCUSSION OF RESULTS

It was noticed that the higher the return period, the lower the risk level but the lower the expected project life the higher the risk. Furthermore, on comparison

with the results of the risk levels estimated using the lognormal distribution, it was seen to be just the same. Hence, no matter the type of distribution used for your design the risk level is the same. The next thing to do is to draw a general conclusion.

7.4 SUMMARY

It is observed that the simple risk of failure involved in all hydrologic design estimated using either of the two distributions were the same.

CHAPTER 8

CONCLUSIONS AND RECOMMENDATIONS

8.1 CONCLUSIONS

(i) Relationship between actual measured stages and discharges for rive Benue at Numan is shown in the rating curve number II of Fig. 3.3.

(ii) As a result of critical analysis of data the maximum discharges for the period (1955-1987) were estimated and are presented in Table 3.5.

(iii) It was proved by applying probability plotting tests and the two goodness-of-fit test statistics for the two periods (1955-1980) and (1982-1987) respectively, that the Two-parameter lognormal distribution fits better into the sample data than the Gumbel extreme value distribution.

(iv) It was found that Bayes's theorem may be applied to transformation of prior distributions (assumed for the initial period) to posterior distributions considering the data for the second period (1982-1987), with modified conditions of flow in the case of :-

(a) The same type of initial distributions (Lognormal) but with different parameters.

(b) The different types of Initial distributions (Lognormal and Gumbel)

(v) From the analysis of the compound distribution it was found that the parameters of posterior probabilities are streaming to parameters of second period (1982-1987).

(vi) From the results of the posterior weights obtained using Bayes's theorem for combined distributions the Lognormal distribution has more influence on the posterior weights than the Gumbel distribution.

(vii) The peak design floods were determined for the three return periods (200, 500 and 1000-yrs) using the parameters of the posterior distributions.

(viii) Basing on the derived distributions the simple risk of failure was computed for the three expected project life (50,100, and 200-yrs) and return period (200,500 and 1000-yrs).

(ix) From the results of the computation, it was found that the simple risk of failure increases as the expected project life increases while it decreases with an increase in the return period.

9.2 RECOMMENDATIONS

(1) It is suggested that the authorities concerned with data collection especially all hydrological divisions keep accurate record of their data.

(2) Rating curve number II of Fig. 3.2 can be used for the estimation of flood discharges for river Benue at Numan.

(3) For design purposes, any hydraulic project on river Benue at Numan for the same design return period and the same expected project life as given in Table 8.1 the risk levels specified in this work will not be exceeded.

(4) The Two-parameter lognormal distribution is recommended for flood probability estimation for river Benue at Numan.

REFERENCES

Achinivu, O.K. (1988). Chief Hydrographic Surveyor "Hydrological records". Inland Waterways Division, Lokoja branch.

Arnold, Z. (1971). An Introduction to Bayesian Inference in Econometrics. John Wiley and Sons Inc. pp. 13-54.

Benjamina, J. R. and Cornell, A. O. (1970). Probability Statistics and Decisions for Civil Engineers. McGraw-Hill Book Coy. pp 134-667.

Bodo, B. and Unny, T.E. (1976). "Model Uncertainty in Flood Frequency Analysis and Frequency-Based Design" Water Resources Research. Vol. 12, No 6 pp 1109-1117.

Box, G.E.P. and Tiao, G.C. (1973). Bayesian Inference in Statistical Analysis. Addison-Wesley Publishing Coy. pp 25-60.

Chow, V.T. (1964). Editor-in-Chief. Handbook of Applied Hydrology. McGraw-Hill Book Coy.

Ejademe, H.O. (1988). Zonal Officer. "Hydrological data". Federal Department of Water Resources, Mando Road, Kaduna, pp 5.79-5.104.

Gumbel, E.J. (1958). "Statistical Theory of Floods and Droughts". Water Resources Research. pp 157-183.

Haan, C.T. (1982). Statistical Methods in Hydrology. Iowa State University Press/Ames. pp 97-348.

Kalbfleisch, J.G. (1979(a)). Probability and Statistical Inference I. Springer-Verlag. pp 65-105.

Kalbfleisch, J.G. (1979 (b)). Probability and Statistical Inference II. Springer-Verlag. pp 99-121.

Kite, G.W. (1978). Frequency and Risk Analysis in Hydrology. Water Resources Publications pp 54-102.

Macdonald, P. (1966). Mathematics and Statistics for Scientists and Engineers. D. Van Nostrand Co. Ltd pp. 197 - 201.

Mood, A.M., Graybial, F.A and Boes, D.C. (1974). Introduction to the Theory of Statistics McGraw-Hill Book Coy. Third Edition. pp 551-552.

Nemec, J. (1972). Engineering Hydrology. McGraw-Hill Book Coy. pp 131-132.

Normand, D. (1984). Modele de "Simulation Prospective" Applique Ala Gestion d'un Barrage a but Multiple. Proceedings of the International Seminar on Water Resources Management Practice, Ilorin, Nigeria. pp. 285-291.

Russell, S.O. (1982). "Flood Probability Estimation". American Society of Civil Engineers, Hydraulics Division, Vol. 108, No. NY1, paper No. 16825 pp 63-73.

Tang, W.H. (1980). "Bayesian Frequency Analysis". American Society of Civil Engineers, Hydraulics Division. Vol. 106, No. HY 7, paper No. 15532. pp 1203-1218.

Tung, Y and Larry, W.M. (1980) "Risk Analysis for Hydraulic Design". American Society of Civil Engineers, Hydraulics Division. Vol. 106, No. HY5, paper No. 15419. pp 893-913.

Wilson, E.M. (1981). Engineering Hydrology. Macmillan Press Ltd. Second Edition. pp 91-117.

Yen, B.C. (1970). Risks in Hydrologic Design of Engineering Projects". American Society of Civil Engineers, Hydraulic Division. Vol. 96, No. HY4 paper No. 1229. pp 959-966.

Hydro-Meteorological Year Book of Upper Benue Basin (1980-1984) Upper Benue River Basin and Rural Development Authority, Yola.

Isu Associates (Nig.) Ltd (Consulting Engineers & Planners) (1985) Benue River Hydrometeorologic Studies Region II & III Report Submitted to the Federal Department of Water Resources, Kaduna, Federal Ministry of Agriculture, Water Resources and Rural Development, Lagos, pp. 5.79 - 5.104.

RIVER BENUE AT YOLA
DISCHARGE AND RUNOFF DATA
1963 – 1987

YEAR	MONTHLY MEAN DISCHARGES m^3/sec MAY	JUN	JULY	AUG	SEP	OCT	NOV	DEC	ANNUAL RUNOFF $10^6 m^3$	REMARKS
1963	107	149	778	2671	4197	2273	594	181	29,668	
1964	92	188	716	1097	2692	1743	515	150	19,632	
1965	47	218	717	2558	2995	1184	266	107	21,805	
1966	119	513	711	1550	3780	—	—	—	—	
1967	48	154	932	1540	2543	1620	286	115	—	
1968	30	243	939	2090	3590	1325	252	109	23,044	
1969	55	237	1340	3516	4103	2512	824	227	35,281	
1970	48	123	500	2367	5267	1579	574	242	28,726	
1971	18	90	962	1848	2985	1307	689	457	22,746	
1972	—	400	1069	1456	1781	1048	161	165	—	
1973	63	160	400	1430	2389	231	240	208	18,801	
1974*	161	179	622	1432	2015	1876	871	652	—	
1975	—	—	870	2446	3340	2197	—	310	—	
1976	965	256	867	1853	1547	1893	809	295	—	
1977	60	190	834	1603		725	368	201	—	
1978	196	380	841	1529	3550	1503	539	308	24,257	UBR BRDA IN OPERATION
1979	187	338	697	1721	2298	812	375	191	18,562	
1980	90	223	824	2336	2972	1203	497	274	23,022	
1981	34	95	780	1616	2900	1247	310	127	19,032	
1982	43	132	407	1232	2123	888	277	120	14,158	IMPOUNDMENT AT LAGDO DAM STARTS
1983*	50	153	302	538	1720	409	138	102	10,011	BENUE RIVER IMPOUNDED AT LAGDO
1984	111	116	282	503	1055	897	209	115	9,742	LAGDO DAM (CAMEROUN) IN OPERATION
1985	83	147	629	1271	1675	547	162	104	13,018	
1986	90	136	317	1113	1590	822	173	99	12,280	
1987*	80	125	201	472	974	760	132	79	8,152	
LIVE STORAGE CAPACITY OF LAGDO DAM									4,700	

*DROUGHT YEARS

APPENDIX I— CONSTRUCTION OF A DAM AT LAGDO IN THE CAMEROUNS

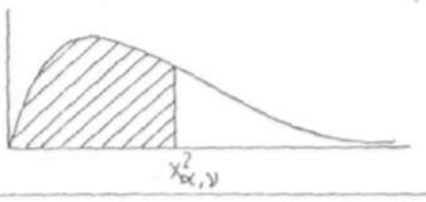
$$\Phi(x)=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}e^{-t^2/2}\,dt$$

x	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998

x	1.282	1.645	1.960	2.326	2.576	3.090	3.291	3.891	4.417
$\Phi(x)$	.90	.95	.975	.99	.995	.999	.9995	.99995	.999995
$2[1-\Phi(x)]$	.20	.10	.05	.02	.01	.002	.001	.0001	.00001

COURTESY OF MOOD GRAYBIAL & BOES (1974).

APPENDIX II— CUMULATIVE NORMAL DISTRIBUTION



F	$\chi^2_{.995}$	$\chi^2_{.99}$	$\chi^2_{.975}$	$\chi^2_{.95}$	$\chi^2_{.90}$	$\chi^2_{.75}$	$\chi^2_{.50}$	$\chi^2_{.25}$	$\chi^2_{.10}$	$\chi^2_{.05}$	$\chi^2_{.025}$	$\chi^2_{.01}$	$\chi^2_{.005}$
1	7·88	6·63	5·02	3·84	2·71	1·32	·455	·102	·0158	·0039	·0010	·0002	·0000
2	10·6	9·21	7·38	5·99	4·61	2·77	1·39	·575	·211	·103	·0506	·0201	·0100
3	12·8	11·3	9·35	7·81	6·25	4·11	2·37	1·21	·584	·352	·216	·115	·072
4	14·9	13·3	11·1	9·49	7·78	5·39	3·36	1·92	1·06	·711	·484	·297	·207
5	16·7	15·1	12·8	11·1	9·24	6·63	4·35	2·67	1·61	1·15	·831	·554	·412
6	18·5	16·8	14·4	12·6	10·6	7·84	5·35	3·45	2·20	1·64	1·24	·872	·676
7	20·3	18·5	16·0	14·1	12·0	9·04	6·35	4·25	2·83	2·17	1·69	1·24	·989
8	22·0	20·1	17·5	15·5	13·4	10·2	7·34	5·07	3·49	2·73	2·18	1·65	1·34
9	23·6	21·7	19·0	16·9	14·7	11·4	8·34	5·90	4·17	3·33	2·70	2·09	1·73
10	25·2	23·2	20·5	18·3	16·0	12·5	9·34	6·74	4·87	3·94	3·25	2·56	2·16
11	26·8	24·7	21·9	19·7	17·3	13·7	10·3	7·58	5·58	4·57	3·82	3·05	2·60
12	28·3	26·2	23·3	21·0	18·5	14·8	11·3	8·44	6·30	5·23	4·40	3·57	3·07
13	29·8	27·7	24·7	22·4	19·8	16·0	12·3	9·30	7·04	5·89	5·01	4·11	3·57
14	31·3	29·1	26·1	23·7	21·1	17·1	13·3	10·2	7·79	6·57	5·63	4·66	4·07
15	32·8	30·6	27·5	25·0	22·3	18·2	14·3	11·0	8·55	7·26	6·26	5·23	4·60
16	34·3	32·0	28·8	26·3	23·5	19·4	15·3	11·9	9·31	7·96	6·91	5·81	514
17	35·7	33·4	30·2	27·6	24·8	20·5	16·3	12·8	10·1	8·67	7·56	6·41	5·70
18	37·2	34·8	31·5	28·9	26·0	21·6	17·3	13·7	10·9	9·39	8·23	7·01	6·26
19	38·6	36·2	32·9	30·1	27·2	22·7	18·3	14·6	11·7	10·1	8·91	7·63	6·84
20	40·0	37·6	34·2	31·4	28·4	23·8	19·3	15·5	12·4	10·9	9·59	8·26	7·23
21	41·4	38·9	35·5	32·7	29·6	24·9	20·3	16·3	13·2	11·6	10·3	8·90	8·03
22	42·8	40·3	36·8	33·9	30·8	26·0	21·3	17·2	14·0	12·3	11·0	9·54	8·64
23	44·2	41·6	38·1	35·2	32·0	27·1	22·3	18·1	14·8	13·1	11·7	10·2	9·26
24	45·6	43·0	39·4	36·4	33·2	28·2	23·3	19·0	15·7	13·8	12·4	10·9	9·89
25	46·9	44·3	40·6	37·7	34·4	29·3	24·3	19·9	16·5	14·6	13·1	11·5	10·5
26	48·3	45·6	41·9	38·9	35·6	30·4	25·3	20·8	17·3	15·4	13·8	12·2	11·2
27	49·6	47·0	43·2	40·1	36·7	31·5	26·3	21·7	18·1	16·2	14·6	12·9	11·8
28	51·0	48·3	44·5	41·3	37·9	32·6	27·3	22·3	18·9	16·9	15·3	13·6	12·5
29	52·3	49·6	45·7	42·6	39·1	33·7	28·3	23·6	19·8	17·7	16·0	143	13·1
30	53·7	50·9	47·0	43·8	40·3	34·8	29·3	24·5	20·6	18·5	16·8	15·0	13·8
40	66·8	63·7	59·3	55·8	51·8	45·6	39·3	33·7	29·1	26·5	24·4	22·2	20·7
50	79·5	76·2	71·4	67·5	63·2	56·3	49·3	42·9	37·7	34·8	32·4	29·7	28·0
60	92·0	88·4	83·3	79·1	74·4	67·0	59·3	52·3	46·5	43·2	40·5	37·5	35·5
70	104·2	100·4	95·0	90·5	85·5	77·6	69·3	61·7	55·3	51·7	48·8	454	433
80	116·3	112·3	106·6	101·9	96·6	88·1	79·3	71·1	64·3	60·4	57·2	53·5	51·2
90	128·3	124·1	118·1	113·1	107·6	98·6	89·3	80·6	73·3	69·1	65·6	61·8	592
100	140·2	135·8	129·6	124·3	118·5	109·1	99·3	90·1	82·4	77·9	74·2	70·1	673

Source: Catherine M. Thompson, „Table of percentage points of the χ^2 distribution, Biometrika, Vol. 32 (1941), by permission of the author and publisher.

COURTESY OF HAAN (1987).

APPENDIX III—PERCENTILE VALUES($\chi^2_{\alpha,\nu}$) FOR THE CHI-SQUARE DISTRIBUTION WITH ν DEGREE OF FREEDOM (Shaded area=α)

APPENDIX IV - CRITICAL VALUES FOR THE KOLMOGOROV-SMIRNOV TEST STATISTIC

Sample Size (n)	Significance Level				
	.20	.15	.10	.05	.01
1	.900	.925	.950	.975	.995
2	.684	.726	.776	.842	.929
3	.565	.597	.642	.708	.829
4	.494	.525	.564	.624	.734
5	.446	474	.510	.563	.669
6	.410	.436	.470	.521	.618
7	.381	.405	.438	.486	.577
8	.358	.381	.411	.457	.543
9	.339	.360	.388	.432	.514
10	.322	.342	.368	.409	.486
11	.307	.326	.352	.391	.468
12	.295	.313	.338	.375	.450
13	.284	.302	.325	.361	.433
14	.274	.292	.314	.349	.418
15	.266	.283	.304	.338	.404
16	.258	.274	.295	.328	.391
17	.250	.266	.286	.318	.380
18	.244	.259	.278	.309	.370
19	.237	.252	272	.301	.361
20	.231	.246	.264	.294	.352
25	.21	.22	.24	.264	.32
30	.19	.20	.22	.242	.29
35	.18	.19	.21	.23	.27
40				.21	.25
50				.19	.23
60				.17	.21
70				.16	.19
80				.15	.18
90				.14	
100				.14	
Asymptotic Formula:	$\frac{1.07}{\sqrt{n}}$	$\frac{1.14}{\sqrt{n}}$	$\frac{1.22}{\sqrt{n}}$	$\frac{1.36}{\sqrt{n}}$	$\frac{1.63}{\sqrt{n}}$

Journal American Statistical Association 47:425-441,1952, Z.W. Birnbaum.

COURTESTY OF HAAN (1982)

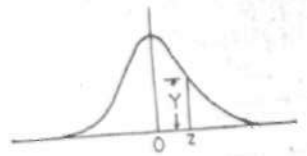
w	CDF	PDF	w	CDF	PDF	w	CDF	PDF
		0.00000 00	−0.50	0.19229 56	0.31704 19	3.5	0.97025 40	0.02929 91
−3.0		0.00000 02	−0.45	0.20839 66	0.32638 10	3.6	0.97304 62	0.02658 72
−2.9	0.00000 00	0.00000 12	−0.40	0.22496 18	0.33560 36	3.7	0.97557 96	0.02411 98
−2.8	0.00000 01	0.00000 51	−0.35	0.24193 15	0.34332 85	3.8	0.97787 76	0.02187 59
−2.7	0.00000 03	0.00001 91	−0.30	0.25927 69	0.34998 72	3.9	0.97996 16	0.01983 63
−2.6	0.00000 14							
		0.00006 24	−0.25	0.27692 03	0.35557 27	4.0	0.98185 11	0.01798 32
−2.5	0.00000 51	0.00017 99	−0.20	0.29481 63	0.36008 95	4.2	0.98511 63	0.01477 24
−2.40	0.00001 63	0.00029 29	−0.15	0.31291 17	0.36355 15	4.4	0.98779 77	0.01212 75
−2.35	0.00002 79	0.00046 47	−0.10	0.33115 43	0.36598 21	4.6	0.98999 85	0.00995 13
−2.30	0.00004 66		−0.05	0.34949 32	0.36741 21	4.8	0.99180 40	0.00816 23
		0.00071 89						
−2.25	0.00007 58	0.00108 63	0.0	0.36787 94	0.36787 94	5.0	0.99328 47	0.00669 27
−2.20	0.00012 04	0.00160 46	0.1	0.40460 77	0.36610 42	5.2	0.99449 86	0.00548 62
−2.15	0.00018 69	0.00232 09	0.2	0.44099 10	0.36105 29	5.4	0.99549 36	0.00449 62
−2.10	0.00028 41	0.00328 66	0.3	0.47672 37	0.35316 56	5.6	0.99630 90	0.00368 42
−2.05	0.00042 31		0.4	0.51154 48	0.34289 88	5.8	0.99697 70	0.00301 84
		0.00456 63						
−2.00	0.00061 80	0.00622 81	0.5	0.54523 92	0.33070 43	6.0	0.99752 43	0.00247 29
−1.95	0.00088 61	0.00834 67	0.6	0.57763 58	0.31701 33	6.2	0.99797 26	0.00202 51
−1.90	0.00124 81	0.01100 04	0.7	0.60850 53	0.30222 45	6.4	0.99833 98	0.00165 88
−1.85	0.00172 97	0.01426 93	0.8	0.63805 62	0.28669 71	6.6	0.99864 06	0.00135 80
−1.80	0.00235 87		0.9	0.66593 07	0.27074 72	6.8	0.99888 68	0.00111 25
		0.01823 15						
−1.75	0.00316 82	0.02296 12	1.0	0.69220 06	0.25464 64	7.0	0.99908 85	0.00091 11
−1.70	0.00419 46	0.02852 48	1.1	0.71686 26	0.23862 28	7.2	0.99925 37	0.00074 60
−1.65	0.00547 82	0.03497 81	1.2	0.73993 41	0.22286 39	7.4	0.99938 89	0.00061 09
−1.60	0.00706 20	0.04236 34	1.3	0.76144 92	0.20751 91	7.6	0.99949 97	0.00050 02
−1.55	0.00899 15		1.4	0.78145 56	0.19270 46	7.8	0.99959 03	0.00040 96
		0.05070 71						
−1.50	0.01131 43	0.06001 78	1.5	0.80001 07	0.17850 65	8.0	0.99966 46	0.00033 54
−1.45	0.01407 84	0.07028 48	1.6	0.81737 95	0.16498 57	8.5	0.99979 66	0.00020 34
−1.40	0.01733 20	0.08147 77	1.7	0.83363 17	0.15218 12	9.0	0.99987 66	0.00012 34
−1.35	0.02112 23	0.09354 65	1.8	0.84764 03	0.14011 40	9.5	0.99992 51	0.00007 48
−1.30	0.02549 44		1.9	0.86107 93	0.12879 04			
		0.10642 20				10.0	0.99995 46	0.00004 54
−1.25	0.03049 04	0.12001 76	2.0	0.87342 30	0.11820 50	10.5	0.99997 25	0.00002 75
−1.20	0.03614 86	0.13423 10	2.1	0.88474 45	0.10834 26	11.0	0.99998 33	0.00001 67
−1.15	0.04250 25	0.14894 68	2.2	0.89511 49	0.09918 16	11.5	0.99998 99	0.00001 01
−1.10	0.04958 01	0.16403 90	2.3	0.90430 32	0.09069 45	12.0	0.99999 39	0.00000 61
−1.05	0.05740 34		2.4	0.91327 53	0.08285 05			
		0.17937 41				12.5	0.99999 63	0.00000 37
−1.00	0.06598 80	0.19481 41	2.5	0.92119 37	0.07561 62	13.0	0.99999 77	0.00000 23
−0.95	0.07534 26	0.21021 95	2.6	0.92841 77	0.06895 69	13.5	0.99999 86	0.00000 14
−0.90	0.08546 89	0.22545 23	2.7	0.93500 30	0.06283 74	14.0	0.99999 92	0.00000 08
−0.85	0.09636 17	0.24037 84	2.8	0.94100 20	0.05722 24	14.5	0.99999 95	0.00000 05
−0.80	0.10800 90		2.9	0.94646 32	0.05207 75			
		0.25487 04				15.0	0.99999 97	0.00000 03
−0.75	0.12039 23	0.26880 94	3.0	0.95143 20	0.04736 90	15.5	0.99999 98	0.00000 02
−0.70	0.13348 68	0.28208 67	3.1	0.95595 04	0.04306 48	16.0	0.99999 99	0.00000 01
−0.65	0.14726 22	0.29460 53	3.2	0.96005 74	0.03913 41	16.5	0.99999 99	0.00000 01
−0.60	0.16168 28	0.30628 08	3.3	0.96378 87	0.03554 76	17.0	1.00000 00	0.00000 00
−0.55	0.17670 86		3.4	0.96717 75	0.03227 79			

† *Source:* National Bureau of Standards [1953], "Probability Tables for the Analysis of Extreme Value Data," Applied Math Series 22, Washington, D.C.

COURTESY OF BENJAMIN AND CORNELL (1970).

APPENDIX V — VALUES OF THE STANDARDIZED TYPE I EXTREME-VALUE DISTRIBUTION

(largest value) $F_W(w) = e^{-e^{-w}}$



APPENDIX VI – ORDINATES Y OF THE STANDARD NORMAL DENSITY

u	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.3989	0.3989	0.3986	0.3988	0.3986	0.3984	0.3984	0.3980	0.3977	0.3973
0.1	0.3970	0.3965	0.3961	0.3956	0.3951	0.3945	0.3939	0.3932	0.3925	0.3918
0.2	0.3910	0.3902	0.3894	0.3885	0.3876	0.3867	0.3857	0.3847	0.3836	0.3825
0.3	0.3814	0.3802	0.3790	0.3778	0.3765	0.3752	0.3739	0.3725	0.3712	0.3697
0.4	0.3683	0.3668	0.3653	0.3637	0.3621	0.3605	0.3589	0.3572	0.3555	0.3538
0.5	0.3521	0.3503	0.3485	0.3467	0.3448	0.3429	0.3410	0.3391	0.3372	0.3352
0.6	0.3332	0.3312	0.3292	0.3271	0.3251	0.3230	0.3209	0.3187	0.3166	0.3144
0.7	0.3123	0.3101	0.3079	0.3056	0.3034	0.3011	0.2989	0.2966	0.2943	0.2920
0.8	0.2897	0.2874	0.2850	0.2827	0.2803	0.2780	0.2756	0.2732	0.2709	0.2685
0.9	0.2661	0.2637	0.2613	0.2589	0.2565	0.2541	0.2516	0.2492	0.2468	0.2444
1.0	0.2420	0.2396	0.2371	0.2347	0.2323	0.2299	0.2275	0.2251	0.2227	0.2203
1.1	0.2179	0.2155	0.2131	0.2107	0.2083	0.2059	0.2036	0.2012	0.1989	0.1965
1.2	0.1942	0.1919	0.1895	0.1872	0.1819	0.1826	0.1804	0.1781	0.1758	0.1736
1.3	0.1714	0.1691	0.1669	0.1626	0.1604	0.1582	0.1582	0.1561	0.1539	0.1518
1.4	0.1497	0.1476	0.1456	0.1435	0.1415	0.1394	0.1374	0.1354	0.1334	0.1315
1.5	0.1295	0.1276	0.1257	0.1238	0.1219	0.1200	0.1182	0.1163	0.1145	0.1127
1.6	0.1109	0.1092	0.1074	0.1057	0.1040	0.1023	0.1006	0.0983	0.09728	0.09566
1.7	0.09405	0.09246	0.09089	0.08933	0.08780	0.08628	0.08478	0.0829	0.08183	0.08038
1.8	0.07895	0.07754	0.07614	0.07477	0.07341	0.07206	0.07074	0.06943	0.06814	0.06687
1.9	0.06562	0.06438	0.06316	0.06195	0.06077	0.05959	0.05844	0.05730	0.05618	0.05508

COURTESY OF HAAN (1982)