REGIONAL FLOOD-FREQUENCY ANALYSIS OF OGUN-OSHUN RIVER BASINS

BY

ODUSOTE, ABIMBOLA OLADIPUPO (B.Sc.)

A project report submitted to the Postgraduate School, Ahmadu Bello University, in partial fulfillment of the requirements for the award of the degree of Master of Science in Water Resources and Environmental Engineering.

Department of Water Resources and
Environmental Engineering,
Ahmadu Bello University,
Zaria, Nigeria,

August, 1989.

DECLARATION

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

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

Odusote, A.O.

Date: 5/1/90

This declaration is confirmed by

Dr. S. Mustafa

Date: 5/1/90

CERTIFICATION

This project report entitled "Regional Flood-Frequency Analysis of Ogun-Oshun Basins" by Abimbola Oladipupo ODUSOTE meets the regulations governing the award of the degree of Master of Science of Ahmadu Bello University, and is approved for its contribution to knowledge and literary presentation.

Chairman, Supervisory Committee — Date: 5/1/90

Member, Supervisory Committee — Date: 5/1/90

Head of Department — Date: 5/1/90

Dean, Postgraduate School — Date: 8/3/90

D E D I C A T I O N

To the Lord Jesus, and my dear parents,

Mr. and Mrs. A. O. Odusote.

ACKNOWLEDGEMENTS

First of all, I lavish encomium of praises on the Almighty God who had mercy on my soul during the course of my study here in Ahmadu Bello University. He has also been bombarding me with blessings in recent times. I thank you Lord.

I am deeply indebted to my able supervisor, Dr. Salihu Mustafa for his valuable advice and assistance not only during the execution of this project but throughout my period of postgraduate studies. I am mostly grateful for the care he has for me by granting me audience at any time I called on him. Thank you Sir.

I gratefully acknowledge the help rendered by the following people in the course of data collection, without which the study could not have been undertaken: Dr. E. S. Oyegoke of the Hydraulics Research Unit, University of Lagos, Mr. Augustine Okojie of Hydrological Section, Iju Water Works, Lagos State Water Cooperation, Mr. Yinusa Yussuf and Co. of the Ogun-Oshun River Basin Authority.

I acknowledge the concern shown by my course mates during the execution of the project and the cordial relationship that exists between us.

I cannot but thank my auntie, Mrs. Oloworaran, and children who are here in Zaria and have given me the unquantifiable care which I mostly needed during my sojourn in this beautiful side of the country. I also thank people like Mr Yemi Adejumo who I knew through my auntie and have also

been concerned.

Finally, I thank my parents, brothers and sisters for their moral and financial support. Without them I will be no where today.

ABSTRACT

This project describes a procedure for estimating the design floods from regional frequency curves for ungauged areas drained by rivers Ogun and Oshun of Western Nigeria. Streamflow records at nine gauging stations were used for the study.

For estimating purposes, two relations have been derived. The first relates graphically the mean annual flood to basin physiographic factors namely: the size of drainage area and the length of main channel. The second relates the mean annual flood to the physiographic factors empirically. From these, the mean annual floods for a basin is estimated and used to obtain design flood from a regional flood frequency quency curve derived for the basin. The regional flood frequency curve expresses the flood discharge-time relation, showing variation of peak discharge, expressed as a ratio to the mean annual flood, with recurrence interval.

TABLE OF CONTENTS

		PAGE
TITLE PAGE		i
DECLARATION		ii
CERTIFICATION		iii
DEDICATION		iv
ACKNOWLEDGEMENT		v
ABSTRACT		vii
TABLE OF CONTENTS		viii
LIST OF TABLES		xi
LIST OF FIGURES		xii
LIST OF APPENDICES		xiv
1.	INTRODUCTION	1
2.	REVIEW OF LITERATURE	5
2.1	Introduction	5
2.2	Flood-Flow Computation	5
2.2.1	Deterministic Approach	7
2.2.2	Probabilistic Approach	10
	Frequency Distribution	11
	Probability Distribution Function	13
	Plotting Position	19
3.	REGIONAL FLOOD FREQUENCY STUDY	25
3.1	Regional Flood Frequency Analysis	25
3.2	Basic Frequency Curve	26
3.3	Study Area	27
3.3.1	Location	27

TABLE OF CONTENTS (CONT'D)

		PAGE
3.3.2	Climate	27
	Rainfall	27
	Temperature	29
	Humidity	29
	Wind System	30
	Cloud	30
3.3.3	Vegetation	30
3.3.4	Hydrology and Hydrogeology	31
3.3.5	Human Activities	35
3.4	Characteristics of the Data	36
4.	ANALYSIS OF DATA	42
4.1	Selection of the Base Period	42
4.2	Adjustment of Records	42
4.3	Preliminary Frequency Curve	43
4.3.1	Log-Normal Distribution	50
4.3.2	Gumbel Extreme-Value Distribution	52
4.3.3.	Log Pearson Type III Distribution	52
4.4	Selection of the Best Fitted Distribution	52
4.4.1	Confidence Band	53
4.4.2	Goodness of Fit Test	66
	Kolmogorov-Smirnov Test	66
4.5	Mean Annual Flood	68
4.6	Homogoeneity Test	70
4.7	Computation of Median Flood Ratios.....	74
4.8	Definition of Regional Frequency Curve .	74
4.9	Estimation of Mean Annual Flood	74

TABLE OF CONTENTS (CONT'D)

		PAGE
4.9.1	Graphical Correlation	77
4.9.2	Mathematical Relationship Between Flood and Physiographic Factors	81
4.10	Determination of a Frequency Curve	90
5.	DISCUSSION OF RESULTS	91
	Determination of Missing Records ...	91
	Goodness-of-fit Test	92
	Homogeneity Test	92
	Regional Frequency Curve	92
	Estimation of Annual Flood	92
	Design Flood and Risk Analysis	94
	Advantages and Disadvantages of the Study	97
6.	CONCLUSION AND SUGGESTION	99
6.1	Conclusion	99
6.2	Suggestion	100
	REFERENCES	102
	APPENDICES	105

LIST OF TABLES

		PAGE
Table 3.1	Characteristics of the Rivers under study area	38
Table 3.2	Maximum annual peak discharges (m^3/s)	39
Table 4.1	Coefficients for the calculation of the 90 percent confidence intervals around the flood frequency curve	56
Table 4.2	Critical statistics for the Kolomogorov-smirnov goodness-of-fit test	68
Table 4.3	Data for homogeneity test	69
Table 4.4	Confidence limits of regional flood homogeneity test	71
Table 4.5	Median flood ratios	75
Table 4.6	Observed and graphically estimated mean annual floods	81
Table 4.7	Data for regression analysis	84
Table 4.8	Observed and mathematically estimated mean annual floods.....	85
Table 4.9	Data for significance test.....	87

LIST OF FIGURES

		PAGE
Figure 3.1	Location of the study area	28
Figure 3.2	Geographical distribution of the gauging stations in the study area..	34
Figure 3.3	Bar graph showing period of record of maximum annual peak flood at gauging stations	41
Figure 4.1	Correlation of annual floods, station 3 versus station 9	44
Figure 4.2	Correlation of annual floods, station 4 versus station 9	45
Figure 4.3	Correlation of annual floods, station 5 versus station 9	46
Figure 4.4	Correlation of annual floods, station 6 versus station 9	47
Figure 4.5	Correlation of annual floods, station 7 versus station 9	48
Figure 4.6	Correlation of annual floods, station 8 versus station 9	49
Figure 4.7	Frequency distribution of discharge at streamflow gauging station 1	57
Figure 4.8	Frequency distribution of discharge at streamflow gauging station 2	58
Figure 4.9	Frequency distribution of discharge at streamflow gauging station 3	59
Figure 4.10	Frequency distribution of discharge at streamflow gauging station 4 ...	60

LIST OF FIGURES (CONT'D)

		PAGE
Figure 4.11	Frequency distribution of discharge at streamflow gauging station 5	61
Figure 4.12	Frequency distribution of discharge at streamflow gauging station 6	62
Figure 4.13	Frequency distribution of discharge at streamflow gauging station 7	63
Figure 4.14	Frequency distribution of discharge at streamflow gauging station 8	64
Figure 4.15	Frequency distribution of discharge at streamflow gauging station 9	65
Figure 4.16	Homogeneity test chart	73
Figure 4.17	Regional frequency curve	76
Figure 4.18	Variation of mean annual flood with drainage area	78
Figure 4.19	Variation of mean annual flood with length of main channel	80

LIST OF APPENDICES

		PAGE
Appendix A-1	Exceedence probability analysis of maximum annual peak discharges at station 1	105
Appendix A-2	Exceedence probability analysis of maximum annual peak discharges at station 2	106
Appendix A-3	Exceedence probability analysis of maximum annual peak discharges at station 3	107
Appendix A-4	Exceedence probability analysis of maximum annual peak discharges at station 4	108
Appendix A-5	Exceedence probability analysis of maximum annual peak discharges at station 5	109
Appendix A-6	Exceedence probability analysis of maximum annual peak discharges at station 6	110
Appendix A-7	Exceedence probability analysis of maximum annual peak discharges at station 7	111
Appendix A-8	Exceedence probability analysis of maximum annual peak discharges at station 8	112

LIST OF APPENDICES (CONT'D)

		PAGE
Appendix A-9	Exceedence probability analysis of maximum annual peak discharges at station 9	113
Appendix B-1	Regional frequency curve on log-Pearson type III paper.....	114
Appendix B-2	Regional frequency curve on log-normal paper.....	115

Chapter 1

INTRODUCTION

KASHINI IBRAHIM
AHMADU BELLO UNIVERSITY
ZARIA, NIGERIA.

Frequent flood disasters in some parts of Nigeria in recent times have been a major concern to hydrologists, environmental engineers and allied professionals. For example, the flood disasters of 1978 and 1980 in Ibadan, which were considered to be unsurpassed prior to their occurences were nothing to compare with the major floods that occurred in almost every part of the country in 1988.

The need to know the frequency associated with magnitudes of such floods in order to find ways of predicting further occurrence of the event is usually of great concern to water resources planners.

Generally speaking, forecasts of floods events can reduce the damages caused by floods. There are two ways in which this beneficial effect may be achieved, Kite (1977). Firstly, and most obviously, warning of an event enables people to evacuate a danger area. If sufficient lead time is provided, vulnerable possessions may be removed from the danger zone and preparations can be made, such as sandbagging, to minimize property damage.

Knowledge of the magnitude and probable frequency of recurrence of floods is necessary to the proper design and location of flood control structures, dams, bridges, culverts, highways, waterworks, sewage-disposal plants, and industrial

buildings. An engineer often must design a structure which may be damaged or destroyed by occasional floods of varying magnitude. The frequency with which such damage may occur must be considered in determining the site or strength of the structure, its location, or the feasibility of building it at all. The problem is an economic one, involving computation of the total annual cost of maintainng a structure of a given design compared to the cost for other design, Dalrymple (1960).

Frequency analysis is not only of use as an aid in averting disaster but is also a means of introducing efficient designs. When a hydraulic structure, through inadequate or inaccurate data or methods, is underdesigned, the results are regretably obvious; the dam may fail, the highways may flood or the bridge collapse. This does not happen very often and so the hydrologist, equating non failure with success, is satisfied with his design techniques. Non failure, however, does not necessarily mean an efficient design techniques. Often, structures are overdesigned and hence very safe, but also very expensive. A truly efficient design will be achieved only as the result of studies relating cost to risk and frequency analysis.

In this study, a flood frequency relationships for a region drained by rivers Ogun, Oshun, Ofiki and Oba in western Nigeria have been determined. The objective of the study is two fold:-

1. To find a relationship between the physiographic factors, for example drainage areas, of the region and an index flood which is the mean annual flood.

2. To find the regional frequency curve. That is, discharge, in ratio to mean annual flood plotted against the recurrence interval, in years(i.e. the interval, in years, in which a flood of a given magnitude will be equalled or exceeded).

With these two relationships the design flood for any structure to be constructed at any site within the region can be determined simply by obtaining the mean annual flood at such site if its physiographic factor(s) is/are known and then using the mean annual flood to obtain the design flood having known the recurrence interval used for the type of structure to be constructed.

This approach is prefarable to those developed for specific sites for two reasons:-

(a) Because of the sample variation possible at a single station, any single station is subject to large error. This error could, however, be reduced by combining data from many sites.

(b) There are generally, many more sites where hydrologic data are needed than there are sites at which data are collected. This means that some form of analysis is required which can transfer data from gauged sites to ungauged sites.

Both single station and regional frequency analyses involve risks. In determining the design flood for a project, the length of data available, the project life and the allowable probability of failure are all factors to be considered.

Chapter two of the thesis reviews the various ways of flood computation if data available are so small as to make the use of flood frequency analysis very risky. Also reviewed are the various probability distributions the data at a gauging station can take, and the plotting formulae.

The rational behind the development of the regional flood frequency analysis, and the characteristics of the study area are provided in Chapter Three.

Chapter four deals with data analysis.

Discussion of results and the risk associated with the design of water resources structures are provided in Chapter five. Conclusion and Recommendation is the theme of Chapter six.

Chapter 2

LITERATURE REVIEW

2.1 Introduction

One of the ways in the hydrologic design of hydraulic structure like dams, weirs, and so forth, is the statistical analysis of peak discharges and/or maximum flood volumes in the series or array of flood-flow data for the site. The purpose of such analysis is to obtain a relation between the discharge magnitude and the probability of exceedence. Because the data used are usually annual maximum values, the exceedence probability obtained usually refers to probability of annual exceedence. The methods of probability analysis are graphical and functional or analytical methods.

This chapter reviews the various ways of determining the flood flows at a site under the sub-heading "Flood-flow Computation". The various ways of functional analysis and graphical procedure are also reviewed under the sub-heading "Frequency Distribution" and "plotting Positions" respectively.

2.2 Flood-Flow Computation

There are many situation in which an estimate of future flood condition is required and many different categories of individual investigator, industry, government agency or other group requiring such information.

Essentially, however, this information is needed for either design or forecasting purposes.

In the design situation engineers (involved in the design of, e.g., dams, spillways, river channel improvements, storm sewers, bridges and culverts), planners, farmers and so forth, need information on flood magnitude and frequency. For example, there are trials to answer the question: What is the maximum flood which is likely to occur at a given place and what is the probability that a flood of a certain magnitude will be exceeded at a particular place once in any given numbers of years? In the forecasting situation local government agencies, industrialists, farmers and home owners require more immediate information on flood magnitude and timing so that appropriate evasive action may be taken. It is necessary to forecast that a flood of given volume, depth and duration will occur within a specified time.

Various methods are available for estimating design floods and some of these are reviewed in this section. Broadly speaking these methods fall into two main categories: those which are deterministically based, which treat floods as the product of specified precipitation falling upon a specified catchement area, and those which are probabilistically based, which treat floods as almost purely random events susceptible to statistical analysis.

2.2.1 Deterministic Approach

Ward (1978) wrote that "Floods are physical phenomena which result from an input of precipitation into a drainage basin, the flood magnitude varying with the nature of both the precipitation and the drainage basin". There have been many attempts to relate precipitation and catchment variables in order to estimate maximum flood discharges.

Chow (1964) has reported 107 empirical formulae still in use in the estimation of flood discharges, although there are other similar techniques involving tables and simple rules of thumb. The simplest empirical formulae for flood discharge use the single parameter of drainage area and take one of the following forms:-

$$Qm = CA^{n} \qquad 2.1$$

$$Qm = CA^{m}, A^{-n} \qquad 2.2$$

$$Qm = \frac{CA}{(a+bA)} \; m + dA \qquad 2.3$$

where Qm is the maximum flood discharge, A is the catchment area, n and m are exponents and c, a, b, d are co-efficients which must be evaluated for an area, depending upon its geographical and climatological characteristics.

Other, more elaborate, empirical formulae relying on a broader or more detailed data base have incorporated basin characteristics in addition to area, such as precipitation, length of main channel, and mean catchment slope, and have often also included an explicit frequency term. One of the oldest and certainly one of the most widely used formula which maintains much of its popularity among practising engineers, particularly in terms of flood discharges into urban storm sewers and from small, relatively homogenous rural catchments, is the rational method. The origins of this method are somewhat obscure. Chow (1964) pointed out that it was first mentioned in American literature by Kurichling E, in 1889, although some other believe that the principle may be ascribed to the earlier work of Mulvaney J.J., in 1851, while in England, the rational method has often been referred to as the Lloyd-Davis method. The method assumes that the maximum flood discharge from a catchment will occur when the entire area is contributing to run-off and may be expressed in the form

$$Qm = CIA \qquad 2.4$$

where Qm is as described above, C is a runoff coeficient indicating the percentage of rainfall which appears as quickflow, I is the mean rainfall intensity during the period of concentration, Tc, that is, the time required for the most distant part of the catchment to contribute to outflow from the catchment area, and A is the area of the catchment.

The United States Development Agency Soil Conservation Service, Ward (1978), proposed a formula incorporating area, rainfall characteristics and a time parameter to describe a simple, triangular flood hydrograph

$$Qm = \frac{65CARe}{T_p} \qquad 2.5$$

where Qm and A are as described above, Re is the rainfall excess in millimetres, and Tp is the time to peak or the time of rise of the hydrograph which can be estimated empirically from catchment characteristics.

The weakness in the use of empirical methods is their 'black-box' approach and their failure to incorporate a proper understanding of flood peaks was given by the works of Sherman (1932) on the unit hydrograph and Horton (1933) on the role of infiltration in the run-off process. In the case of the unit hydrograph, some of the most valuable work concerned the development of the instantaneous unit hydrograph and synthetic hydrographs, particularly in terms of flood prediction from ungauged catchments. Horton's infiltration theory has subsequently manifested itself in flood prediction methods both, directly, in the form of infiltration indices, and indirectly in the form of indices of antecedent precipitation, soil moisture and the proportion of the catchment under different land-use types.

2.2.2 Probabilistic Approach

The probabilistic approach regards floods as random streamflow events whose distribution can be identified and whose probability of occurrence during a given period of time can be calculated using standard statistical techniques. The longer the period under consideration the larger is the maximum flood experienced likely to be and the more frequent is a flood of modest magnitude liable to occur. The average time elapse between two events which equal or exceed a particular level is the return period, Tr, or recurrence interval. Thus the N-year flood, a flood which is expected to be equalled or exceeded on average every N-years, has a return period of N-years. It has been emphasized over and over again by many authors and hydrologists that the N-year return period flood can occur in any year, but, that the probability of its occurrence in any number of years from now differ. For example, naturally the 100-year flood could occur within fifty years, five years or even one year, but, the probability of its occurrence during 100-years is obviously much greater than during a five-year period, and greater still during 1,000 years of record.

The statistical relationships between the magnitude of a flood and the frequency with which it is likely to occur have always been investigated in two ways. First, given a return period, the instantaneous flood peak which will be equalled or exceeded once on average during that time period can be estimated.

Secondly, given a particular flood magnitude, the average frequency with which it will be equalled or exceeded (i.e., the return period) can be estimated.

Chow (1964) noted that graphical techniques of probability analysis involving the use of flow-duration curves were in use in the United States during the latter part of the ninetenth century, although the sparcity of adequate data hindered development for some decades. It was during the 1920s and 1930s that the probability a approach became most popular in the United States, with standard methods based upon either the Gaussian or Galton (log-probability) laws. Later, Gumbel (1941) began to apply the Fisher-Tippet theory of extreme values to flood frequency studies. Gumbel's numerous publications in this field culminated in the appearance of the first book devoted entirely to the statistics of extreme, Gumbel(1958).

The remaining part of this chapter is devoted to probabilistic approach to flood estimation.

KASHIM IBRAHIM LIBRARY
AHMADU BELLO UNIVERSITY
ZARIA, NIGERIA.

Frequency Distribution

One of the problems encountered in hydrology is the estimation of a design flood from a fairly short record of stream-flows. Plotting the magnitudes of the measured events (e.g. annual maximum), some kind of pattern is generally apparent. The question is how to use this pattern to extend the available data and estimate the design event reliably.

The time series of flood discharges at a site provides a statistical sample of the flood population that can be described by probability distribution functions. These probability distribution functions are also presented in form of graphs so that the statistical sample can be plotted out on the graph that best fits the distribution, thereby avoiding the rigours of parameter estimation in using the functions directly.

When the distribution functions are used directly, Sokolov et al (1976) found that "records of annual maximum discharges are invariably too short to permit the computation of more than three parameters, by the method of moments, because the standard errors of parameters of the higher orders are large for short arrays". Consequently, the theoritical probability distributions used in hydrologic analysis are usually limited to those that can be defined by the three Parameters: arithmetic mean: standard deviation or coefficient of variation, which is the standard deviation divided by the arithmetic mean; and the coefficient of skewness. The theoritical distribution curves selected for use also, are those that do not have negative values of the random variable, because flood discharges is always a positive value.

Of the many theoritical probability distributions that might be used in flood-frequency studies, those most widely used are: Normal distribution, logarithmic-normal distribution, Pearson type III distribution, log-Pearson

type III distribution, and extreme-value (Gumbel) distribution.

Probability Distribution Function

In his book, Kite (1977) explained that a distribution is said to be normal if the variable can take any value form - 00 to + 00. The probability function is defined as:

$$P(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp - \left[\frac{(x-\mu)^2}{2\sigma^2}\right] \qquad \text{---- 2.6}$$

where μ and σ are the distribution parameters namely: the population mean and standard deviation of the variable. The normal distribution is applicable if:-

1. The variable is continuous
2. Consecutive values are independent
3. Probabilities are stable.

One of the feature of the normal distribution is that the mean, mode and median are all the same.

If the variable, x, is standardized, that is, forced to a mean of zero and unit variance by substracting the mean and dividing by standard deviation, and is denoted by t then equation (2.6) becomes

$$P(t) = \frac{1}{\sqrt{2\pi}} \left[\exp - \frac{(t^2)}{2}\right] \qquad \text{---- 2.7}$$

which is known as the standard normal distribution.

A related distribution to the normal distribution which is used for flood-flow analysis is the logarithmic normal distribution. The distribution involves the transformation of the random variable from natural units to logarithmic units. If the logarithmic, ln x, of a variable x are normally distributed, then the variable x is said to be logarithmic-normally distributed so that from equation (2.6),

$$P(x) = \frac{1}{\sigma_y \sqrt{2\pi}} \exp - \frac{(\ln x - \mu_y)^2}{2\sigma_y^2} \quad \text{---- 2.8}$$

where μ_y and σ_y are the mean and standard deviation of the natural logariths of x.

Chow (1964) has provided a theoritical justification for the use of the lognormal distribution. The causative factors for many hydrologic variables act multiplicatively rather than additively, and so the logarithms of these factors will satisfy the three basic conditions for normal distributions. The hydrologic variable will then be the product of these causative factors.

The Pearson type III distribution is another type of distribution but it is not directly used in the probability studies of food-flow. Instead, the log-Pearson type III distribution is used. The probability density function of the Pearson type III distribution is of the form:

$$P(x) = \frac{1}{\alpha \Gamma(\beta)} \left(\frac{x - \gamma}{\alpha}\right)^{\beta - 1} \exp - \frac{(x - \gamma)}{\alpha} \quad \text{---- 2.9}$$

where α, β and γ are the scale, shape and location parameters respectively and $\Gamma(\beta)$ is the gamma function.

If the logarithms, ln x, of a variable x are distributed as a Pearson type III variate then the variable x will be distributed as a log-Pearson type III with probability density function,

$$P(x) = \frac{1}{\alpha \Gamma(\beta)} \left(\frac{\ln x - \gamma}{\alpha}\right)^{\beta - 1} \exp -\left(\frac{\ln x - \gamma}{\alpha}\right)$$

---- 2.10

where α, β, γ and $\Gamma(\beta)$ are as defined above.

Being a 3-parameter distribution operating upon logarithms of the variable the log-Pearson type III would appear to be an extremely versatile distribution. However, its application in hydrology is strictly limited.

The last, but not the least, type of distribution is the type I extremal distribution. Suppose that from N samples each containing n events the maximum or minimum event in each sample is selected. As n increases, the distribution of the N maxima or minima approaches a limiting or assymptotic form, Kite (1977).

The type of the limiting form depends on the type of the initial distribution of the Nn values. The distribution of the maxima or minima is given by the functional equation.

$$P(x) = P(a_n + b_n) \qquad \text{---- 2.11}$$

where a_n and b_n are functions of n.

Fisher and Tippet, (Verma et al 1973) have shown that there are three possible solutions to the functional equation. These are known as types I, II and III extremal distributions. The type I distribution is unbounded, the type II has a lower limit and the type III has an upper limit.

The type I distribution Gumbel (1959), is often used for maximum type events and results from any initial unlimited distribution of exponential type which converges to an exponential function.

The derivation of the type I distribution for a simple exponential function has been described, Samuelson (1972) as follows:

(a) Let $\varepsilon_1, \varepsilon_2 \ldots\ldots\ldots, \varepsilon_n$ be a series of independent random variables with cummulative probability distribution given by:

$$P(x) = P(\varepsilon_v \leq y) \qquad \text{---- 2.12}$$

(b) Define x_n as the maximum value of ξ in a sample length n, that is, maximum ξ_v so that $1 \leq v \leq n$

$$P(x_n < y) = P(\xi_1 < y, \xi_2 < y, \ldots\ldots, \xi_n < y) \quad \text{----2.13}$$

Or $$P(x_n \leq y) = (P(y))^n \quad \text{---- 2.14}$$

in the case where all the ξ_v are identically distributed.

(c) Now assume that the tail of the distribution is exponential such that

$$P(y) = 1 - \propto e^{-y} \quad \text{---- 2.15}$$

(d) From equation 2.14 if $\ln(\propto n)$ is a normalising constant

$$P(x_n \leq y + \ln(\propto n)) = (P\ (y+\ln(\propto n))^n \quad \text{----2.16}$$

and from equation (2.15)

$$P(y+\ln(\propto n)) = 1-\propto e^{-(y+\ln(\quad n))} \quad \text{----2.17}$$

so that

$$P(x_n \leq y + \ln(\propto n)) = (1-\propto e^{-(y+\ln(\propto n)}{}^n \quad \text{----2.18}$$

or $$P(x_n \leq y + \ln(\propto n)) = (1-e^{-y/n}{}^n \quad \text{----- 2.19}$$

(e) If $n \to \infty$ then

$$\lim_{n\to\infty} p(x_n \leq y+\ln(\propto n)) = \lim_{n\to\infty} (1-e^{-y/n}{}^n \quad \text{----2.20}$$

or

$$\lim p(x_n \leq y+\ln(\propto n)) = e^{-e^{-y}} \quad \text{---- 2.21}$$

This is the reduced form of the cummulative probability distribution. Substituting the expression

$$y = \alpha(x - \beta) \qquad \text{---- } 2.22$$

where α is a concentration parameter and β is a measure of central tendency, the cummulative probability of the type 1 extremal distribution becomes

$$P(x) = e^{-e^{-\alpha(x-\beta)}} \qquad \text{---- } 2.23$$

The above distribution functions are the major ones out of a handful that are commonly used in the application of probability theory to flood frequency analysis. The familiarity of these distributions, in fact, leads to their frequent adoption simply for reasons of computational expediency. Even though there may exist no argument suggesting that a particular distribution is appropriate, it is often convenient to have a simple mathematical function to describe a variable.

In practise, both the plotting and comparison of cummulative curves can be simplified by scale changes, that is, by special plotting paper. Such graph paper is called probability paper. This probability paper provides properly scaled ordinates such that the cummulative distribution function of the probability law plots as a straight line.

With such paper, comparison between the function and data is reduced to a comparison between the cummulative frequency polygon of the data (plotted on the paper) and a straight line. Fitting a straight line to a set of plotted points aids in the comparison of curves between stations and in extrapolation.

Also the data plotted on different probability papers enables us to determine the function that best fits the data because the more the data plotted on a probability paper assumes a straight line the more the model fits the data.

The only criteria, Benjamin and Cornell (1970), for choosing one of the above probability distributions of annual maximum floods are convenience and goodness-of-fit.

Plotting Position

As stated earlier, given a series of N annual maximum discharges X, it is imperative to plot these annual maxima on a graph in order to better interpret the data, perhaps detect errors, or to get an idea of which probability distribution to use in describing the data. The event magnitude is usually plotted against the probability of occurrence or exceedence of such magnitude or average time interval between occurrence (since this is simply the inverse of the probability of such occurrence).

It has been realized, Yevjarich (1972), that the probability of occurrence of the event of a certain magnitude varies with the sample size N. The problem of apparent return period varying with sample length is approached in practise by defining a plotting position for the frequency of occurrence.

Gumbel (1958) summarizes five conditions that a plotting Position $P(X_m)$ of the m-th ordered value should satisfy (m is the ordered value of an event of the series by arranging the event in descending order so that m = 1 for large event and m = N for the smallest event):

1. All observations can be plotted on graph paper between P(X) = 0 and P(X) = 1, without the impossible one.
2. The plotting position of the m-th value is distribution-free, or it should be independent of the underlying distribution P(X), and it should lie between the observed frequencies M/N and (m-1)/N, if the first condition is to be satisfied that $P(X_m)$ is never either one or zero.
3. Probabilities of exceedence of the largest observed value or of non-exceedence of the smallest observed value should be close to the relative frequency 1/N, or the inverse of the sample size.

4. Observations plotted on the frequency scale should be equally spaced, that is, the difference between the plotting positions of the (m + 1) - th and the m-th ordered value should be a function of N only, independent of m.

5. The plotting position should be analytically simple, with a good intuitive meaning.

Various hydrologists have developed different methods of determining the plotting positions for the graphical methods of probability analysis.

California department of Public Works, Dalrymple (1960), developed the simplest method in which the probability

$$P(X) = \frac{M}{N} \qquad \text{---- 2.24}$$

It is at present known as the California method. The second plotting position developed alongside that of equation (2.24) is

$$P(X) = \frac{M - 1}{N} \qquad \text{---- 2.25}$$

The disadvantages of equation (2.24) and (2.25) is that the first gives $P(X) = 1$ for $M = N$, and the second give $P(X) = 0$ for $m = 1$, which are the sure and the impossible events, respectively. Neither should appear in a finite sample. Since the frequencies zero and unity do not exist for an unlimited variate the largest observation of the series cannot be plotted using the function $(m-1)/N$ and the smallest observation

cannot be plotted using the function M/N. These two functions therefore fail Gumbel's conditions and are not acceptable as plotting positions unless an upper or lower limit to the population can be envisaged.

Hazen (1930) proposed a function for plotting the probability as

$$P(X) = \frac{m - 0.5}{N} \qquad \text{---- 2.26}$$

which is between the two plotting positions of equations (2.24) and (2.25). The exceedence probability of the largest discharge in the series, when computed using equation (2.26) is 1/2N, and the recurrence interval (which is the inverse of probability) is 2N. The disadvantage of this function is that it gives the recurrence interval of the highest flood to be 2N, which is an artificial lengthening of the period of record. To overcome this, Gumbel (1958) proposed a variate correlation factor C to be multiplied with the function. The advantage of the function is that it satisfies all of the five conditions indicated by Gumbel (1958).

The most practical plotting position which fully satisfies all five of Gumbel condition is

$$P(X) = \frac{m}{N + 1} \qquad \text{---- 2.27}$$

Equation 2.27 has been adopted by Gumbel and many adherents of his theories. The property of equation (2.27) is that for m = 1 it gives $P(X_1) = 1/(N + 1)$, and for m = N, it yields $P(X_n) = N/(N + 1)$, the first value

been close to the frequency 1/N, and the last value been smaller than one. The use of equation (2.27) dominates at present the hydrologic applications.

The literature given by Chow (1964) summarizes other expressions proposed for plotting positions. Yevjevich (1972) gave the opinion that these expressions are not especially useful in hydrology because they do not have more advantages than those of equations (2.24) to (2.27). Expressions for the plotting position, given by Chow (1964) and elsewhere produce relatively small differences in comparison with the four-plotting positions reviewed, when they are properly applied.

In this study, equation (2.27) is adopted as the plotting position on graph papers representing the log-normal, log-pearson and the Gumbel extreme-value distribution for the arrays of discharge data at nine streamflow gauging stations within the study area. The best fitting distribution out of the three is selected for each of the nine stations, and regional flood frequency analysis is performed. Beard (1962) presented a method for constructing bands within which, at least, 90% of the plotted positions must fall in order to fit a theoretical distribution.

The regional flood frequency analysis is one such method which facilitates the computation of flood potential and design floods. The technique was first developed by Dalrymple (1960) of the United State Geological Survey.

The computational expediency and reliability resulting from it warranted its application in preparing about 25 regional studies for all parts of the United States. The method provides two relationships. One relating the physiographic factors of the region with the mean annual flood, Benson (1964), and the other representing the flood-frequency curve applicable to the region.

If the physiographic factors at any site in the region on which an hydraulic structure is to be constructed is known, then the mean annual flood at such site can be obtained. The design flood for such a structure can then be determined from the flood frequency curve applicable to the region.

Chapter 3

REGIONAL FLOOD FREQUENCY STUDY

3.1 Regional Flood-Frequency Analysis

Flood records at a gauging station are generally short, the sampling errors correspondingly large, and the record represent different period of time. In addition it is only rarely that flood-frequency information is needed at a gauging-station site. More often it is required at an ungauged site.

Dalrymple (1960) developed a method of combing records within a region. This method reduces the sampling error, bases the result on a uniform period of experience, and produces flood frequency relations generally applicable within the region. Dalrymple (1960) noted that "experience has shown that such result is not only possible but leads to result of acceptable accuracy for design and planning.

Regional flood-frequency study consists of two major parts. The first is the development of basic, dimensionless frequency curve representing the ratio of the flood at any frequency to an index flood (the mean annual flood). The second part is the development of relations between physiographic factors (e.g. the drainage areas or the length of main channel) and the mean annual flood, to enable the mean annual flood to be predicted within the region. The combination of the mean annual

flood with the basic frequency curve which is in term of the mean annual flood, provides a frequency curve for any location.

3.2 Basic Frequency Curve

Throughout large regions, which are homogenous with respect to flood-producing characteristics, individual streams have a wide range of drainage area will have frequent curve of about equal slope or steepness. If the peak flood at each gauging station are reduced to dimensionless ratios by dividing by an index flood which is the 2.33-year flood, the curves plotted, using the flood ratios can be superimposed and will nearly coincide.

These curves will all pass through the 2.33-year frequency at the ratio 1.0, but will have somewhat different slopes. The variation of these slopes can be tested to see whether the spread could occur by chance among samples from the same population (this is the "homogeneity test", which uses the ratio of the 10-year flood to the 2.33-year flood as the slope). If so, the assumption of a homogenous region containing all the stations is satisfactory. Within such a homogenous region, the best representation of the flood-frequency relation can be obtained by combining all dimensionless curves. The resulting average-frequency curve is then applicable throughout the region, and is called the

regional frequency curve.

If the spread in slope is greater than can be attributed to chance alone, then further separation of the region is made to produce two or more homogenous region.

Combining individual curves to obtain the regional frequency curve is accomplished by taking the median of station flood ratios at the same recurrence intervals (see next chapter). These median values are then plotted to define the regional frequency curve.

3.3 Study Area

3.3.1 Location

The region covers 44,000 km^2 and lies between 6 and 9 degrees north latitude and 2 and 5 degrees east longitude. It comprises a flat humid coastal belt broken by rivers and lagoons, which is mainly mangrove swamp, and a forest belt 160 to 240 km inland from the coast, at an altitude of less then 300m. Figure 3.1 indicates the location of the study area.

3.3.2 Climate

Rainfall: Rainfall as a whole is of average in the region, averaging about 1700mm a year. There is a principal rainy season in May, June and first half of July, and a secondary rainy season in the latter half of September and October.



NOTES:

Regional drainage area boundary

• Towns and Cities

Scale : 1 : 9,000,000

FIG. 3·1 : Location of the study Area .

In individual months, departures of the rain-fall from the average can be considerable.

Rainfall is brought into the region by the south-westerly wind. This is very moist, and when it prevails in sufficient depth it gives cloudy weather, frequently with afternoon and evening thunderstorm, period of moonsoon rain near the coast, and periods of mist in the early mornings.

Temperature: The annual mean of daily maximum temperatures is about 32^{o}C in the region. Maximum temperatures are highest from February to April. They are lowest in July and August.

Humidity: Near the coast at dawn-normally the time of minimum temperature-relative humidity is generally between 95 and 100%, and decreases to between 70 and 80% in the early afternoon-normally the time of meximum temperature. The seasonal variation is slight.

Moving further up from the region from the coast the relative humidity decreases steadily, and changes abruptly with the season, according to whether the locality is under the influence of the very moist south westerly wind current in the summer or the very dry north to north easterly current at the time of the northern winter. The period of northerly flow is short in the region, which is near the coast, mainly November to February.

Wind Systems: The two major wind current affecting Nigeria is generally applicable to the region. The currents are the northeast and the southwest currents.

The northeasterly wind current is very dry, and gives normally cloudless weather. There is consequently considerable radiation of heat from the ground during the night, and surface temperatures are low in the night and early in the morning.

As earlier mentioned, the southwesterly current is very moist, and when it prevails in sufficient depth it gives cloudy weather, frequently with afternoon and evening thunderstorm, period of moonsoon rain near the coast, and periods of mist in the early mornings. Federal Surveys (1978).

Cloud: The weather is normally cloudy breaking somewhat and lifting in the midday and afternoon. In January and February, however, when the southwesterly current is very shallow, clear skies are not uncommon, while in July and August when the southwesterly current is deep and strong, it gives days of unbroken cloud. Federal Surveys (1978).

3.3.3 Vegetation

The vegetation is especially the high forest, or rather, rain forest - except some part of a belt of bush country further north of the forest. The forest is one

of the most important natural resources of the region, for it provides not only timber and fuel, especially important in a country like Nigeria, but also influences the climatic and water regimes of the region.

Unfortunately, the forest has not been subjected to optimum use in the past, and it is accepted that measures will now have to be put in hand first to stop degeneration and then regenerate the forest. Federal Surveys (1978).

3.3.4 Hydrology and Hydrogeology

The area of study is predominantly the basement complex region of western Nigeria Nigeria, drained by a system of rivers communicating with the sea through a system of lagoons and creeks that characterize the coastal region of West Africa.

Oyegoke et al (1983) found that the basement complex, a geological formation which dominates the north-central parts of Nigeria, extends to these southwestern areas. It covers some 33,000 km^2 in the Ogun River basin and about 11,000 km^2 in the basin area of River Oshun.

Oyegoke et all (1983) also found that the geological formation is generally regarded as comprising poor aquifiers because of its limited storage capacity and the low permeability of much of its laterized surface layers. The area is generally well drained and streamflow responds quickly to rainfall inputs.

Most of the rivers in the region are perennial rivers with low flows occuring in the dry season. A few of intermittent rivers are also present. All the rivers provide well drained agricultural land and the recording and proper documentation of stage and discharge records for a long time where not given the proper attention they deserve. This attitude changed with the pressing need for the development of the region's water resources for water supply, irrigation and to some extent, inland waterways. Attempts at recording storage as well as disemination of the region's hydrological data thus began in the middle to late 1960's, Oyegoke et al (1983).

GKW-Consulting Engineers (1975) reported that River Ogun is the largest river in the western states and it is the best investigated resources of surface water. The River Ogun basin north of Abeokuta is situated in the basement complex which contributes little or nothing to the river's discharge during the dry season.

The Ibaragun gauging station on the Ogun River South of Abeokuta is the oldest station in the region, GKW-Consulting Engineers (1975), having been established in 1962 by the Federal Ministry of Transport, Lokoja. Discharge measurements is carried out on this station during dry season by wading and during flood by boat. It has a cross section of about 150m and at high flood both banks are over-flooded. The western bank is of sandy clay while the eastern bank is of a muddy clay.

The river bottom is of sand with little silt on the eastern side. The discharge varies between $10m^3/s$ and $400m^3/s$. This station enjoys a total catchment of about 21,644 km^2.

GKW-Consulting Engineers (1975) also reported that River Oshun is the second major river in the region. Its source is in Ondo State. It runs westerly into Oyo State to intercept the River Oba which runs southerly from Ogbomoso Fig. 3.2 shows the orientation of the rivers.

Another major station is the Igangan on Ofiki river. The River Ofiki is a tributary to River Oyan which is also tributary to Ogun River.

Generally speaking, in some of the rivers flood carry a considerable load of silt during the rainy season from the catchment area. This silt is transported according to the velocity of the rivers. There are, therefore, changes in the cross-section from period to period due to erosion and sedimentation. The water level in the lower course of some of the rivers suffer from tidal fluctuation especially during the dry season.

As should be expected the drainage is towards the coast, and this is achieved via the system of Lagoons and Creeks. There is a network of minor streams which are tributary to the major rivers.

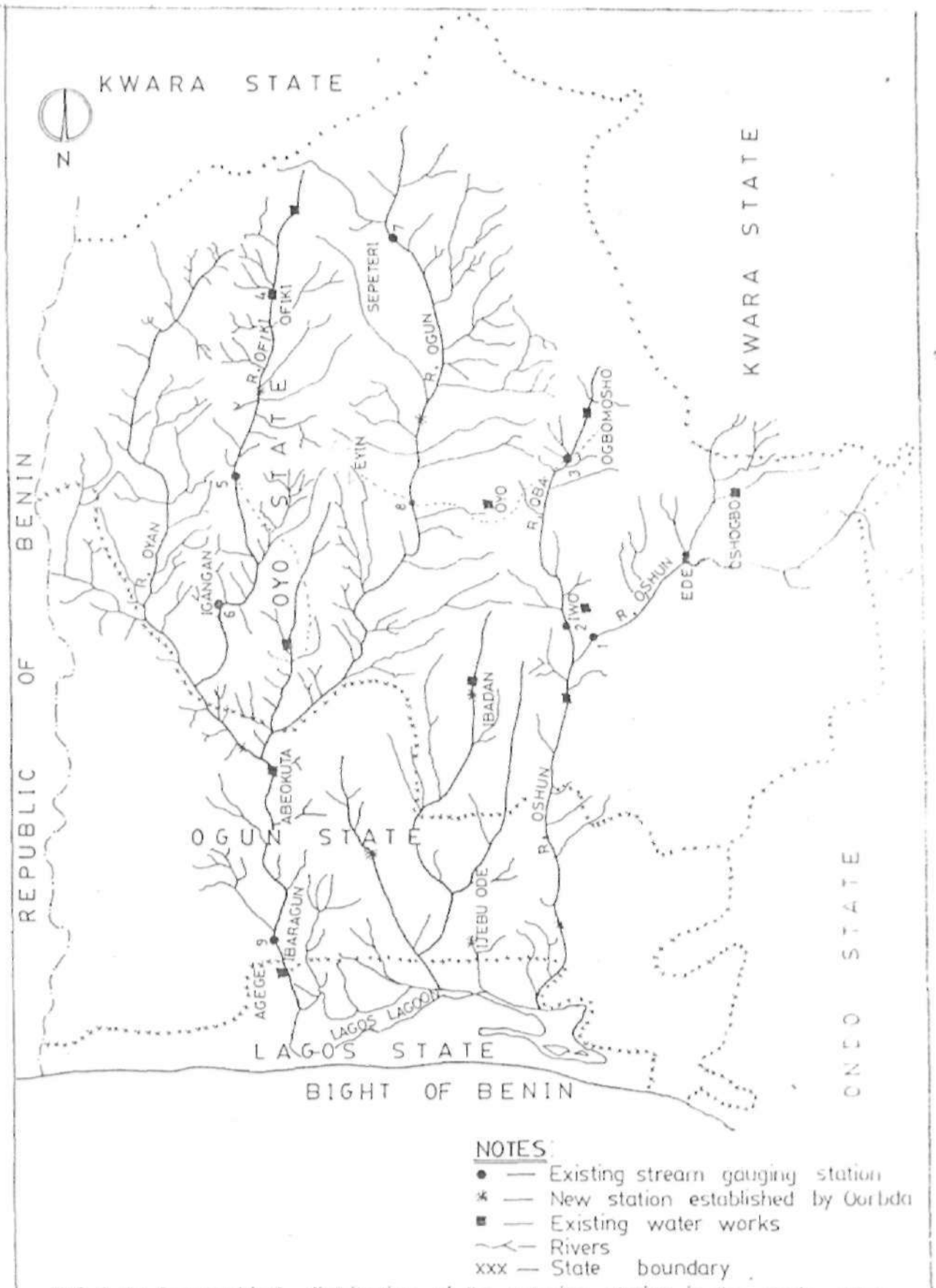


FIG. 3.2: Geographical distribution of the gauging station in the study area

3.3.5. Human Activities

The major activity noticeable in the river basin is farming. Peasant farmers were highly involved in the cultivation of maize, tubers, citrus and kolanut - with few ones engaged in cocoa cultivation - prior to the establishment of Ogun/Oshun River Basins Authority in 1977 with area of jurisdiction encompassing the whole of Lagos, Ogun and Oyo States.

In its annual reports (1980), the Ogun/Oshun River Basins Authority reported that with the establishment of the authority, operations of small sizes of pilot farms were commenced. With the successful operations of these farms larger hectares of land were acquired in subsequent years. Major grains grown by the authority include rain-fed maize and rice with cocoa, citrus and kolanuts as secondary crops. With the advantage of the existence of sprinckler irrigation system on the farmlands, tomatoes, okro and vegetables are grown after maize harvest. A lot of farmers still grow vegetables/tomatoes in the basin.

One other agricultural activity in the basin is fish production. Fish multiplication ponds are constructed where fish is raised and produced.

Forest resources in the region yield highly - rated class of woods that are used in making domestic and industrial furnitures.

3.4 Characteristics of the Data

The data used in the study are the annual flood discharges at nine gauging stations within the region, the drainage areas at these stations and the lengths of main channels.

For the study, the annual flood discharge is defined as the maximum value of the mean daily discharge for each year. The annual flood is considered to be a random variable. The data were collected at the Hydraulics Research Unit, University of Lagos, Oyegoke et al (1983). They were checked and confirmed from the Oyo State Water Corporation, Ibadan and the Ogun-Oshun River Basin Development Authority, Abeokuta, where every other information is got. A sample of nine observation stations has been selected for the study. These stations are the one that have available and reliable data. For each station the following conditions were satisfied:

- the series of values, considered sufficiently reliable, from a period of observation of not less than 15 years.
- the series of values did not present any substantial inhomogeneity due to human activity on the catchment area upstream from the station during the period of observation.
- the values were not significantly influenced by the regulation by reservoirs of considerable capacity.

The climatic situations and hydrological conditions of the selected observation station is fairly the same. Their geographical distribution is illustrated in Fig. 3.2, while Table 3.1 shows the names of the rivers, the gauging stations and their locations, drainage areas of the basin, the lengths of main channels and the length of record at the gauging stations. The stations selected rerer to drainage basins of various sizes. The size distribution of the basins depends mostly on the geomorphological characteristics of the basin.

The average of the annual flood discharges in the region varies from a minimum of $16.80m^3/s$ on River Ofiki at Ofiki town to a maximum of $202.78m^3/s$ on River Oshun at Iwo railway station while the coefficient of variation range from 0.33 on River Oba at Iwo town to 0.59 on River Ofiki at Igangan town. The length of the period of observations varies from 17 years to 24 years with an average of 21 years.

The observed value of each water year at each station is shown on Table 3.2. A bar graph showing graphically the length of record and the missing record is presented as Fig. 3.3.

Table 3.1. Characteristics and length of record of the gauging stations

	River	Gauging Point	Location	Drainage area (km²)	Length of Main Channel (km)	Length of record (years)
1	Oshun	Iwo railway Station	07° 51'N 03° 56'E	4325	123.5	24
2	Oba	Iwo town	07° 38'18"N 04° 11'E	2354	75.5	24
3	Oba	Oyo-Obgbomoso Road	07° 51'N 03° 56'40"E	578	18.0	20
4	Ofiki	Ofiki town	07° 37'45"N 03° 12'20"E	715	38.0	17
5	Ofiki	Iganna-Ilere Road	07° 57'N 03° 14'E	2732	95.0	19
6	Ofiki	Igangan	07° 40'45"N 03° 11'E	3978	145.0	17
7	Ogun	Shepeteri	03° 38'N 03° 39'E	1077	46.0	18
8	Ogun	Oyo-Isehin Road	07° 51'N 03° 56'E	5737	132.5	23
9	Ogun	Ibaragun	06° 46'N 03° 20'E	21644	292.5	24

Table 3.2: Maximum Annual Peak Discharges (m^3/s)

	GAUGE NO								
Year	1	2	3	4	5	6	7	8	9
1965	152.40	61.78	-	-	-	-	-	-	164.10
1966	165.51	43.24	13.76	13.97	62.18	63.49	15.93	142.89	199.94
1967	137.27	81.10	18.46	16.68	42.14	92.54	7.99	69.83	162.06
1968	416.57	115.98	36.93	35.03	89.39	201.56	40.31	305.49	313.31
1969	345.57	101.81	17.10	10.28	54.00	72.09	15.31	140.12	163.11
1970	147.33	72.87	15.38	11.61	65.45	26.68	11.92	75.54	164.10
1971	98.66	42.05	8.01	27.50	37.01	36.05	17.13	75.72	157.85
1972	88.40	43.10	11.24	5.13	5.27	12.83	1.49	82.98	33.67
1973	163.51	34.83	20.53	13.20	92.00	91.74	22.61	114.59	202.57
1974	112.06	44.35	19.51	28.26	84.35	84.48	16.28	112.62	153.11
1975	201.02	54.90	13.71	12.74	33.98	95.92	11.43	66.06	203.36
1976	231.52	58.50	16.90	9.68	20.93	28.91	14.29	72.38	147.32
1977	130.04	27.47	10.47	6.88	55.82	49.68	6.69	66.68	144.69
1978	202.47	60.60	21.82	-	-	-	19.51	123.09	214.67
1979	251.24	74.79	25.01	-	105.98	-	26.82	183.11	244.33
1980	288.31	80.63	30.68	-	138.74	172.01	-	232.06	283.05

Table 3.2 (Cont'd)

Year	GAUGE NO 1	2	3	4	5	6	7	8	9
1981	214.03	67.55	22.50	-	-	100.54	22.47	153.45	228.65
1982	126.28	46.18	-	-	-	-	-	14.57	70.77
1983	178.29	53.05	-	-	41.97	-	-	58.21	128.91
1984	219.61	93.93	38.93	22.93	162.72	187.08	-	246.44	351.28
1985	276.30	82.47	27.43	21.08	128.83	138.31	29.46	220.09	254.10
1986	191.47	54.05	-	10.50	-	-	-	73.35	144.69
1987	250.72	78.87	26.25	21.00	115.90	128.87	28.40	203.87	253.26
1988	227.83	67.40	23.75	19.10	97.01	-	24.21	166.47	228.65
$\bar{X}$	202.78	64.23	20.92	16.80	75.46	93.10	18.48	130.41	193.37
σ	77.92	21.37	8.12	8.00	41.36	55.16	9.17	71.85	71.50
C.V.	0.38	0.33	0.39	0.48	0.55	0.59	0.50	0.55	0.37

Note $\bar{X}$ = arithmetic mean, σ = standard deviation

C.V = coefficient of variation

The blank spaces (-) are missing **records**

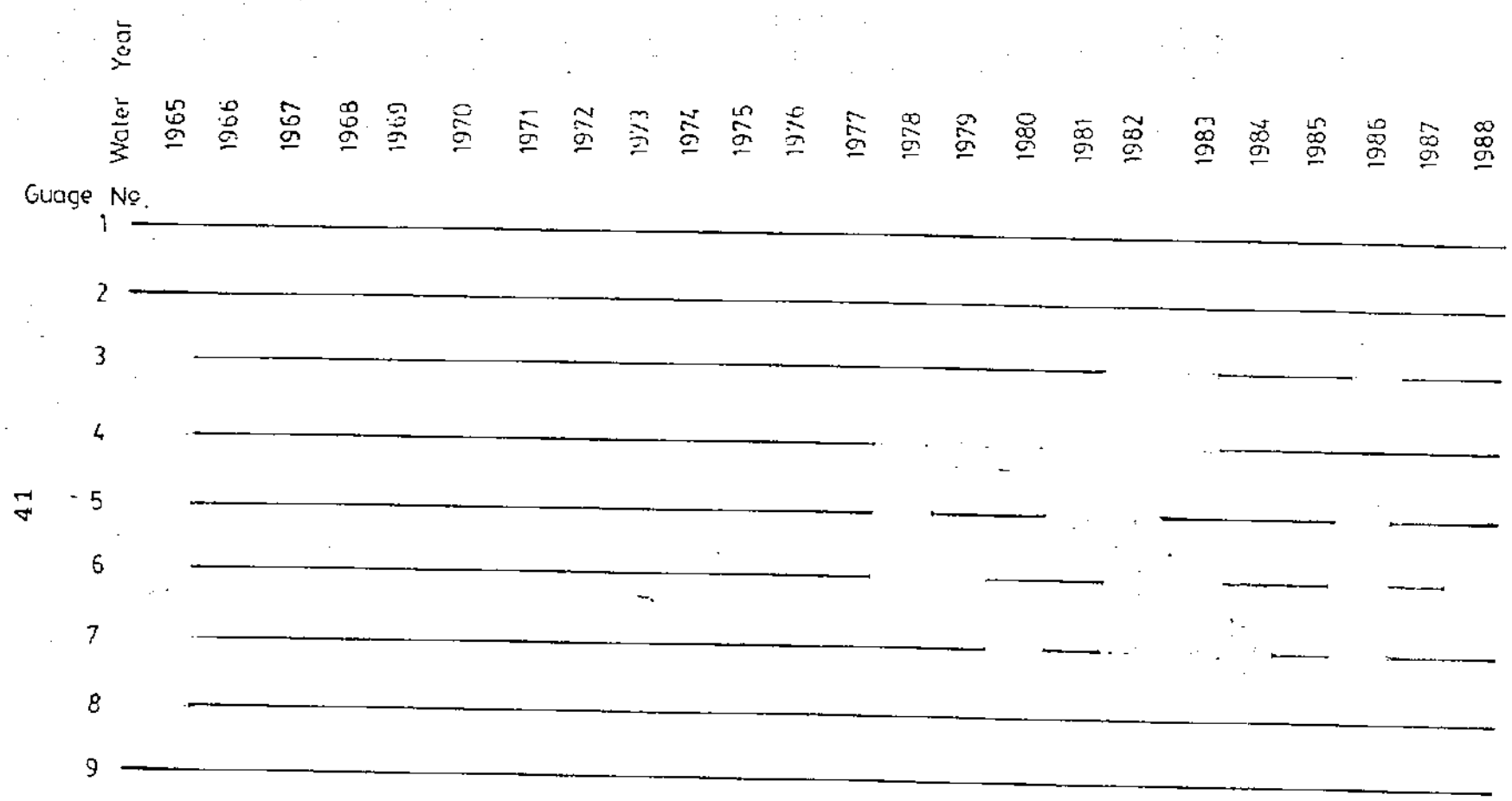


FIG. 3.3: Bar graph showing period of record of maximum annual peak flood at gauging stations.

Chapter 4

ANALYSIS OF DATA

4.1 Selection of Base Period

In order to combine records, comparable data must be analysed. This eliminates the variability with time, so that the effect of other factors on flood peak may be analysed more easily. Yevjevich (1972) propounded that for the purpose of defining the long term relation the period of time used should be as long as possible - a period not less than 20 years is suggested; for the purpose of separating the various factors affecting flood peaks, a common period of time for all records is desirable.

The base-or-common-time period which is chosen is derived from the longest available records. The bar graph of Fig. 3.3 furnishes a compact record of what data are available, and is an aid in selecting a base period. The figure shows that stations numbered 1, 2 and 9 have a a continous record from 1965 to 1988, and so this period has been selected as the base period.

4.2 Adjustment of Records

In deriving a frequency curve for each station the data of the base period must be complete so that the sample can be arranged in order of magnitude with the largest value having an order m = 1 and the smallest

having an order n = number of years of base period = 24 in this analysis. The problem then is what to do concerning the stations that do not have complete record of the base period. The stations are those numbered 3, 4, 5, 6, 7 and 8. The solution is to estimate the discharges of the years of no record.

Dalrymple (1960) explained how the floods of the years of no record can be estimated. The method is to fill missing years by correlation procedures. The procedure is by plotting annual floods of record of a station with missing years of record during the base period against the annual floods at another station where the record is complete. From the curve, the estimates for the years of the missing record is then recorded from those of the years of complete record.

The correlation of annual floods of stations 3 to 8 with station 9 are shown on Figs 4.1 to 4.6, while column 2 of appendix A-3 to A-8 show the estimates of the missing records in parenthesis. Station 9 is chosen for correlation because it has the largest drainage area.

4.3 Preliminary Frequency Curve

Identification of probability distributions of the flood flows is connected to the necessity of characterising in a synthetic form, the data collected and provides a tool for making forecasts.

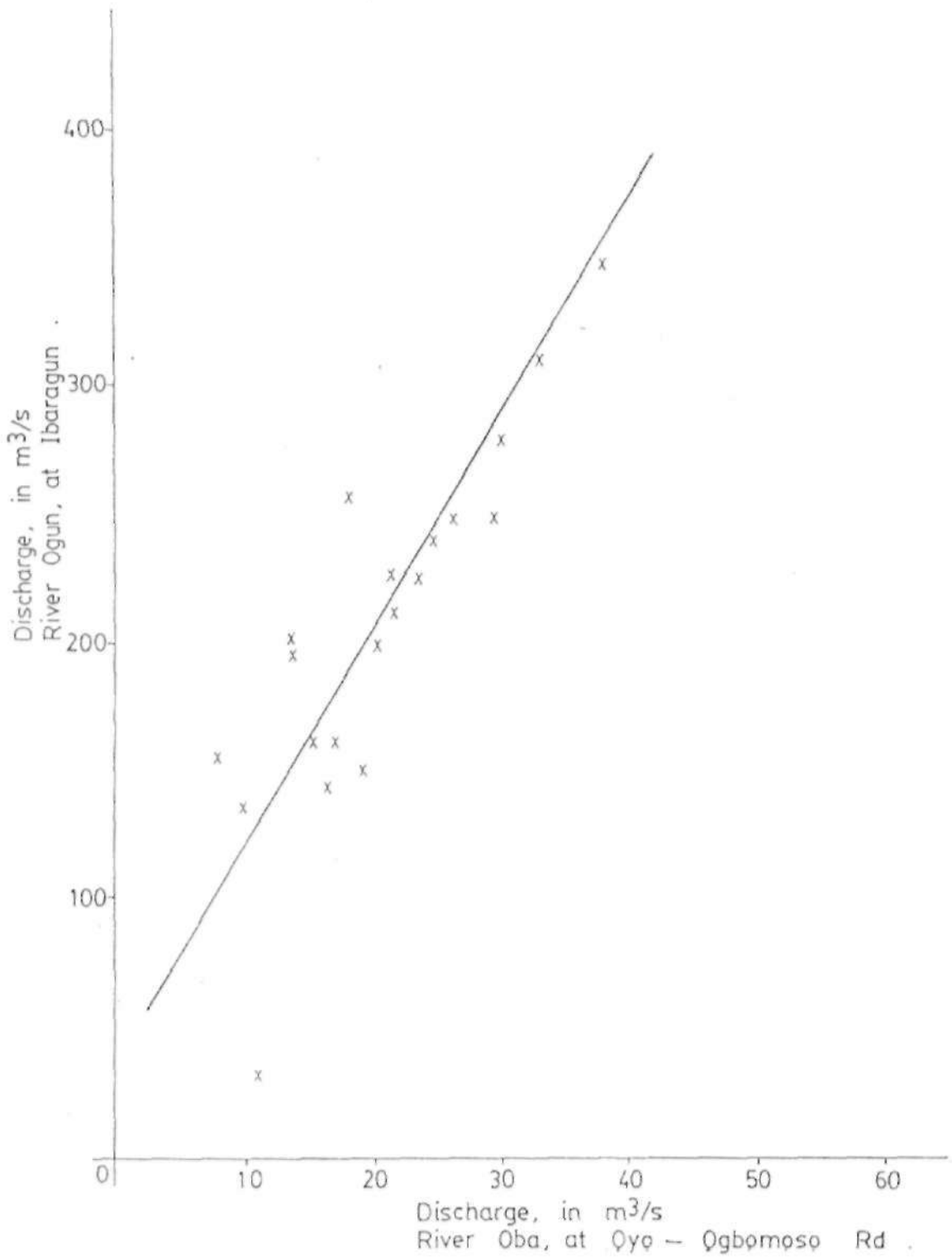


FIG. 4·1 :— Correlation of annual floods, station 3 versus station

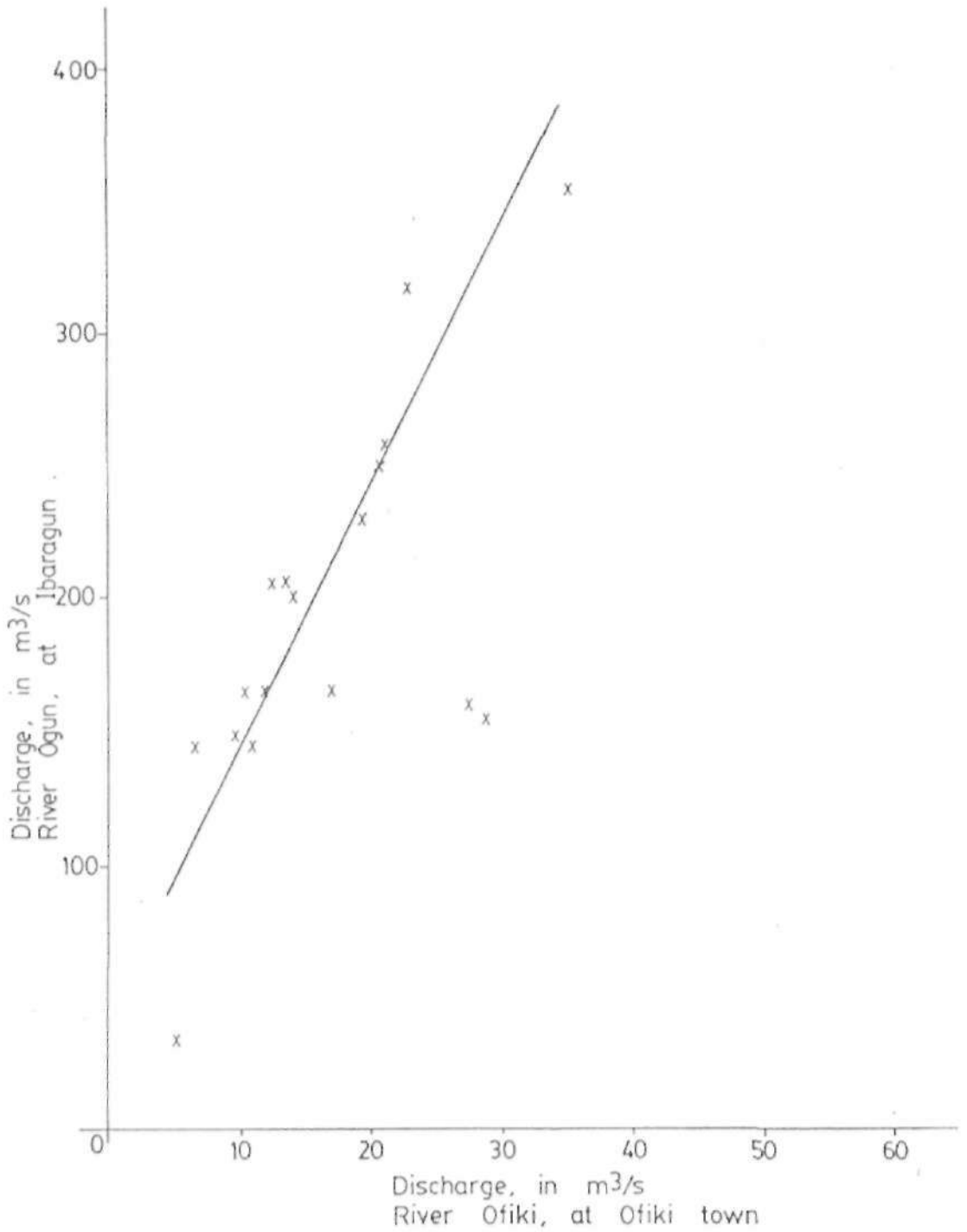


FIG. 4·2 :— Correlation of annual floods, station 4 versus station 9

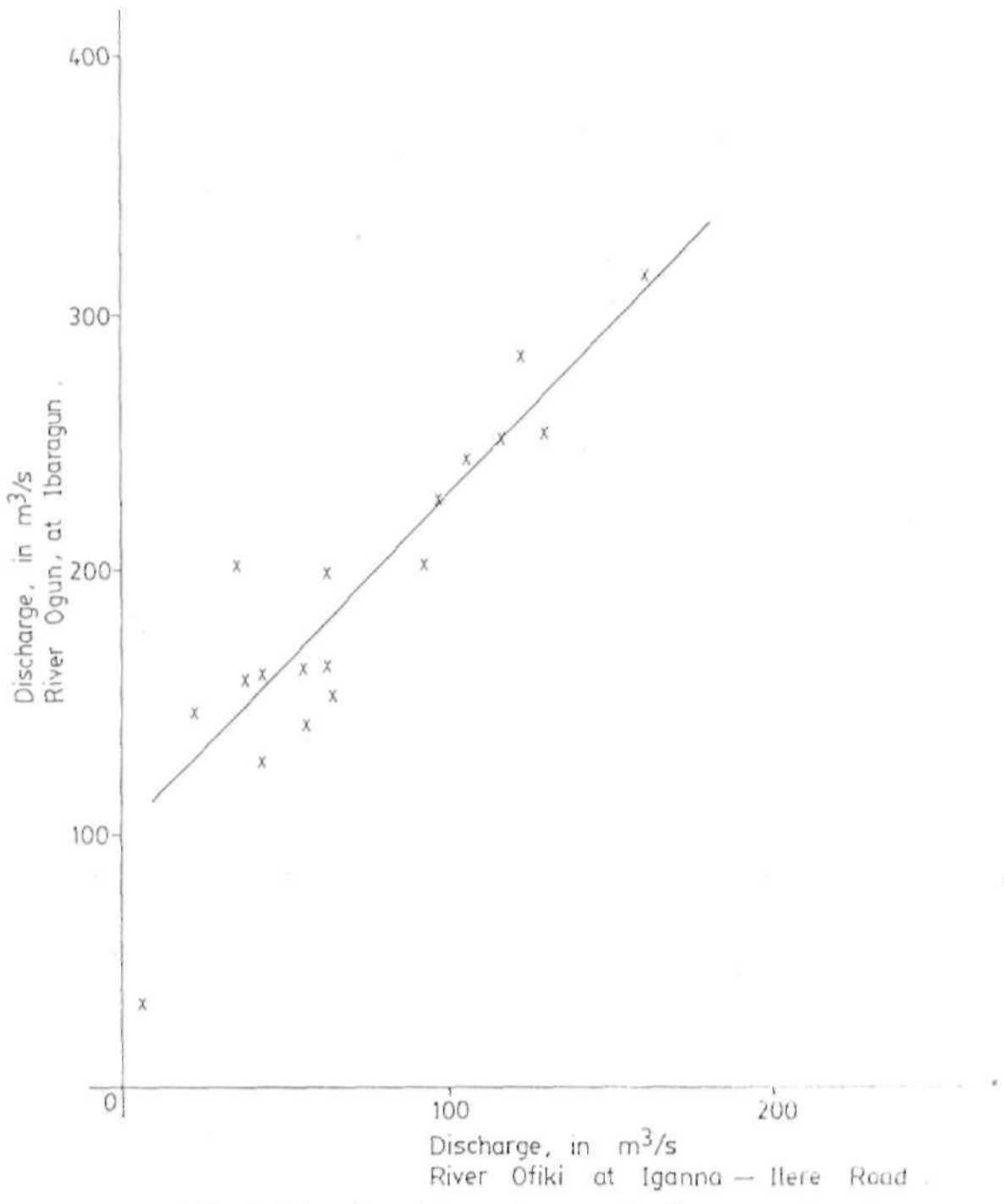


FIG. 4·3 :— Correlation of annual floods, station 5 versus 9

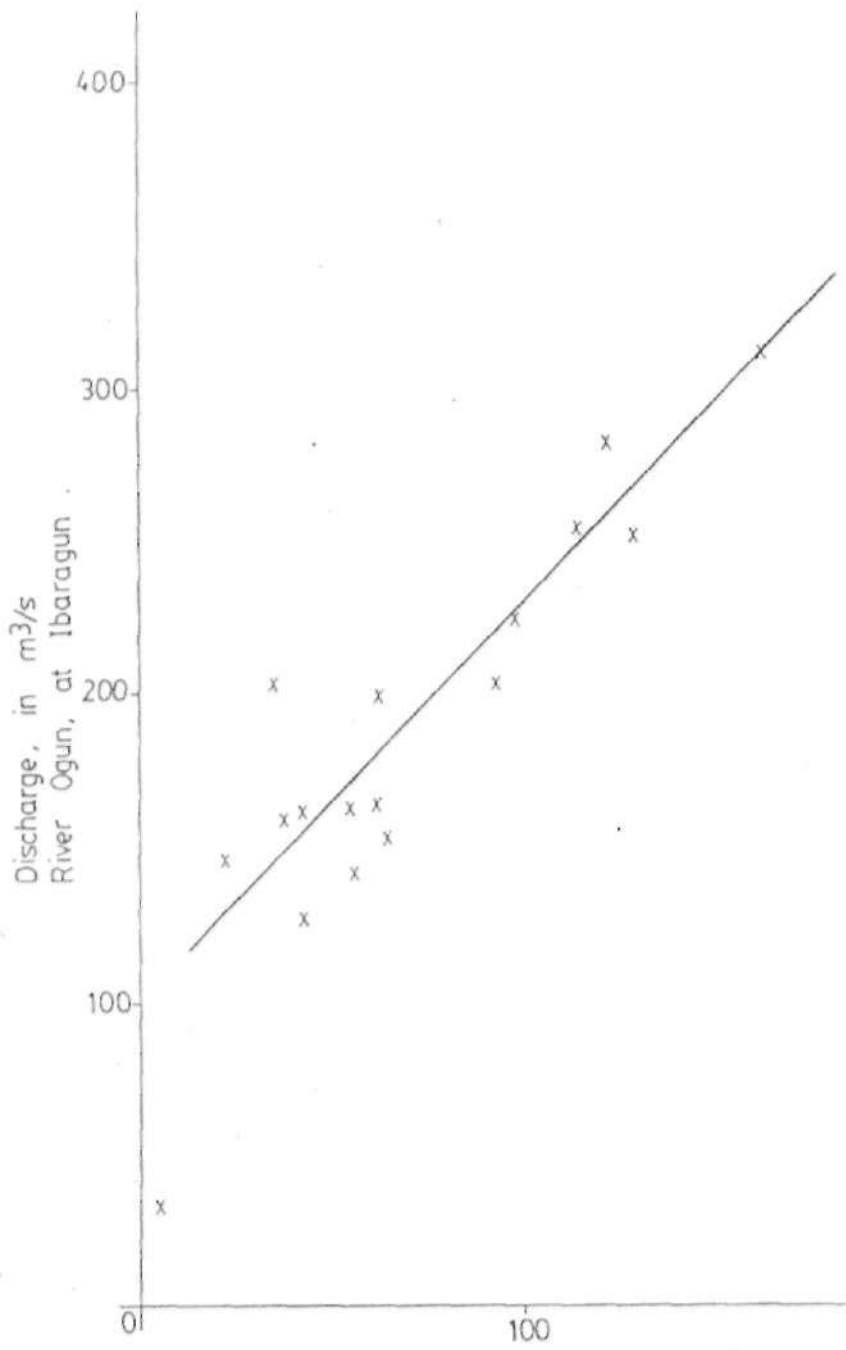


FIG. 4·4 :— Correlation of annual floods, station 6 versus 9

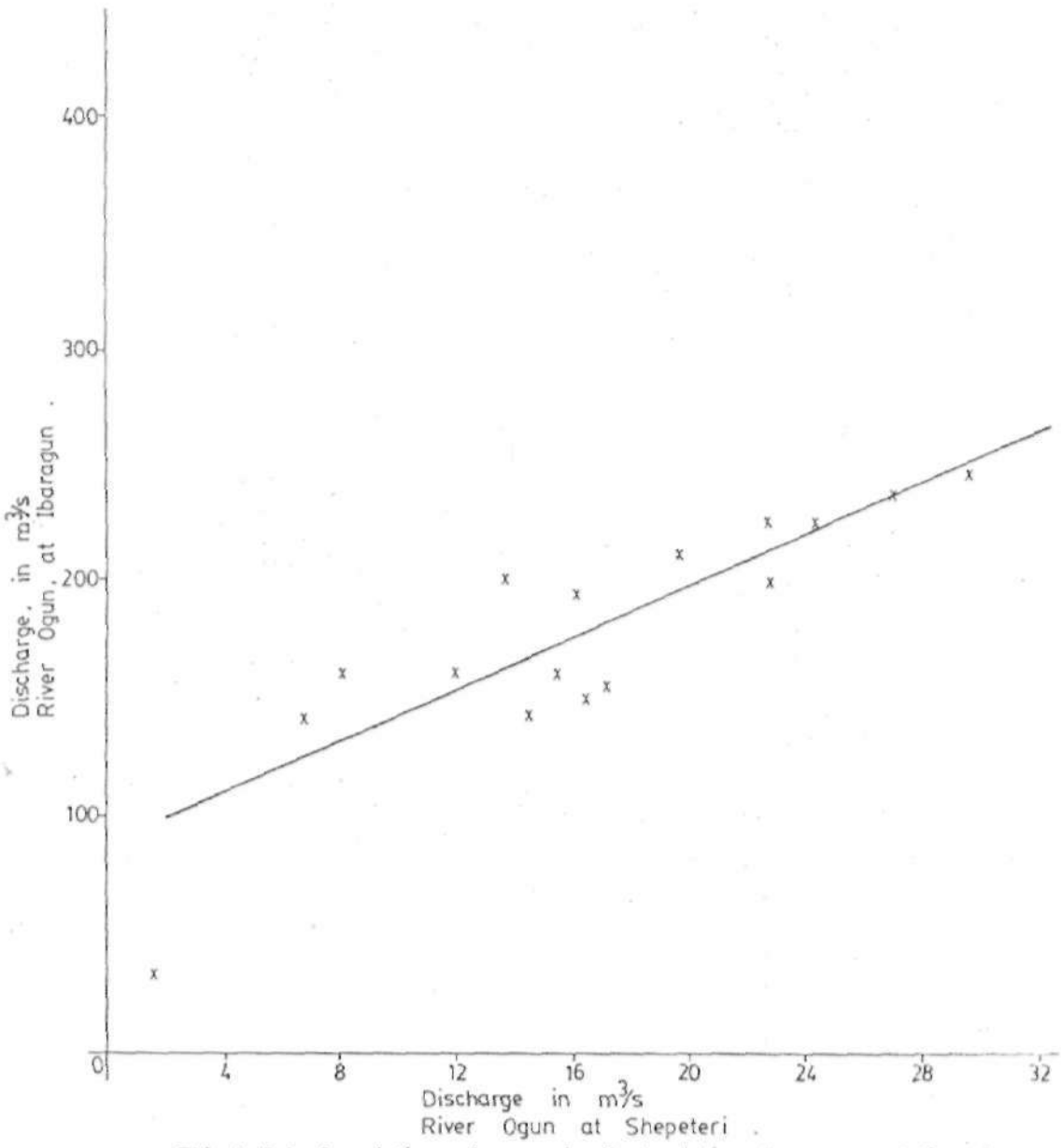


FIG. 4·5 : Correlation of annual floods, station 7 versus station 9 .

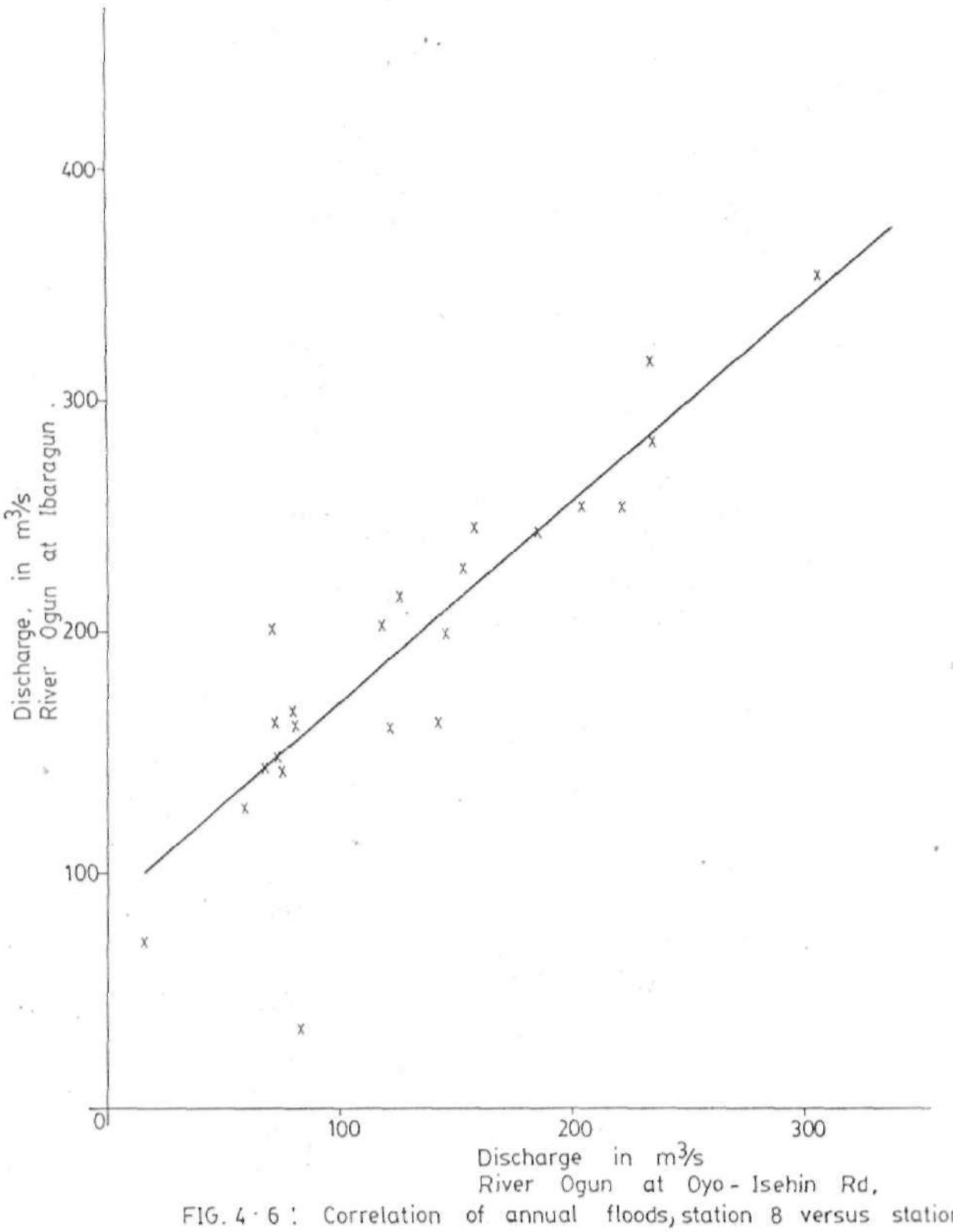


FIG. 4 · 6 : Correlation of annual floods, station 8 versus station 1

The problem of identification of a probability distribution is dealt with using empirical methodologies. This consists of the application of statistical methods on sets of observed data.

In this study, following the statistical approach some theoretical distributions have been adapted to the flood data concerning the nine observation stations. The intention was to single out one distribution for each station to be used for descriptive purposes and in making forecasts needed for the analysis. The statistical methodologies are discussed in the subsequent sections. In brief, three distributions-the log-normal, the Gumbel extreme-value and the log-Pearson type III distributions - are initially assumed for the stream-flow data at each station, then the best fitting distribution out of the three is determined.

4.3.1. Log-Normal Distribution

Log-normal distribution is fitted to the stream-flow data of the nine stations by the following procedures:

(a) The flow magnitudes at each station is ordered with the largest value having an order m = 1 and the smallest having an order n = 24 column 3 of appendix A-1 to A-9 show the flow in order of magnitude.

(b) The probability of a flood being equalled to or exceeded at a station is determined using the formula, Gumbel (1958).

$$P(X) = \frac{m}{n + 1} \qquad \text{---- 4.1}$$

where m and n are as defined above.

(c) The return period, T, of this flood is the inverse of P(X), so it is calculated from

$$T = \frac{n + 1}{m} \qquad \text{---- 4.2}$$

The return period and the exceedence probability of floods at each station as determined from equations 4.2 and 4.1 are shown in columns 4 and 5, respectively, of appendix A-1 to A-9.

(d) The observed discharges are then plotted against the exceedence probability on a log-normal probability paper to obtain the log-normal distribution, that is, the exceedence probabilities on column 5, are plotted against the observed discharges of column 2 of appendix A-1 to A-9.

(e) A frequency curve is drawn by visual aid as the line of best fit Yevjevich (1977).

4.3.2 Gumbel Extreme-value Distribution

The process of fitting the Gumbel Extreme-value distribution, Yevjevich (1977), is the same as that of fitting the log-normal distribution except that stage (d) of the previous section is modified. Here, the observed discharges are plotted against the return period on a Gumbel Extreme-value probability paper to obtain the Gumbel Extreme-value distribution, that is, the return period on column 4 are plotted against the observed discharges of column 2 of appendix A-1 to A-9.

4.3.3 Log-Pearson Type III Distribution

In a similar procedure as stated above, the observed discharges of column 2 of appendix A-1 to A-9 are plotted against the exceedence probabilities on column 5 of appendix A-1 to A-9 but this time on log-Pearson type III probability paper to obtain the log-Pearson type III distribution for the nine stations.

4.4 Selection of the Best-fitted Distribution

For computational expediency, one out of the three distributions assumed for each station must now be selected. Benjamin and Cornell (1970) found that the best-fitting distribution out of the three assumed for the observed data must have the plotted positions closiest to a straight line. The question then is which one of these distribution plots very close to a straight line.

4.4.1 Confidence Band

The first attempt in determing the best fitted distribution was by defining a region around the fitted line of each distribution in which we can say that the true theoretical distribution wholly lies within the upper and lower limits of the region with a certain confidence level. If the fitted line of any of the distribution is to be taken as the true theoretical distribution, then according to Beard (1962), at least, 90% of the plotted positions must fall with the confidence region. It is imperative to have a very tight interval to be able to have a well-defined straight line representing the distribution. In practise, 90, 95 or 99 percent confidence intervals are reported. Increasing the percent confidence intervals requires increasing the interval width, that is, making a less precise statement about the actual position of the fitted line, Benjamin and Cornell (1970).

Beard (1962) presented a method for constructing confidence bands within which, at least, 90% of the plotted positions must fall in order to fit a theoretical distribution. He published coefficients for various levels of confidence, which should be multiplied by the standard deviation of the sample and then added to the discharges of the flood-frequency curve for various recurrence intervals to get the upper-limit of the band.

To get the lower limit of the band, coefficients for various levels of confidence, which should be multipled by the standard deviation of the sample are subtracted from the discharges of the flood-frequency curve for various recurrence intervals.

In this study the procedure taken is as follows:

For a start, the 99% confidence level was applied on the three distributions for each station. In this case, all the plotted positions for the three distributions for each station fall within the confidence band. Consequently a distribution could not be selected out of the three as the one of best fit. The remedy then is to narrow the confidence band which is by reducing the confidence level.

So the confidence level of 95% was subsequently applied on the three distributions for each station. In this case, only one out of the three distributions have all the plotted positions fall within the 95% confidence band for stations numbered 2, 3, 4 and 9. The remaining five station still have the plotted positions of the three distributions falling within the confidence band. The confidence band would still have to be reduced to be able to select a distribution for these other five stations.

Subsequently, a 90% confidence level was applied on the five stations numbered 1, 5, 6, 7 and 8. At this level, one distribution out of the three for each of these

five stations was okay.

At this stage, a distribution has been selected for the nine stations, four of them at 95% confidence level and five at 90% confidence level.

For uniformity in the project it was decided to check whether the four stations having a distribution at 95% confidence level could still satisy Beard's condition at 90% confidence level for the selected distribution of the four stations. Surprisingly, the selected distribution for the four stations are okay at 90% confidence level. Therefore, it was decided to use the 90% confidence level for all the nine stations.

The coefficients of the 90% confidence level, Beard (1962), are reproduced as Table 4.1.

Figure 4.7 to 4.15 show the distribution for each station that satisfy Beard's condition at 90% confidence level. The frequency curves are thus presented.

KASHIM IBRAHIM LIBRARY
AHMADU BELLO UNIVERSITY
ZARIA, N.GERIA

Table 4.1: Coefficients for the calculation of the 90% confidence intervals around the flood frequency curve Beard (1962)

Confidence band	Years of record	Recurrence Interval, Tr, (years) 1,000	100	10	2	1.1	1.01
90%, Upper	5	4.41	3.41	2.12	0.95	0.76	1.00
	10	2.11	1.65	1.07	0.58	0.57	0.76
	15	1.52	1.19	0.79	0.46	0.48	0.65
	20	1.23	0.97	0.64	0.39	0.42	0.58
	30	0.93	0.74	0.50	0.31	0.35	0.49
	40	0.77	0.61	0.42	0.27	0.31	0.43
	50	0.67	0.54	0.36	0.24	0.28	0.39
	70	0.55	0.44	0.30	0.20	0.24	0.34
	100	0.45	0.36	0.25	0.17	0.21	0.29
90%, Lower	5	-1.22	-1.00	-0.76	-1.95	-2.12	-3.41
	10	-0.94	-0.76	-0.57	-0.58	-1.07	-1.65
	15	-0.80	-0.65	-0.48	-0.46	-0.79	-1.19
	20	-0.71	-0.58	-0.42	-0.39	-0.64	-0.97
	30	-0.60	-0.49	-0.35	-0.31	-0.50	-0.74
	40	-0.53	-0.43	-0.31	-0.27	-0.42	-0.61
	50	-0.49	-0.39	-0.28	-0.24	-0.36	-0.54
	70	-0.42	-0.34	-0.24	-0.20	-0.30	-0.44
	100	-0.37	-0.29	-0.21	-0.17	-0.25	-0.36

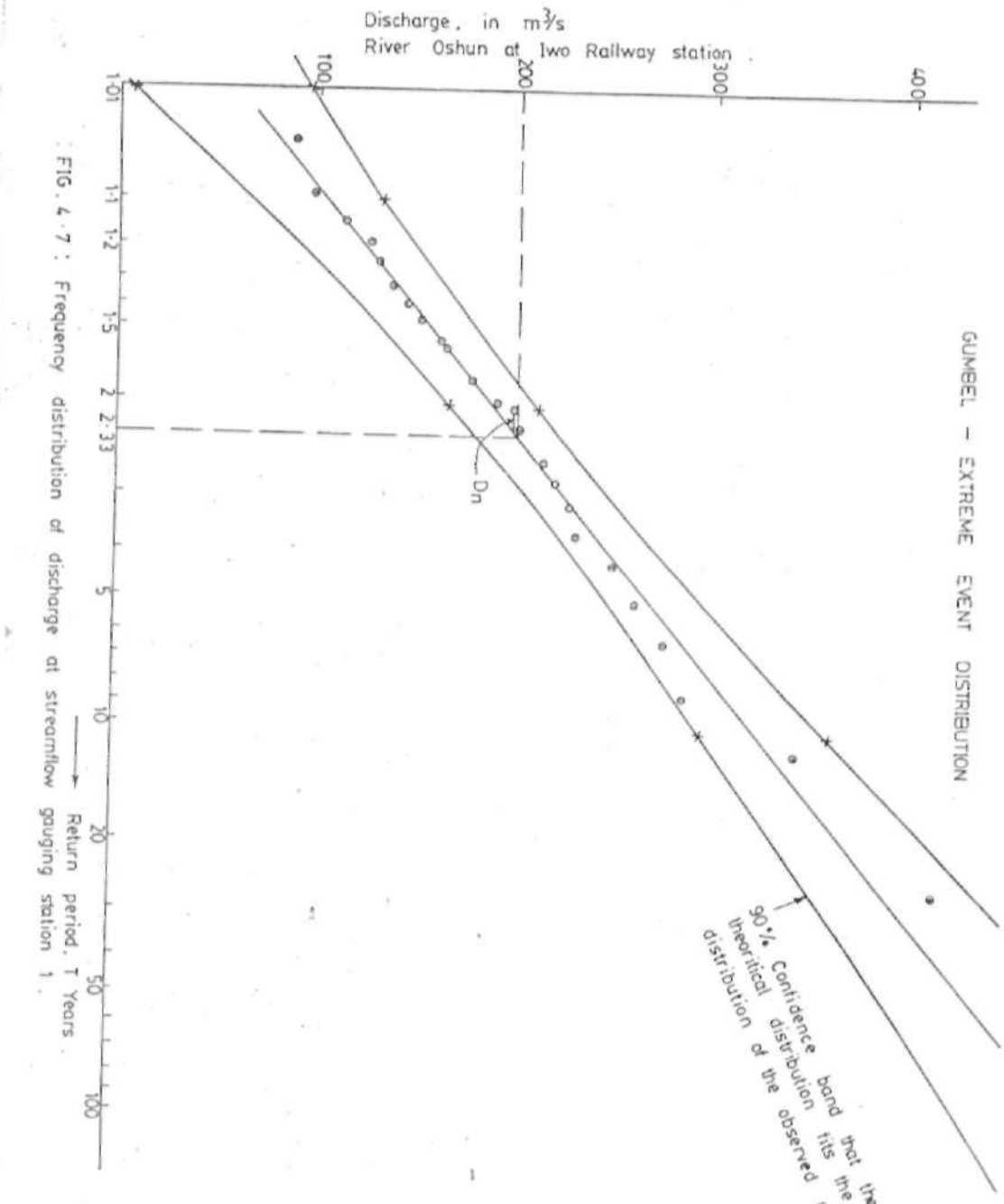
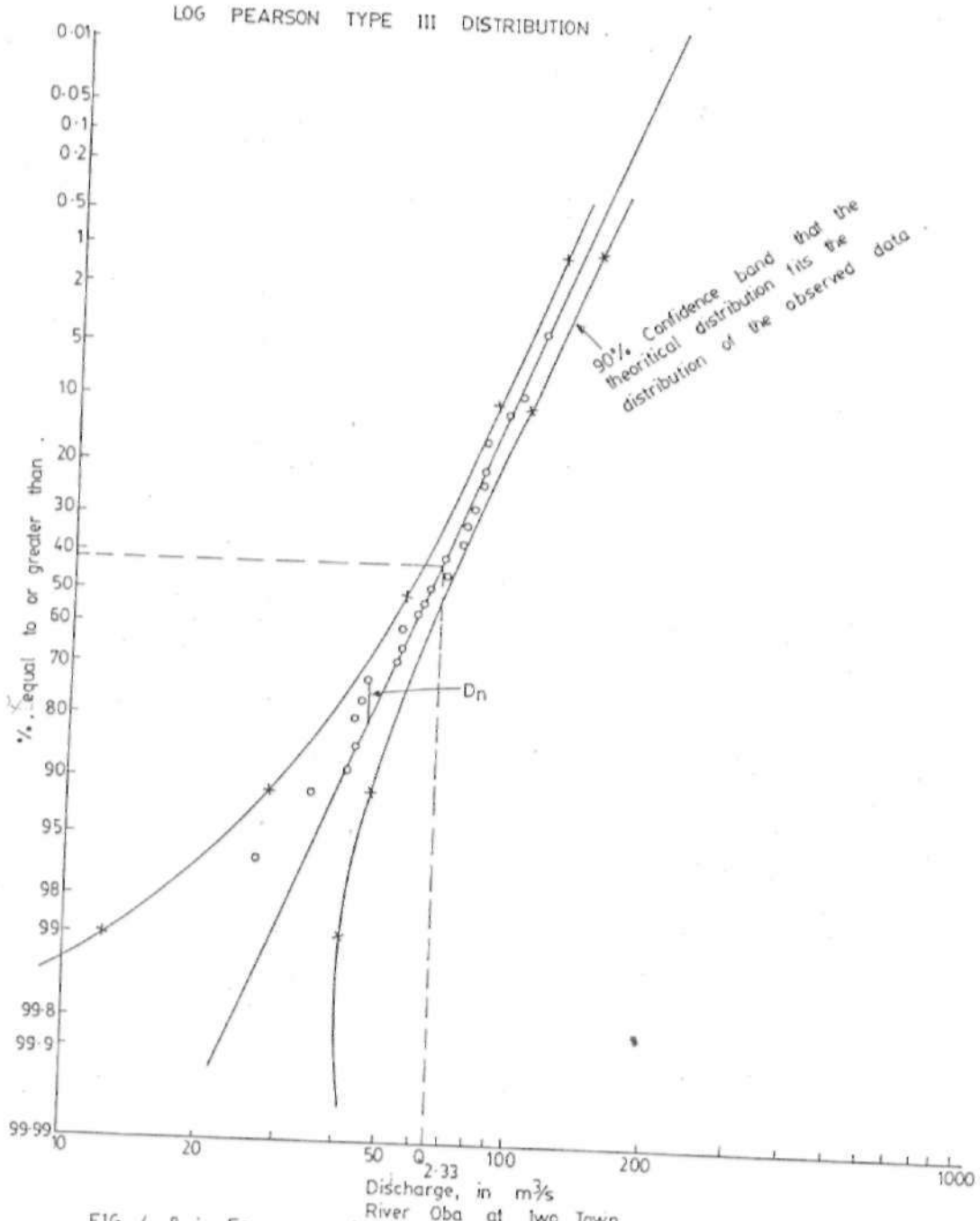
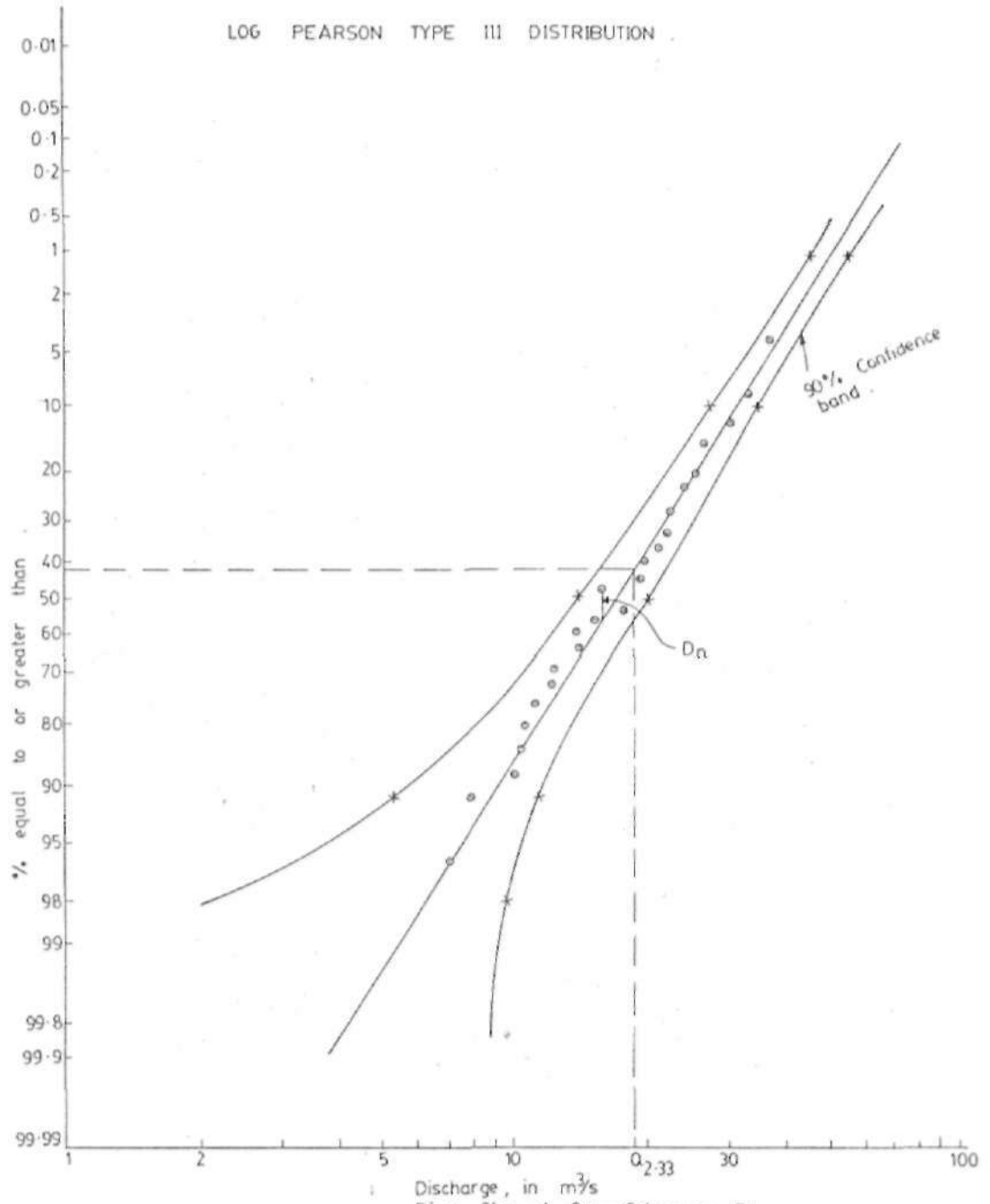


FIG. 4·7 : Frequency distribution of discharge at streamflow gauging station 1.



River Oba at Iwo Town .

FIG. 4.8 : Frequency distribution of discharge at streamflow gauging station 2 .



River Oba at Oyo – Ogbomoso Rd

FIG. 4·9 : Frequency distribution of discharge at streamflow gauging station 3.

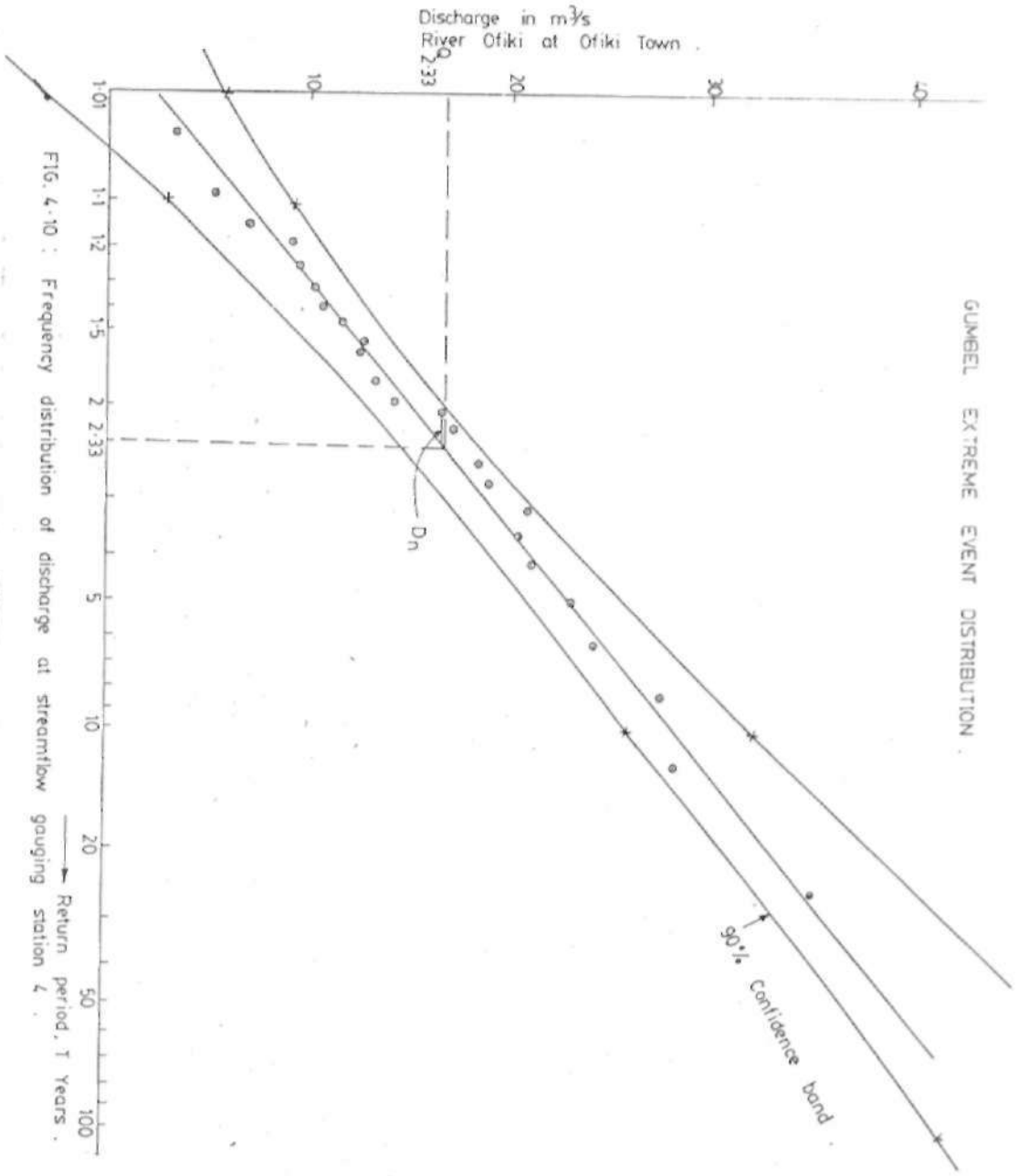


FIG. 4·10 : Frequency distribution of discharge at streamflow gauging station 4 .

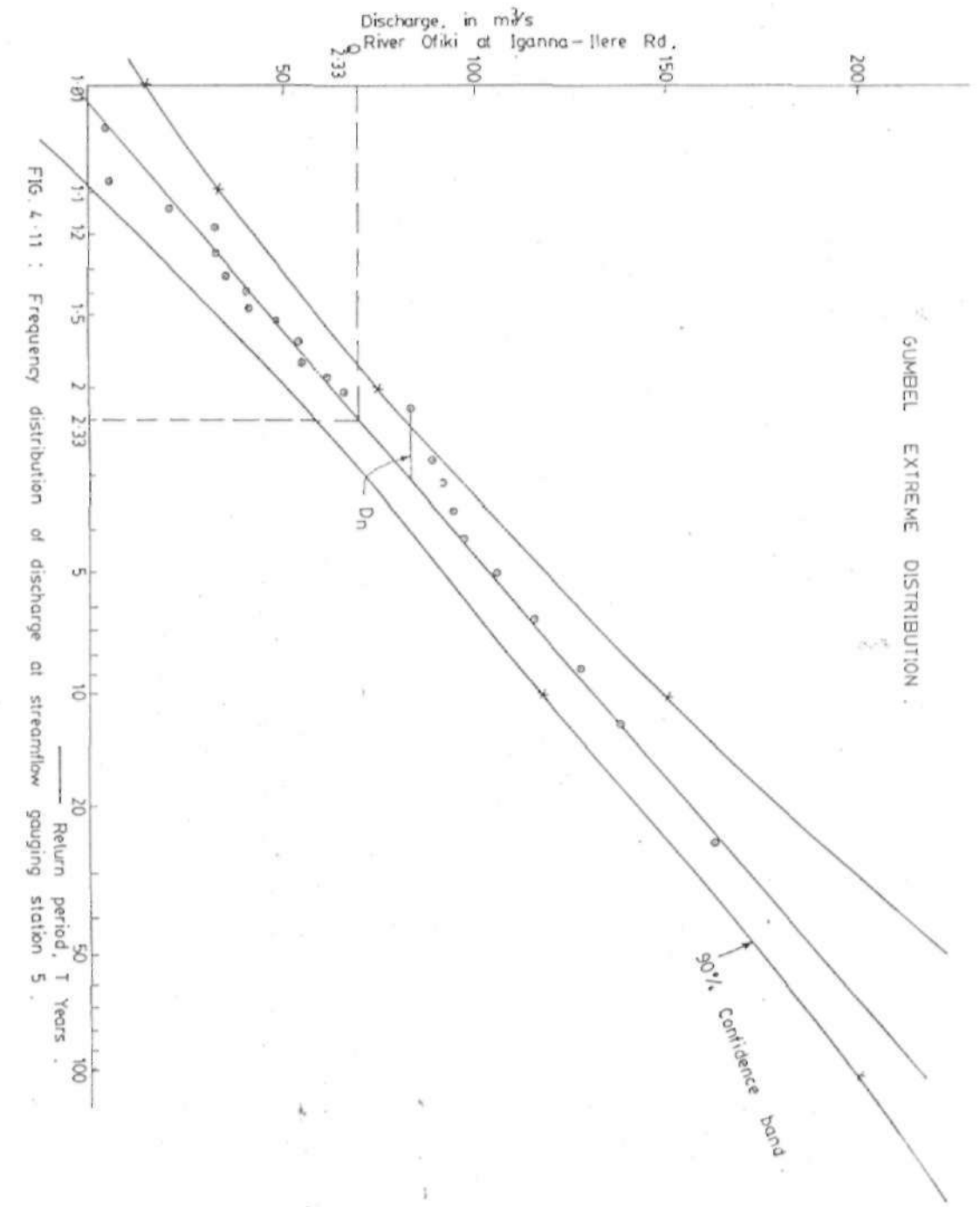


FIG. 4.11 : Frequency distribution of discharge at streamflow gauging station 5.

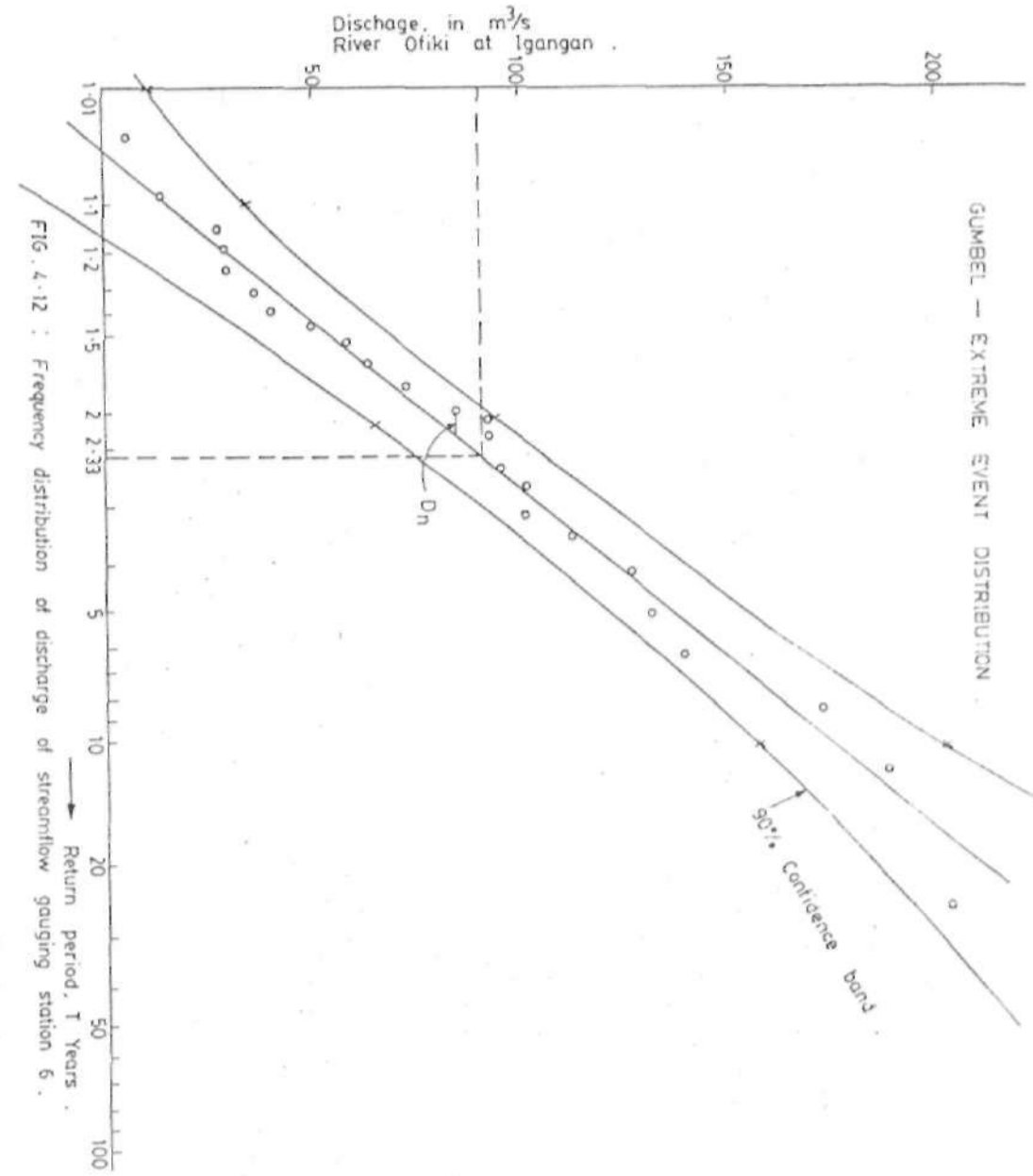


FIG. 4·12 : Frequency distribution of discharge of streamflow gauging station 6 .

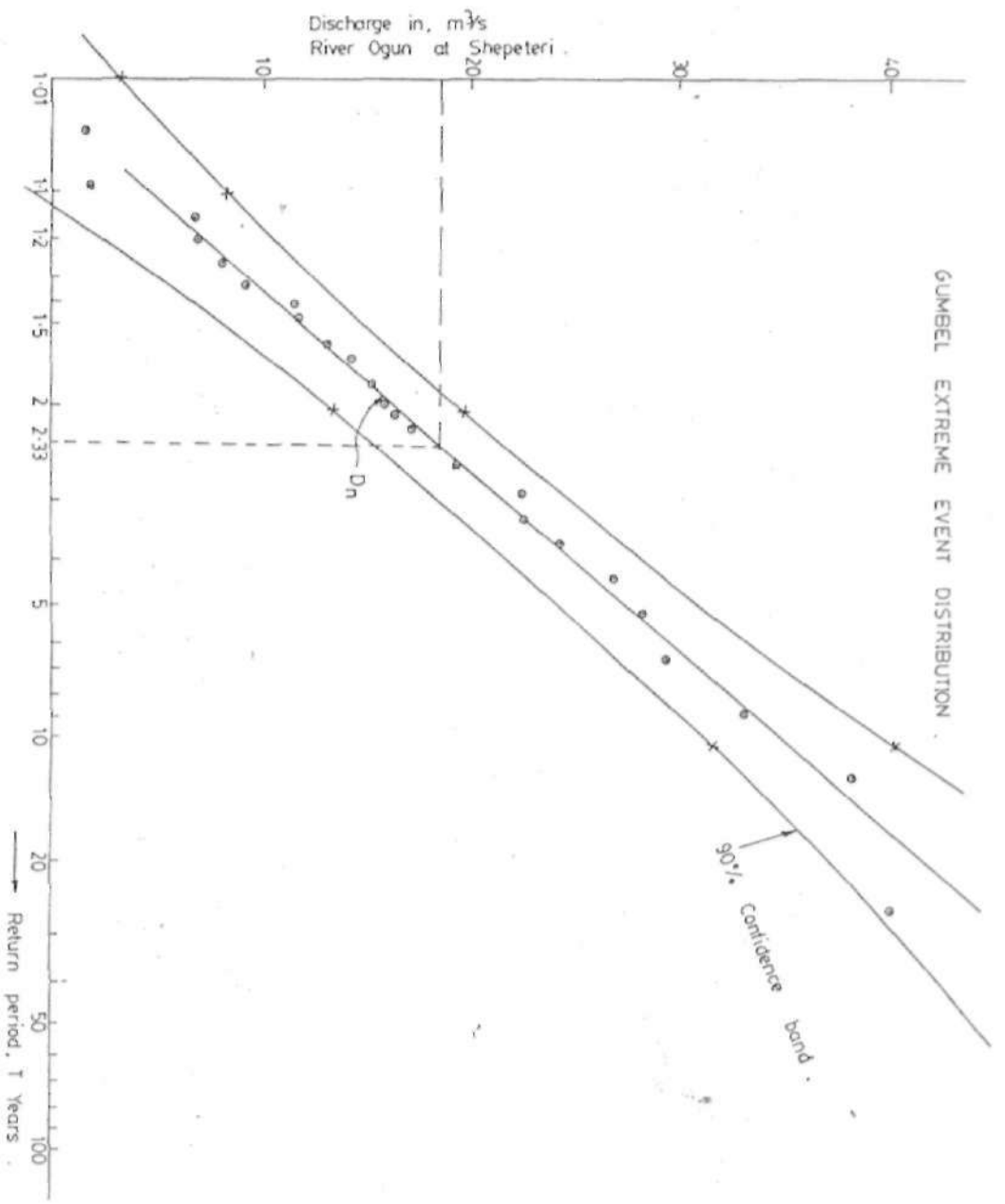


FIG. 4·13 : Frequency distribution of discharge at streamflow gauging station 7 .

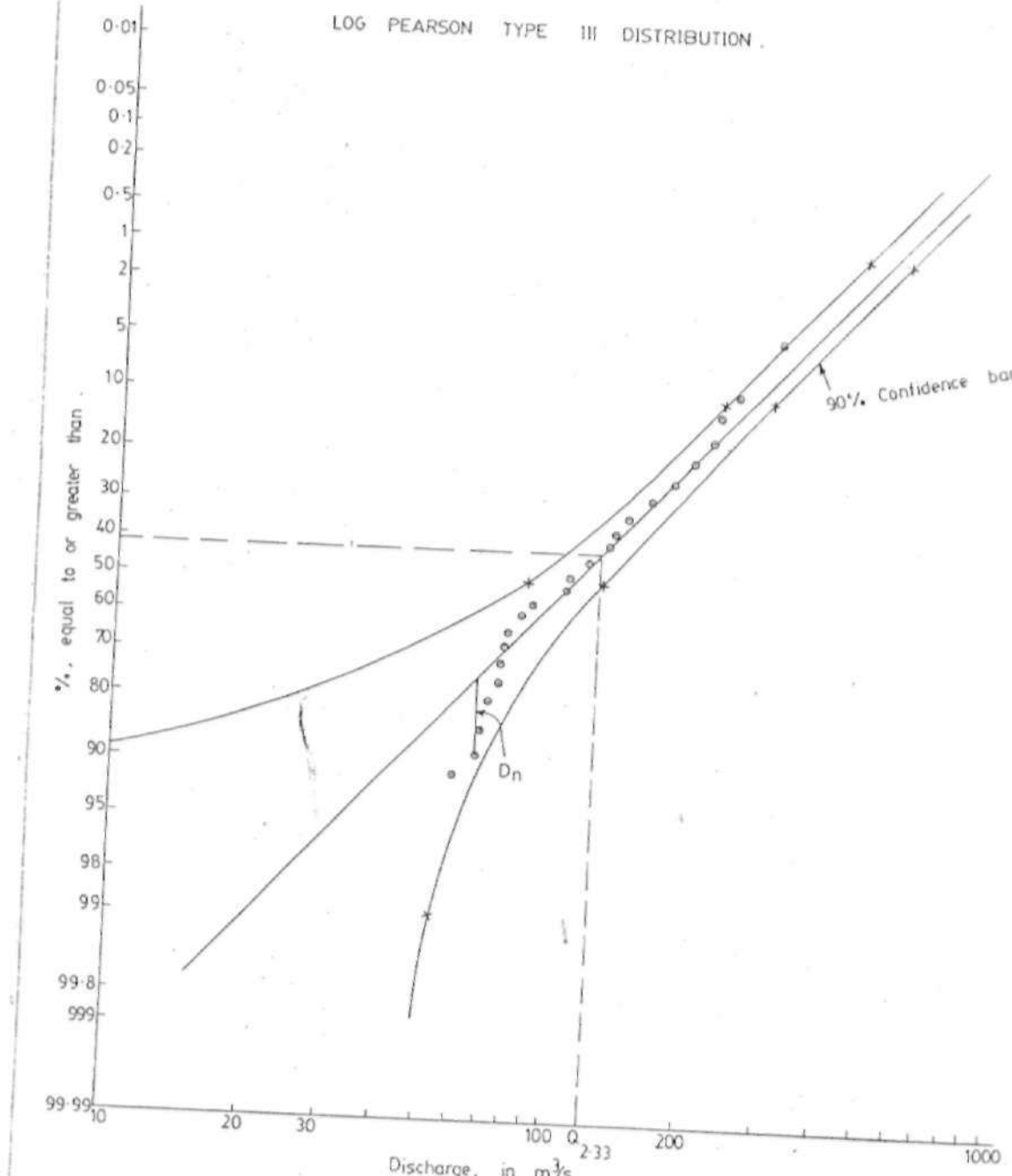


FIG. 4·14 : Frequency distribution of discharge at streamflow gauging station 8.
River Ogun, at Oyo – Isehin Rd.

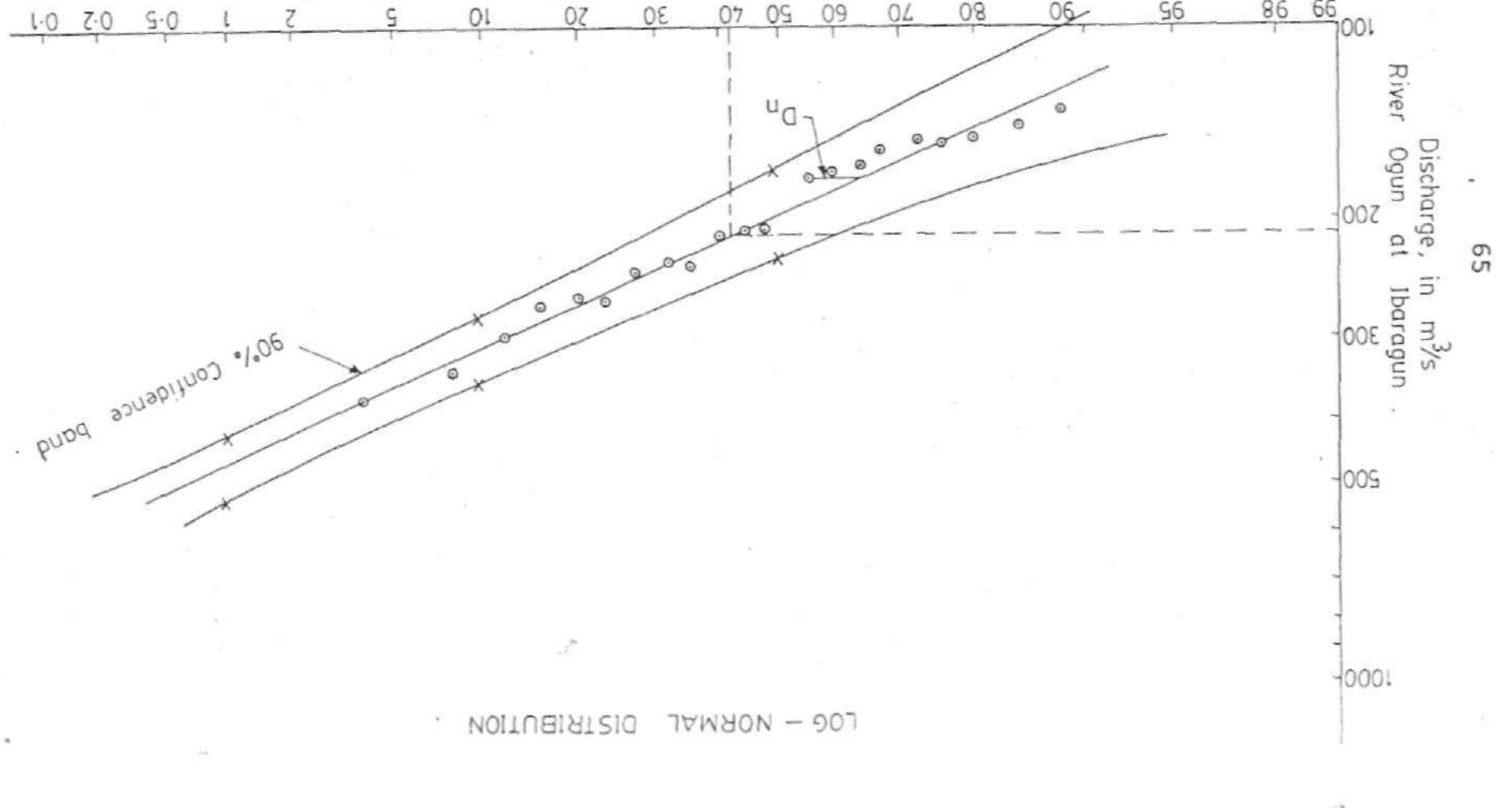


FIG. 4·15: Frequency distribution of discharge of streamflow gauging station 9.

4.4.2 Goodness of Fit Test

Next is to test whether the selected distribution actually fit the sample. Yevjevich (1977) propounded that in a goodness-of-fit type of problem like this the non parametric statistics of the Kolmogorov-Smirnov Test is applicable.

Kolmogorov-Smirnov Test:

$$Dn = \max/Fn(x) - f(x)/ \quad \text{---} \quad 4.3$$

in which Fn (x) is the cummulative function of a sample of n observations arranged in order of size. This is shown on column 5 of appendix A-1 to A-9 for the 9 stations. f(x) is the theoretical cummulative distribution function which is determined from the fitted distribution of Figs. 4.7 to 4.15. The values are shown on column 6 of appendix A-1 to A-9.

The variable Dn is distributed asymptotically as the K-S statistics, Yevjevich (1977).

Assume we have the random sample X_1, X_2, X_n (discharges in this study) and the hypothesistesting situation

$$H_o : Fx(x) = f(x) \text{ for all } X$$

Fx (x) is the population distribution and Ho is termed the null hypothesis.

Since Fn (x) is the statistical image of the population distribution Fx (x), if the null hypothesis is true, the difference between Fn (x) and f(x) should be small for all X except for sampling variation.

For the usual two-sided goodness-of-fit, the alternative hypothesis is:

$$H_1 : Fx(x) \neq f(x) \quad \text{for some } X.$$

Large absolute values of these deviations tend to discredit the null hypothesis. Therefore, the Kolmogorov-smirnov goodness-of-fit test with significance level α is to reject Ho when $Dn > Dn,\alpha$, Yevjevich (1977). The significance level, α, is the probability that the K-S statistics, Dn will lie in the acceptance region, given that Ho holds. Dn,α is the critical value, that is, the start of the rejection region of the K-S statistics at significance level α .

Since a 90% confidence interval had been used previously, the significance level for this study is 1-90%, that is 10%. The critical level of different sample sizes for different significance levels can be found in statistical texts, Benjamin and Cornell (1970), and is reproduced as Table 4.2. For this study, the sample size is 24 and the critical level at 10% significance level is found to be approximately 0.245.

Column 7 of appendix A-1 to A-9 shows $Fn(x) - f(x)$ values for the discharges of the nine stations of study.

The max $Fn(x) - f(x) = Dn$ is circled for each station, and it was found that non of these exceeds the K-S critical critical values $Dn,\alpha = D24, 0.10 = 0.245$. Therefore, the the distribution selected for each station fits well. The max $Fn(x) - f(x)$ are also shown on Figs 4.7 to 4.15.

Table 4.2: Critical Statistics for the Kolmogorov-Smirnov goodness-of-fit test Benjamin and Cornell (1970).

Sample sice	$\alpha = 0.10$	$\alpha = 0.05$	$\alpha = 0.01$
5	0.51	0.56	0.67
10	0.37	0.41	0.49
15	0.30	0.34	0.40
20	0.26	0.29	0.35
25	0.24	0.26	0.32
30	0.22	0.24	0.29
40	0.19	0.21	0.25
Large n	$1.22/\sqrt{n}$	$1.36/\sqrt{n}$	$1.63/\sqrt{n}$

4.5 Mean Annual Food

The mean used for each frequency distribution is the graphical mean determined by the intersection of the visually best fitting frequency line with the line corresponding to the 2.33-year recurrence interval. The mean annual flood is defined as the flood having a recurrence interval of 2.33 years, Dalrymple (1960). The mean annual floods as determined graphically are shown in column 5 of Table 4.3.

Table 4.3: Data for Homogeneity Test. Base Period: 1965 - 1988

No	Stream	Drainage area (km^2)	Length of main channel (km)	Mean annual flood Q2.33 (m^3/s)	10-year flood Q10 (m^3/s)	Ratio $\frac{Q10}{Q2.33}$	Q2.33 X1.70 (m^3/s)	T_E for Q of column 8 (years)	Period of record N_E adjusted (years)
1	2	3	4	5	6	7	8	9	10
1	Oshun River	4325	123.5	200.75	325.11	1.62	341.28	13.5	24
2	Oba River	2354	75.5	67.50	96.50	1.43	114.75	17.0	24
3	Oba River	578	18.0	19.73	32.15	1.63	33.54	12.0	22
4	Ofiki River	715	38.0	16.73	28.56	1.71	28.44	9.5	$20\frac{1}{2}$
5	Ofiki River	2732	95.0	71.44	126.31	1.77	121.45	7.50	$21\frac{1}{2}$
6	Ofiki River	3978	145.0	91.74	176.56	1.92	155.96	7.0	$23\frac{1}{2}$
7	Ogun River	1077	46.0	18.83	35.48	1.88	32.01	7.0	21
8	Ogun River	5737	132.5	138.96	258.59	1.86	236.23	6.7	$23\frac{1}{2}$
9	Ogun River	2144	292.5	209.35	308.75	1.47	355.90	18.0	24
	Average ratio					1.70			

4.6 Homogensity Test

Dalrymple (1960) describe a test which should be used at this stage to check for regional hydrologic homogeneity. If the standard error of estimate of the reduced variable, Y, of the distribution at a station is given by:

$$y = \frac{e^{Y}}{\sqrt{n}} \sqrt{\frac{1}{T-1}} \qquad \text{---- 4.4}$$

then assuming a normal distribution of the estimates, 95% of the estimates will lie within $\pm$ $2\sigma y$ of the most probable value. That is, 95% of the estimates will lie within $y \pm 2\sigma y$. If T the return period of the estimate, is taken as 10 years then from equation 4.4

$$2\sigma y = \frac{0.666\ e^{Y}}{\sqrt{n}} \qquad \text{---- 4.5}$$

Since for T = 10 years the reduced variable is 2.25, Dalrymple (1960), then the confidence limits are given by

$$2.25 \pm 6.33 \Big/ \sqrt{n} \qquad \text{---- 4.6}$$

Table 4.4 after Dalrymple (1960) gives the upper and lower confidence limits with the corresponding return periods for various values of n.

Table 4.4: Confidence Limits of Regional - flood Homogeneity Test Dalrymple (1960)

Sample size n	Lower Limit		Upper Limit	
	$y - 2\sigma$	T_L	$y+2\sigma$	Tu
5	-0.59	1.20	5.09	160
10	0.25	1.85	4.25	70
20	0.83	2.80	3.67	40
50	1.35	4.4	3.15	24
100	1.62	5.6	2.88	18
200	1.80	6.5	2.70	15
500	1.97	7.7	2.53	13
1000	2.05	8.3	2.45	12

The procedure used for the test is to first of all plot T_L and Tu from Table 4.4 versus n on a semi-log scale paper shown as Fig. 4.16. Then, the 10-year flood is estimated at each station from the frequency curves (Fig. 4.7 to 4.15), and presented in column 6 of Table 4.3. Then, for each station in the region to be tested, the ratio of the 10-year event to the mean annual event is computed, this is shown in column 7 of Table 4.3, and an average ratio for the region is multiplied by the mean annual event for each station to give a modified 10-year event magnitude for each station.

This is presented in column 8 of Table 4.3. The return period corresponding to these modified 10-year events are then found for each station from the individual station frequency curves, say T_E. This is presented in column 9 of Table 4.3. The effective period of record of each gauging station is determined as the number of recorded annual events plus one half the number of events computed for that station by inter-correlation (Section 4.2), say N_E. The effective period of record is shown in column 10 of Table 4.3. Next, the co-ordinate pairs (^{T}E, ^{N}E) of columns 9 and 10 of Table 4.3 for each station are plotted on the test graph, Fig. 4.16, showing curves of ^{T}L and ^{T}u. Any station for which the plotted point is outside the confidence limit curves is then exchulded from the homogenous region.

For this study, all the nine co-ordinate pairs (^{T}E, ^{N}E) fall within the limits of the test graph, therefore the records are acceptably homogenous, and records from the nine stations may be grouped together to define a regional flood-frequency curve.

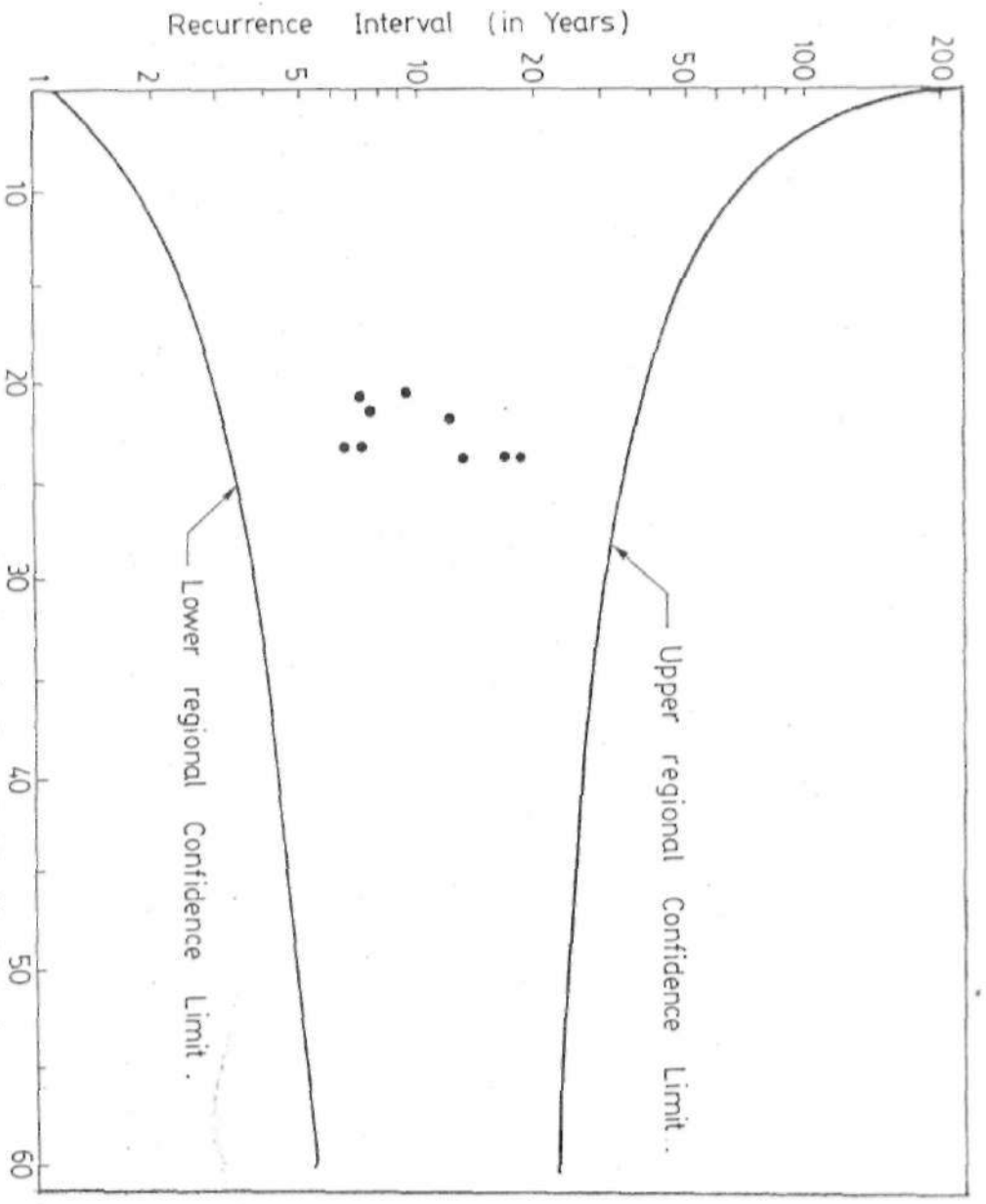


FIG. 4:16 : Homogeneity Test Chart.

4.7 Computation of Median Flood Ratios

The preliminary curves (Figs. 4.7 to 4.15) for all the nine stations in an hydrological homogenous region are now assembled. For each station, ratios of events of different return periods, T, to the mean annual event are computed for T values of say 1.1, 1.5, 5, 10, 20, 50 as shown on Table 4.5.

After tabulating the flood ratios, the median ratios are determined, as recorded in the bottom line of Table 4.5.

4.8 Definition of Regional Frequency Curve

Each median flood ratio is plotted to its corresponding recurrence interval on a frequency chart and an average frequency curve is drawn. The three frequency distribution papers were used but the Gumbel extreme-value distribution was selected because all the points on it plot closiest to a straight line.
Fig. 4.17 (see Appendix B-1 and B-2).

A plot of these median ratios versus, return period is then the regional frequency curve and represents the most likely relationship for all parts of the region.

4.9 Estimation of Mean Annual Flood

In order to apply the regional flood-frequency curve to an ungauged drainage basin the mean flood for that basin will have to be determined.

Table 4.5: Median Flood Ratios

Station	Recurrence Intervals, in years														
	1.1	1.3	1.4	1.5	2	3	4	5	7	10	20	30	40	50	100
1	0.51	0.68	0.75	0.78	0.95	1.12	1.25	1.34	1.48	1.62	1.87	2.00	2.10	2.18	2.43
2	0.59	0.71	0.75	0.79	0.89	1.04	1.16	1.20	1.33	1.41	1.70	1.78	1.85	1.93	2.07
3	0.47	0.63	0.71	0.76	0.89	1.10	1.22	1.29	1.43	1.60	1.90	2.10	2.14	2.28	2.53
4	0.40	0.60	0.65	0.72	0.90	1.14	1.26	1.37	1.53	1.70	2.00	2.17	2.28	2.37	2.69
5	0.26	0.51	0.59	0.66	0.89	1.16	1.33	1.46	1.64	1.85	2.20	2.42	2.56	2.67	3.06
6	0.17	0.44	0.52	0.59	0.84	1.14	1.34	1.48	1.67	1.90	2.30	2.53	2.68	2.82	2.86
7	0.24	0.50	0.57	0.64	0.88	1.17	1.35	1.49	1.67	1.89	2.26	2.48	2.63	2.74	2.97
8	0.29	0.43	0.50	0.59	0.79	1.12	1.30	1.44	1.71	1.91	2.48	2.95	2.98	3.23	4.32
9	0.56	0.74	0.77	0.82	0.94	1.07	1.16	1.25	1.35	1.40	1.53	1.75	1.95	2.09	2.35
Median	0.40	0.60	0.65	0.72	0.89	1.12	1.26	1.37	1.53	1.70	2.00	2.17	2.28	2.37	2.67

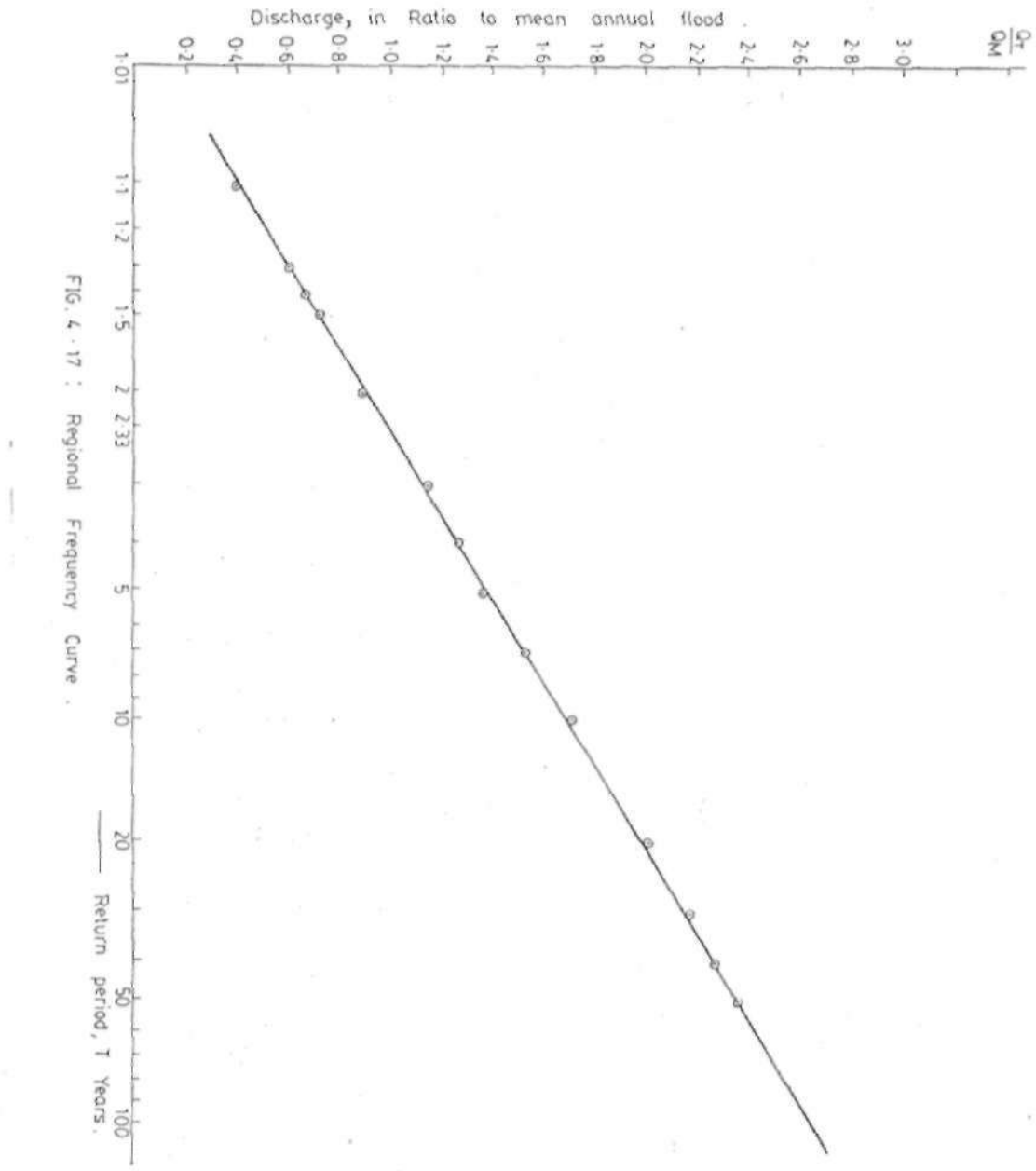


FIG. 4·17 : Regional Frequency Curve.

This involves a correlation analysis of the observed mean floods with drainage basin characteristics, that is, phsiographic factors.

Now that it is assumed that this region is hydrologically homogenous, the factors which should affect the mean flood are size of the basin, slope of the channel, length of main channel and shape of the drainage basin. Of these, size is the most important, and the factors most readily available. Measuring the other factors, except lengths of main channel, was difficult and, in fact, impossible because good topographic maps are not available. The sizes of the drainage basins and length of main channel are again provided in column 3 and 4 of Table 4.3.

The correlation of the mean annual flood with the drainage basin areas and lengths of main channel is achieved either graphically or mathematically.

4.9.1 Graphical Correlation

The graphical correlation of the mean annual flood with both the basin areas and lengths of main channel is achieved by plotting individually, the basin areas and the length of main channel versus the mean annual flood on a log-log scales. That is, the drainage areas in column 3 of Table 4.3 is plotted against the corresponding mean annual event in column 5 of the Table, and a straight line is visually fitted through the points, Fig. 4.18.

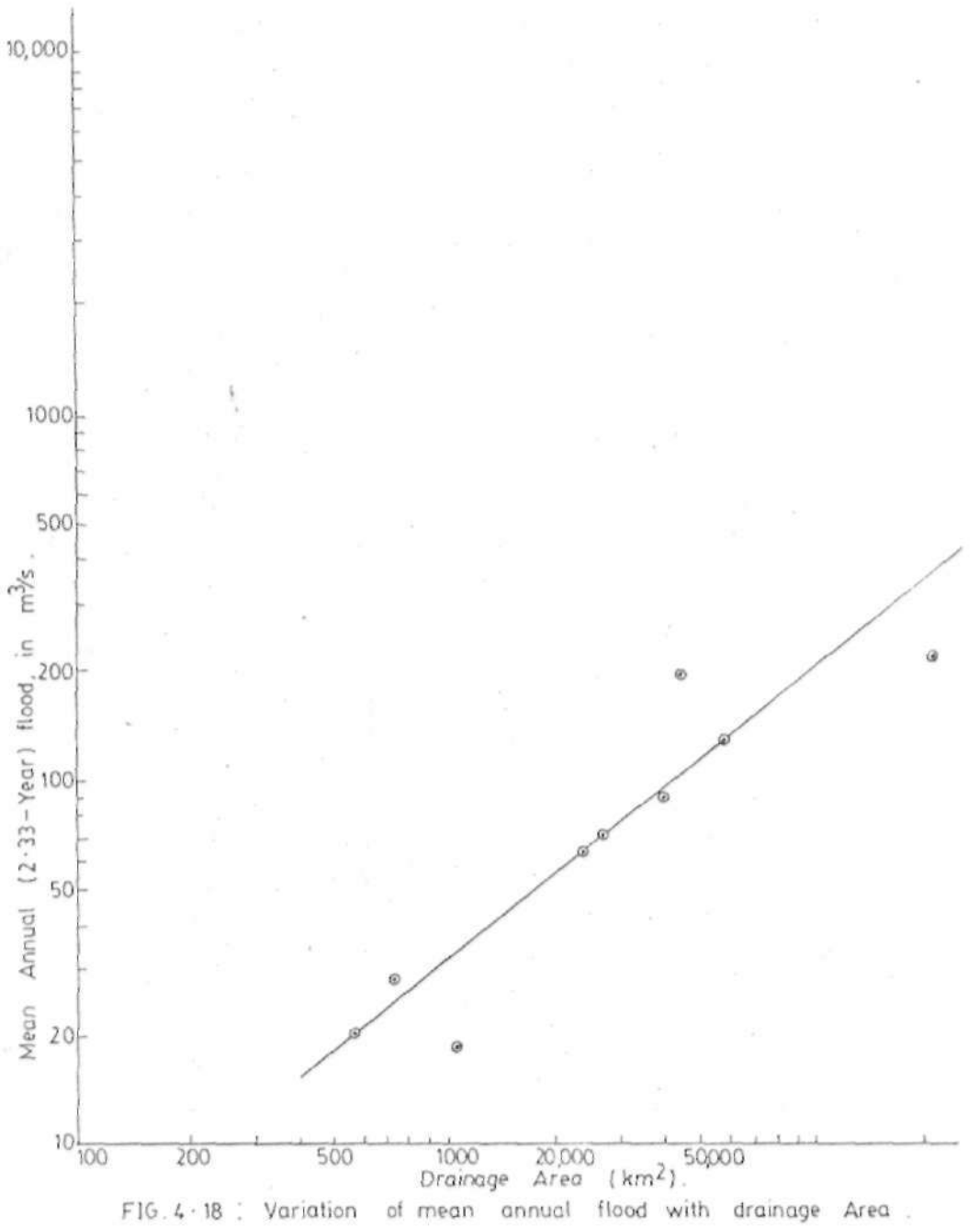


FIG. 4·18 : Variation of mean annual flood with drainage Area .

Similarly, the lengths of main channels in column 4 is plotted against the corresponding mean annual event in column 5, and a straight line is visually fitted through the points, Fig. 4.19.

The mean annual flood discharge for any stream in the region at any site, either gauged or ungauged, can be selected from these curves if the size of its drainage area, the length of its main channel or both are known. If both are known then the mean annual flood can be selected from both curves and an average determined.

Table 4.6 shows the observed mean annual floods for the mine stations used for analysis and the mean annual floods as estimated graphically from curves of Figs. 4.18 and 4.19.

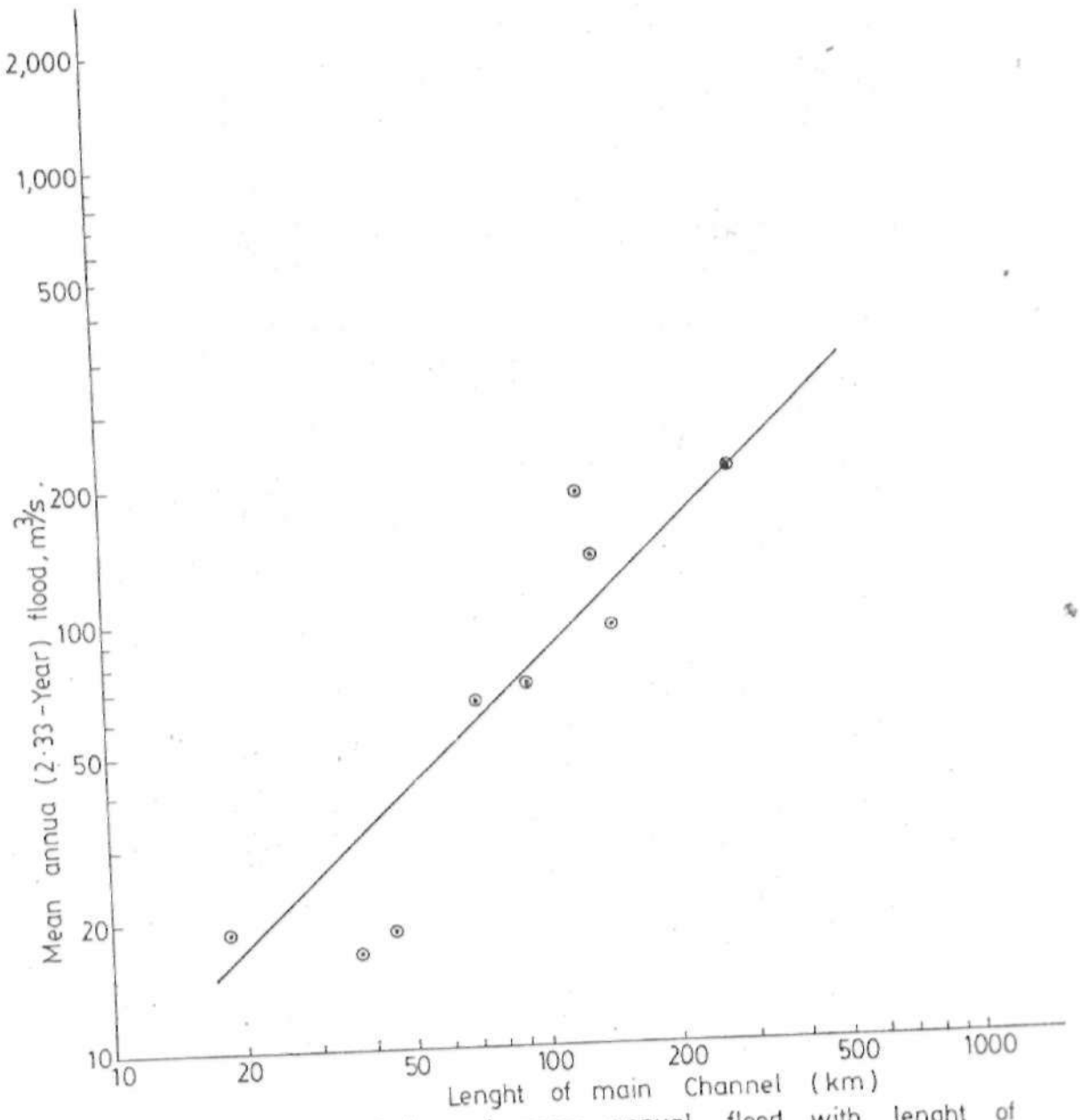


FIG. 4·19 : Variation of mean annual flood with lenght of main Channel.

Table 4.6: Observed and Graphically Estimated Mean Annual Floods

No	Observed mean annual flood (m^3/S)	Estimated mean annual flood from Fig. 4.18 (m^3/S)	Estimated mean annual flood from Fig. 4.19 (m^3/S)	Average of the mean annual flood estimated on columns 3 and 4 (m^3/S)
1	2	3	4	5
1	200.75	112.50	109.95	111.23
2	67.50	63.45	62.44	62.95
3	19.73	19.50	16.85	18.18
4	16.73	24.17	31.23	27.70
5	71.44	70.51	76.35	73.43
6	91.74	93.13	105.00	99.07
7	18.83	30.52	33.25	31.89
8	138.96	125.97	113.68	119.83
9	209.35	324.18	208.77	266.48

4.9.2 Mathematical Relationship Between Flood and Physiographic Factors

Benson (1962) proposed a general correlation formula between the discharge and the physiographic factors of a region as

$$Q = aA^b L^c \qquad \text{---- 4.7}$$

in which Q is the discharge, A and L are drainage area

and length of main channel respectively a, b and C are correlation constants.

Following Benson's proposal, a correlation of the form $Q2.33 = aA^{0.5}L^{b}$ was adopted for the region under analysis (see chapter 5) where Q2.33 is the mean discharge and, A and L are as defined above. Regression analysis was applied in determining the constants a and b.
The procedure is as follows:

(a) The assumed correlation equation is first linearized that is $Q2.33 = aA^{0.5}L^{b}$ becomes

$$\text{in } Q2.33 = \text{Ln } a + 0.5\text{Ln } A + b\text{Ln } L \qquad \text{---- 4.8}$$

(b) Applying the method of least squares, that is minimizing the sum of square of the errors, the following two equations are got from which a and b can be determined.

$$\sum_{i=1}^{9} \text{Ln } Q_{2.33i} = 9\text{Ln } a + 0.5 \sum_{i=1}^{9} \text{LnA}_i + b \sum_{i=1}^{9} \text{LnL}_i \quad \text{---4.9}$$

and

$$\sum_{i=1}^{9} \text{Ln } 0_{2.33i}\text{LnA}_i = n\ a \sum_{i=1}^{9} \text{LnA}_i + 0.5 \sum_{i=1}^{9} (\text{LnAi})^2 + b \sum_{i=1}^{9} \text{LnA}_i\text{LnL}_i \qquad \text{---- 4.10}$$

(c) All the summations in equations 4.9 and 4.10 are then determined from the values of mean discharge, Q2.33, the drainage areas and the lengths of main channel on Table 4.3.

Table 4.7 shows the summation as needed to solve for a and b.

The normal equations are thus:

37.31 = 9 Ln a + 0.5 (70.87) + 39.60 b ---- 4.11

and 302.21 = 70.87 Ln a + 0.5 (568.32) + 319.14 b --4.12

solving equations 4.12 and 4.13 simultaneously gives a = 0.21 and b = 0.4.

Therefore the mathematical relationship between the mean discharge and the two physiographic factors developed for the region is $Q2.33 = 0.21\ A^{0.5}\ L^{0.4}$.

(e) Next is to determine the multiple correlation coefficient. Table 4.8 shows the observed and the estimated values of Q2.33 using the mathematical relationship, and the natural logarithms for the nine stations.

Multiple correlation coefficient

$$= \sqrt{\frac{\sigma_{estimated}}{\sigma_{observed}}} \quad -4.13$$

where $\sigma_{estimated}$ = standard deviation of estimated values of Q2.33 and $\sigma_{observed}$ = standard deviation of observed values of Q2.33, Yevjevich (1972).

Table 4.7: Data for regression Analaysis

No	Q2.33	A	L	Ln Q.2.33	Ln A	Ln L	Ln Q.2.33 Ln A	$(\text{Ln A})^2$	Ln A Ln L
1	200.75	4325	123.5	5.30	8.37	4.82	44.36	70.06	40.34
2	67.50	2354	75.5	4.12	7.76	4.32	32.67	60.22	33.52
3	19.73	578	18.0	2.98	6.36	2.89	18.95	40.45	18.38
4	16.73	715	38.0	2.82	6.57	3.64	18.53	43.16	23.91
5	17.44	2732	95.0	4.27	7.91	4.55	33.78	62.57	35.99
6	91.74	3978	145.0	4.52	8.29	4.98	37.47	68.72	41.28
7	18.83	1077	46.0	2.94	6.98	3.83	20.52	48.72	26.73
8	138.96	5737	132.5	4.93	8.65	4.89	42.64	74.82	42.30
9	209.35	21644	292.5	5.34	9.98	5.68	53.29	99.60	56.69
			Σ	37.31	70.87	39.60	302.21	568.32	319.14
			σ	1.01	1.13	0.84			
			$\bar{X}$	4.15	7.87	4.40			

Table 4.8: Observed and Estimated values of Q2.33

No.	Observed Q2.33	Estimated Q2.33	Ln (observed Q2.33)	Ln (estimated Q2.33)
1	200.75	94.82	5.30	4.55
2	67.50	54.45	4.21	4.05
3	19.73	16.04	2.98	2.78
4	16.73	24.06	2.82	3.18
5	71.44	67.85	4.27	4.22
6	91.74	96.96	4.52	4.57
7	18.83	31.87	2.94	3.46
8	138.96	112.32	4.93	4.72
9	209.36	299.46	5.34	5.70
		σ	1.01	0.89

Since the mathematical relationship was linearized in solving for the parameters a and b then the multiple correlation coefficient is determined from the natural logarithms of observed and estimated values of Q2.33, Benjamin and Cornell (1970). That is, multiple correlation coefficient

$$= \sqrt{\frac{\sigma \text{ Ln (estimated value of Q2.33}}{\sigma \text{ Ln (observed value of Q2.33)}}} \quad \text{---- 4.14}$$

$$= \sqrt{\frac{0.89}{1.01}} = 0.93$$

(f) Next is the significance test.

(i) To test whether there is significance dependence of Ln Q2.33 on Ln L.

H_o : b = 0

H_1 : b ≠ 0

Test statistics : $\hat{b}$, the estimator of b .

Rule form : we accept the null hypothesis, H_o, if $- t_{\alpha/2,n-3}\sigma_{\hat{b}} \leq b \leq t_{\alpha/2,n-3}\sigma_{\hat{b}}$

where $t_{\alpha/2, n-3}$ is the t statistics at significance level α and n-3 degrees of freedom, Benjamin and Cornell (1970).

n = no of pairs of data = 9.

Benjamin and Cornell (1970) gives

$$\sigma_{\hat{b}} = \frac{1}{n-3} \frac{(1-\sigma^2 LnL,LnQ).\sigma^2 LnA}{\sigma^2 LnQ\ \sigma^2 LnL\ \sigma^2 LnL} \quad \text{---- 4.15}$$

where

$$\sigma_{LnL,LnQ} = \frac{1}{n} \sum_{i=1}^{9} (LnLi-(\overline{LnLi}))(LnQi-(\overline{LnQi})) \text{ -- 4.16}$$

From Table 4.7.

$\sigma^2 LnQ = 1.02$, $\sigma^2 LnL = 0.71$ and $\sigma^2 LnA = 1.28$

also $(\overline{LnQ}) = 4.15$, $(\overline{LnLi}) = 4.40$ and $(\overline{LnAi}) = 7.87$

Table 4.9 provides the values for determing $\sigma_{Ln1,LnQ}$

Table 4.9: Table for Significance Test

No	Ln Q-($\overline{LnQ}$)	Ln A-($\overline{LnA}$)	Ln L-($\overline{LnL}$)	Column 2 x Column 3	Column 2 x Column 4
1	2	3	4	5	6
1	1.15	0.50	0.42	0.58	0.48
2	0.06	-0.11	-0.08	-0.01	-0.01
3	-1.17	-1.15	-1.51	1.77	1.77
4	-1.33	-1.30	-0.76	1.73	1.01
5	0.12	0.04	0.15	0.01	0.02
6	0.37	0.42	0.58	0.16	0.21
7	-1.21	-0.89	-0.57	1.08	0.69
8	0.78	0.78	0.49	0.61	0.38
9	1.19	2.11	1.28	2.51	1.52
			$\sum$	8.44	6.08

From Table 4.9: (Ln Li - ($\overline{Ln\ Li}$) (Ln Qi-($\overline{Ln\ Qi}$) = 6.08

Therefore, from equation (4.16):

$$\sigma_{LnL,LnQ} = \frac{6.08}{9} = 0.68$$

and from equation (4.15)

$$\hat{\sigma_b} = \frac{1}{6}\left[1 - \frac{(0.68)^2}{(1.02)(0.71)}\right] \times \frac{(1.02)}{(0.71)} = 0.09$$

Assuming a 5% significance level, $t_{0.025,6} = 2.447$.

The acceptance level for the null hypothesis is

$0 + t_{0.025,6}\ _b = 0 \pm 2.447 \times 0.09 = 0 \pm 0.22$.

Since b falls outside this interval we reject the null hypothesis that the length of main channel does not have a relation to the mean discharge.

(ii) To test whether there is significance dependence of Ln Q2.33 on Ln_A.

H_o: 0.5 is insignificantly different from 0

H_1: 0.5 is significantly different from 0.

Test statistics: The power of A in the developed mathematical relationship.

Rule form: we accept the null hypothesis,

H_o, if $-t_{\alpha/2,\ n-3} \leq \sigma_{\text{estimated power of A}} \leq 0.5 \leq t_{\alpha/2,\ n-3}$.

$\sigma_{\text{estimated power of A}}$

where $t_{\alpha/2,\ n-3}$ is as defined above.

from Benjamin and Cornell (1970)

$$\sigma_{\text{estimated power of A}} = \frac{1}{n-3}\left[\left(1 - \frac{\sigma^2_{LnA,LnQ}}{\sigma^2_{LnQ},\ \sigma^2_{LnA}}\right)\frac{\sigma^2_{LnQ}}{\sigma^2_{LnA}}\right] \quad --4.17$$

where $\sigma_{LnA,LnQ} = \frac{1}{n}\sum_{i=1}^{9}(LnAi-(\overline{LnAi}))(LnQi-(\overline{LnQi}))$ --4.18

from Table 4.7

σ^2 LnQ = 1.02, σ^2 LnA = 1.28

and from Table 4.9 (LnAi - ($\overline{\text{LnA}}$i)) (LnQi-($\overline{\text{LnQ}}$i)) = 8.44

Therefore, from equation (4.18):

$$\sigma_{\text{Ln A, Ln Q}} = \frac{8.44}{9} = 0.94$$

and from equation (4.17)

$$\sigma_{\substack{\text{estimated} \\ \text{power of A}}} = \frac{1}{6}\left[1 - \frac{(0.94)^2}{(1.02)(1.28)}\right] \times \frac{1.02}{1.28} = 0.04$$

Assuming a 5% significance level, $t_{0.025,6} = 2.447$

The acceptance interval for the null hypothesis is $0 \pm t_{0.025,6}$.

$$\sigma_{\substack{\text{estimated} \\ \text{power of A}}} = 2.447 \times 0.04 = 0 \pm 0.10$$

Since 0.5 falls outside this interval we reject the null hypothesis that the size of the drainage area does not have a relation to the mean discharge.

With the above conditions satisfied, and the multiple correlation coefficient being close to unity, the mathematical relationship is accepted as representing the hydrologic system of the region. Therefore the mean discharge at any site within the region can be determined from the mathematical relationship $Q_{2.33}=0.21\ A^{0.5}L^{0.4}$, if the drainage area and the length of main channel at such site are known.

The mean discharge at such site will then be used to determine the design flood of any structure to be constructed at the site from the regional frequency curve of Fig. 4.17.

4.10 Determination of a Frequency Curve

To define a frequency curve at any location, either gauged or ungauged, within the homogenous region, the mean annual event is determined from the curve relating this event to the drainage area and/or the length of main channel, or from the mathematical relationship developed. The mean annual event is then multiplied by the median ratios for the various return periods required (Fig. 4.17). A complete frequency curve can be produced by plotting discharges for several frequencies and drawing the curve defined by the points.

With the regional flood-frequency analysis having been completed the next chapter is a discussion on all the issues raised and the results of the analysis.

Chapter 5

DISCUSSION OF RESULTS

Determination of Missing Record

In determining the missing record of the annual maximum discharges within the base period, correlation of individual data at stations having the missing record with a station having a complete record was made. Dalrymple (1960) found that the correlation will not be perfect. If it were, there would be no independence and there would be only one record that had been measured twice.

It is observed that a certain point repeatedly fall out of correlation for all the six correlations of Figs. 4.1 to 4.6. The water year in which this occured for the continous record of station 9 was 1972. The maximum annual peak discharges of station 9 for the water year was 33.67 m^3/s which is small compared to the other maximum annual peak discharges of the same station. The discharges for all the other stations in the same water year was also small. It can be concluded that a very low precipitation was recorded in this period which altered the trend of correlation of discharges of station 9 with the discharges of other stations.

Correlation was made with the discharges of station 9 because station 9 drains the largest area of the region.

Goodness-of-fit-Test

It was observed in the course of the review of literature that one or two other test of goodness-of-fit of distribution exist. Yevjevich (1970) noted that the K-S test is the most powerful of all these tests, and subsequently it has been adopted in the study.

Homogeneity Test

In testing the hydrological homogeneity of the region the 10-year flood was used. It was decided to make the test on the 10-year flood because this is the longest recurrence interval for which most record will give dependable estimates Dalrymple (1960).

Regional Frequency Curve

The plot of each median flood ratio against the corresponding recurrence interval was made for all the three probability distribution. The Gumbel extreme-event probability distribution was finally chosen because all the points plot closest to a straight line on, Fig. 4.17.

Estimation of Mean Annual Flood

The mean annual flood, which is the 2.33-year flood, has been determined graphically from the preliminary frequency curves. The graphical mean was used because it is more stable and dependable than an arithmetic mean.

This method of determining the mean gives greater weight to the medium floods than to the extreme floods with large sampling errors, and for this reason is not influenced adversely by the chance inclusion or exclusion of a major flood, as is the arithmetic mean.

A mathematical relationship of the form $Q2.33 = aA^{0.5}L^{b}$ was assumed. Previously, $Q2.33 = aA^{b}L^{o}$, Benson (1964), was applied and the constant a, b, and c were determined but because the difference between the observed and the estimated values of Q2.33 is smaller when the power of A is fixed at 0.5 than when it is determined, it was decided to use the relationship with the power of A fixed at 0.5.

The observed and the graphically/analytically estimated mean annual floods compare fairly well for all the stations except station number one and, probably number nine. The discrepancy of the observed and the graphically/analytically estimated values of station numbers one and nine might be a result of incorrect data observation or misinformation on data or it might even be natural.

It is noted that though the mathematical relationship is acceptable, having passed the significance test and having a multiple correlation coefficient close to unity, the graphical estimation of the mean discharge seem to give

a better estimate because it gives a smaller difference between the observed and the estimated values of Q2.33 than if the mathematical relationship is used, Tables 4.6 and 4.8.

Design Floods and Risk Analysis

The recurrence interval of the design flood for any structure to be built on any site is based on the type of structure. For example, the design flood of a moderately costly dam with the return period of 100 years can be determined from the regional frequency curve if the mean discharge at the site on which the dam will be built is known.

Both single station and regional frequency analyses involve risks. In determining the design flood for a project, the length of data available, the project life and the allowable probability of failure are all factors to be considered.

Risk analysis is an aid to assessing the probability that the derived design event will differ from the actual future event. In other words, risk is the calculable probability of failure, for example, occurrence of a certain flood.

The procedure for the risk analysis, Yen (1971), is as follows:

If we have an hydrological system that does not vary with time, the probability of an occurrence of an event, X, greater than the design flood x is $P(X \geqslant x) = \frac{1}{T}$, where T is the time that, on average, elapses between two events that equal or exceed the particular level x. Putting in another way, the N-year event, the event that is expected to be equalled or exceeded on average, every N-years, has a recurrence interval, T of N-years.

If the probability $P(X \leqslant x)_n$ represent the probability that x will not be equalled or exceeded in n such periods. (n is usually the design life of the project), then, for an independent series and applying the multiple probability rule,

$$P(X \leqslant x)_n = (P(X \leqslant x))^n$$
$$= (1 - P(X \geqslant x))^n \quad --- \quad 5.1$$

therefore

$$P(X \geqslant x)_n = 1 - (1 - P(X \geqslant x))^n \quad -- \quad 5.2$$

since

$$T = \frac{1}{P(X \geqslant x)} \quad --- \quad 5.3$$

then

$$P(X \geqslant x)n = 1 - (1 - \frac{1}{T})^n \quad --- \quad 5.4$$

This is the risk of failure and is based on the assumption of independence of annual event.

If the project life of the dam mentioned above is 50 years, then the risk of failure in this period is:

$$P(X \geqslant \text{design flood})_{50} = 1 - \left(1 - \frac{1}{100}\right)^{50}$$

$$= 0.39$$

It is often good to specify the simple risk, as determined from equation (5.4), associated with hydrologic design of structures as a check on the design. In the hydrologic design of engineering projects there are other risks and uncertainties involved in addition to the basic simple risk calculated as above. Some of these risks and uncertainties include the risk due to limited record of data available. A risk is involved because hydrologic data of a limited period may not and usually do not represent exactly the true case of the complete population which normally is of an infinite number of data over an infinite period. The other risks include the uncertainty of the mathematical techniques in handling the data and measurement errors.

Yen (1971) found that the evaluation of the simple risk depends on neither the type of distribution nor the plotting position of the hydrologic data. However, in attempts to evaluate the additional risks and uncertainties mentioned above, knowledge of the characteristics of the hydrologic data is usually necessary.

Nevertheless, the evaluation of the combined risk, if possible, is by no means a simple task, particularly for complex hydrologic systems.

Advantages and Disadvantages of the Study

The regional flood frequency analysis for Ogun/ Oshun basin in Western Nigeria was performed but not without the following merit and demerit.

Since many records from many gauging stations were combined in the analysis the sampling errors have been greatly reduced. Usually, the flood records at individual stations are generally short thereby making the sampling errors correspondingly large if frequency analysis is performed independently for each station.

It is only rarely that flood-frequency information is needed at a gauging station. More often it is required at an ungauged site. This analysis has now provided the means of obtaining this information at an ungauged site.

The most serious defects of the analysis are that the "true" distribution of peak event cannot be known despite the goodness-of-fit test and the confidence region analysis, and all the relationships, except the mathematical correlation of mean discharge with drainage basin and channel lengths, are derived by visual-aid graphical-curve fitting. The method has little place in determining maximum limits of flood design, that is, "the maximum possible flood". With the ordinary stream-flod record

of 25-year length, errors of sampling introduce large errors in judging the magnitude of the greater floods. However, if properly computed and conservatively interpreted, flood-frequency analysis is a valuable hydrologic tool.

Chapter 6

CONCLUSION AND SUGGESTION

6.1 Conclusion

In this study, a regional flood frequency analysis was performed for the Ogun/Oshun river basins of western Nigeria using the annual maximum stream-flow records, the drainage area and the lengths of main channel.

The region was found to be hydrologically homogenous, therefore streamsflow records from the region can be combined for regional flood studies. A relationship of the form $Q_{2.33} = 0.21\ A^{0.5}\ L^{0.4}$ between the mean annual flood, the drainage area and the length of main channel was found to be applicable within the region, and the regional median flood ratio was found to follow the Gumbel extreme event distribution.

To estimate the design flood for a basin the following steps should be taken:

1. Determine the area of the basin and length of the longest stream in the basin.
2. Use either Figure 4.18, 4.19 or the equation $Q_{2.33} = 0.21\ A^{0.5}\ L^{0.4}$ to estimate the mean annual flood for the basin.

3. Determine the ratio of the design flood/mean annual flood from the regional flood frequency curve, Figure 4.17, by fixing the return period of the design flood.

4. Using the value of mean flood for the basin the design flood Q_T is then evaluated.

It is hoped that the results of this study will meet the urgent need of hydrologists and allied profesional engaged in flood control, planning and development of water at any place with inadequate or non-existent hydrological information.

6.2 Suggestion

It is suggested that the authority in charge of the region studied should try and provide better topographical maps with, at most, 10-metre contour intervals so that estimate of some other physiographic factors, for example, the mean slopes of the channels for the basin can be determined. At present the best available topographical maps have contour intervals of not less than 35 metres.

To harness the improvement of studies within the basin, it is suggested that the means of data collection should be greatly improved by the River Basin Authority. Automatic measuring devices should be acquired in order to have continous and more reliable data.

For example, recording raingauges will help in obtaining data on intensity-duration of precipitation. Lack of intensity-duration data on precipitation has made it impossible for the present study to also correlate mean discharge with rainfall.

Likewise, recording devices to streamflow could yield more reliable data. At present, stage height is being measured annually once or twice in a day and there were cases when gauge readers absent themselves from gauging stations due to one reason or the other, for example, industrial action. Who knows, the maximum annual event could have occured in this period.

The flood estimating capabilities of this study will improve with each passing year as more data become available. It is suggested that the relationships derived in the study be updated from time to time when more data will be available. Rating curves used for the estimation of discharges from stage heights should also be modified from time to time.

KASHIM IBRAHIM LIBRARY ABU ZARIA

313564

Finally, it is suggested that the 10,000-year flood from the frequency curve derived for any site be used as the "probable maximum flood", ASCE (1973), by any agency to be involved in the design and construction of structures on sites where failure may result in loss of lives and huge sum of properties.

REFERENCES

ASCE (1973). Re-evaluating spillway adequacy of existing dams. Proceedings American Society of Civil Engineers, Vol. 99, No. Hy 2, pp. 337 - 372.

Beard, L.R. (1962). Statistical Methods in Hydrology. In: Dune, T. and Leopold, L.B. (Ed.) Water in Environmental Planning, W.H. Freeman and Company, San Francisco pp. 308 - 311.

Benjamin, J.R. and Cornel, C.A. (1970). Probability, Statistics, and Decision for Civil Engineers, MoGraw-Hill, U.S.A. pp. 380 - 435.

Benson, M.A. (1964). Factors Affecting the Occurence of Floods in the Soughwest, U.S. Geological Survey Watersupply Paper 1580-D pp. 46 - 52.

Chow, Ven Te (1964). Handbook of Applied Hydrology, McGraw-Hill, New York.

Dalrymple Tate (1960). Flood-Frequency Analysis, U.S. Geological Survey water supply paper 1543-A pp. 1 - 46.

Federal Surveys (1978). The National Atlas of the Federal Republic of Nigeria. Federal Surveys, Lagos. pp. 31-53.

Gumbel, E. J. (1941). The Return Period of Flood Flow. Ann. Math. Stats. Vol. 12, pp. 163 - 190.

Gumbel, E.J. (1958). Statistics of Extremes. In: Yevjerich Vujica (Ed.). Probability and Statistics in Hydrology, Water Resources Publications, Fort Collins, Colorado, U.S.A. pp. 89 - 91.

Reference (Cont'd)

GKW-Consulting Engineers (1975). Lagos Water Supply Expansion Programme. Water Resources for Phase II, Hydrological Report Vol. I.

Hazen, A. (1930). Floodflows, A study of Frequencies and Magnitudes. In: Dalrymple Tate (Ed.) Flood Frequency Analysis, U.S. Geological Survey Water Supply paper 1543-A pp. 15.

Horton, R.E. (1933). The Role of Infiltration in the Hydrologic Cycle. Trans. Am. Geophy. Un. Vol. 14, pp 448 - 460.

Kite, G.W. (1977). Frequency and Risk Analysis in Hydrology, Water Resources Publications, Fort Collins, Colorado, U.S.A. pp. 44 - 128.

Oorbda (180) Activities of the Basin. In: Annual Report 1980 - 81, pp. 4 - 9. Ogun-Oshun River Basin Development Authority, Abeokuta, Nigeria.

Oyegoke , E.S., Ige, E.D. and Oyebande, L. (1983). River Flow Forecasting through a Regression Model: a case study of the Basement Complex of Western Nigeria, Proceedings of the Hamburg Symposium. IAHS Publication No. 140.

Samuelson, B. (1972). Statistical Interpretation of Hydrometeorological Extreme Values, Nordic Hydrology, Vol. 3, No.4,pp. 199-214.

Sherman, L.K. (1932). Streamflow from Rainfall by the Unitgraph Method, Engugu. News Record, 108, pp. 501 - 505.

References (Cont'd)

Verma, R.D. (1973). Flood Control Planning with Inadequate Hydrologic Data, Proceedings of Second International Symposium in Hydrology, Water Resources Publications, Fort Collins, Colorado, U.S.A. pp. 259 - 268.

Ward Roy (1978). Floods: A Geological Perspective. The Macmillan Press Ltd., London, pp. 68 - 92.

Yen, B.C. (1971). Risks in Hydrologic Design of Engineering Projects. Proceeding American Society of Civil Engineers, Vol. 96, No. H14 pp. 959 - 966.

Yevjevich Vujica (1972). Probability and Statistics in Hydrology, Water Resources Publication, Fort Collins, Colorado, U.S.A. pp. 89 - 230.

APPENDIX A-1

EXCEEDENCE PROBABILITY ANALYSIS FOR STATION 1

Water Year	Maximum Annual peak discharges (m^3/s)	Order Number	Recurrence Interval (T years)	Cumulative function $F_n(x) = \frac{1}{T}$	Theoritical distribution function f(x)	Fn(x) - f(x)
1	2	3	4	5	6	7
1965	152.40	17	1.47	0.68	0.68	0.00
1966	165.51	15	1.67	0.60	0.60	0.00
1967	137.27	19	1.32	0.76	0.78	0.02
1968	416.57	1	25.00	0.04	0.03	0.01
1969	345.57	2	12.50	0.08	0.07	0.01
1970	147.33	18	1.39	0.72	0.72	0.00
1971	98.66	23	1.09	0.92	0.93	0.01
1972	88.40	24	1.04	0.96	0.95	0.01
1973	163.51	16	1.56	0.64	0.64	0.00
1974	112.06	22	1.14	0.88	0.88	0.00
1975	201.02	12	2.08	0.48	0.43	0.05
1976	231.52	7	3.57	0.28	0.31	0.03
1977	130.04	20	1.25	0.80	0.80	0.00
1978	202.47	11	2.27	0.44	0.43	0.01
1979	251.24	5	5.00	0.20	0.22	0.02
1980	288.31	3	8.33	0.12	0.16	0.04
1981	214.03	10	2.50	0.40	0.40	0.00
1982	126.28	21	1.19	0.84	0.82	0.02
1983	178.29	14	1.79	0.56	0.56	0.00
1984	219.61	9	2.78	0.36	0.36	0.00
1985	276.30	4	6.25	0.16	0.19	0.03
1986	191.47	13	1.92	0.52	0.48	0.04
1987	250.72	6	4.17	0.24	0.25	0.01
1988	227.83	8	3.13	0.32	0.32	0.00

APPENDIX A-2

EXCEEDENCE PROBABILITY ANALYSIS FOR STATION 2

Water Year	Maximum Annual peak discharges (m^3/s)	Order Number	Recurrence Interval (T years)	Cumulative function $Fn(x) = \frac{1}{T}$	Theoritical distribution function f(x)	Fn(x) - f(x)
1	2	3	4	5	6	7
1965	61.78	12	2.08	0.48	0.48	0.00
1966	43.24	20	1.25	0.08	0.84	0.04
1967	81.10	5	5.00	0.20	0.24	0.04
1968	115.98	1	25.00	0.04	0.05	0.04
1969	101.81	2	12.50	0.08	0.09	0.01
1970	72.87	9	2.78	0.36	0.33	0.03
1971	42.05	22	1.14	0.88	0.85	0.03
1972	43.10	21	1.19	0.84	0.83	0.01
1973	34.83	23	1.09	0.92	0.93	0.01
1974	44.35	19	1.32	0.76	0.81	0.05
1975	54.90	15	1.67	0.60	0.60	0.00
1976	58.50	14	1.79	0.56	0.56	0.00
1977	27.47	24	1.04	0.96	0.98	0.02
1978	60.60	13	1.92	0.52	0.52	0.00
1979	74.79	8	3.13	0.32	0.32	0.00
1980	80.63	6	4.17	0.24	0.24	0.00
1981	67.55	10	2.50	0.40	0.04	0.00
1982	46.18	18	1.39	0.72	0.78	(0.06)
1983	53.05	17	1.47	0.68	0.64	0.04
1984	93.93	3	8.33	0.12	0.14	0.02
1985	82.47	4	6.25	0.16	0.22	0.06
1986	54.05	16	1.56	0.64	0.64	0.00
1987	78.87	7	3.57	0.28	0.27	0.01
1988	67.40	11	2.27	0.44	0.42	0.02

APPENDIX A-3

EXCEEDENCE PROBABILITY ANALYSIS FOR STATION 3

Water Year	Maximun Annual peak discharges (m^3/s)	Order Number	Recurrence Interval (T years)	Cumulative function $Fn(x) = \frac{1}{T}$	Theoritical distribution function $f(x)$	$Fn(x) - f(x)$
1	2	3	4	5	6	7
1965	(15.38)	15	1.67	0.60	0.61	0.01
1966	13.76	17	1.47	0.68	0.73	0.05
1967	18.46	13	1.92	0.52	0.52	0.04
1968	36.93	2	12.50	0.08	0.09	0.01
1969	17.10	12	2.08	0.48	0.63	0.15
1970	15.38	16	1.56	0.64	0.64	0.00
1971	8.01	23	1.09	0.92	0.94	0.02
1972	11.24	21	1.19	0.84	0.84	0.00
1973	20.53	10	2.50	0.40	0.40	0.00
1974	19.51	11	2.27	0.44	0.44	0.00
1975	13.71	18	1.39	0.72	0.72	0.00
1976	16.90	14	1.79	0.56	0.59	0.03
1977	10.47	22	1.14	0.88	0.88	0.00
1978	21.82	9	2.78	0.36	0.36	0.00
1979	25.01	6	4.17	0.24	0.24	0.00
1980	30.68	3	8.33	0.12	0.12	0.00
1981	22.50	8	3.13	0.32	0.32	0.00
1982	(7.37)	24	1.04	0.96	0.94	0.02
1983	(11.63)	20	1.25	0.80	0.81	0.01
1984	38.92	1	25.00	0.04	0.04	0.00
1985	27.43	4	6.25	0.16	0.19	0.03
1986	(12.61)	19	1.32	0.76	0.78	0.02
1987	26.25	5	5.00	0.20	0.21	0.01
1988	23.75	7	3.57	0.28	0.28	0.00

APPENDIX A-4

EXCEEDENCE PROBABILITY ANALYSIS FOR STATION 4

Water Year	Maximun Annual Peak discharges (m^3/s)	Order Number	Recurrence Interval (T years)	Cumulative function $Fn(x) = \frac{1}{T}$	Theoritical distribution function f(x)	Fn(x) - f(x
1	2	3	4	5	6	7
1965	(12.62)	16	1.56	0.64	0.64	0.00
1966	13.97	13	1.92	0.52	0.57	0.05
1967	16.68	12	2.08	0.48	0.41	0.07
1968	35.03	1	25.00	0.04	0.04	0.00
1969	10.28	19	1.32	0.76	0.76	0.00
1970	11.61	17	1.47	0.68	0.68	0.00
1971	27.50	3	8.33	0.12	0.11	0.01
1972	5.13	23	1.09	0.92	0.94	0.02
1973	13.20	14	1.79	0.56	0.62	0.06
1974	28.26	2	12.50	0.08	0.10	0.02
1975	12.74	15	1.67	0.06	0.60	0.00
1976	9.68	20	1.25	0.80	0.80	0.00
1977	6.88	22	1.14	0.88	0.90	0.02
1978	(17.29)	11	2.27	0.44	0.40	0.04
1979	(20.48)	7	3.57	0.28	0.28	0.00
1980	(24.19)	4	6.25	0.16	0.18	0.02
1981	(18.54)	10	2.50	0.40	0.36	0.04
1982	(3.37)	24	1.04	0.96	0.97	0.01
1983	(8.92)	21	1.19	0.84	0.82	0.02
1984	22.93	5	5.00	0.20	0.20	0.00
1985	21.08	6	4.17	0.24	0.26	0.02
1986	10.50	18	1.39	0.72	0.74	0.02
1987	21.00	8	3.13	0.32	0.26	0.06
1988	19.10	9	2.78	0.36	0.34	0.02

APPENDIX A-5

EXCEEDENCE PROBABILITY ANLAYSIS FOR STATION 5

Water Year	Maximum Annual peak discharges (m^3/s)	Order Number	Recurrence Interval (T years)	Cumulative function $Fn(x) = \frac{1}{T}$	Theoritical distribution function f(x)	Fn(x) - f(x
1	2	3	4	5	6	7
1965	(48.40)	16	1.56	0.64	0.64	0.00
1966	62.18	13	1.92	0.52	0.51	0.01
1967	42.14	17	1.47	0.68	0.71	0.03
1968	89.39	9	2.78	0.36	0.29	0.07
1969	54.00	15	1.67	0.60	0.59	0.01
1970	66.45	12	2.08	0.48	0.45	0.03
1971	37.01	19	1.32	0.76	0.78	0.02
1972	5.27	23	1.09	0.92	0.96	0.04
1973	92.00	8	3.13	0.32	0.27	0.05
1974	84.35	11	2.27	0.44	0.32	0.12
1975	33.98	20	1.25	0.80	0.80	0.00
1976	20.93	22	1.14	0.88	0.89	0.01
1977	55.82	14	1.79	0.56	0.57	0.01
1978	(87.12)	10	2.50	0.40	0.29	0.11
1979	105.98	5	5.00	0.20	0.19	0.01
1980	138.74	2	12.50	0.08	0.08	0.00
1981	(97.64)	6	4.17	0.24	0.24	0.00
1982	(4.05)	24	1.04	0.96	0.97	0.01
1983	41.97	18	1.39	0.72	0.72	0.00
1984	162.72	1	25.00	0.04	0.03	0.01
1985	128.83	3	8.33	0.12	0.10	0.02
1986	(32.01)	21	1.19	0.84	0.80	0.04
1987	115.90	4	6.25	0.16	0.16	0.00
1988	97.01	7	3.57	0.28	0.24	0.04

APPENDIX A-6

EXCEEDENCE PROBABILITY ANALYSIS FOR STATION 6

Water Year	Maximum Annual peak discharges (m^3/s)	Order Number	Recurrence Interval (T years)	Cumulative function $F(x) = \frac{1}{T}$	Theoritical distribution function f(x)	Fn(x) - f(x)
1	2	3	4	5	6	7
1965	(57.96)	16	1.56	0.64	0.64	0.00
1966	63.49	15	1.67	0.60	0.59	0.01
1967	92.54	11	2.27	0.44	0.41	0.03
1968	201.56	1	25.00	0.04	0.06	0.02
1969	72.09	14	1.79	0.56	0.56	0.00
1970	26.68	22	1.14	0.88	0.86	0.02
1971	36.05	19	1.32	0.76	0.79	0.03
1972	12.83	23	1.09	0.92	0.92	0.00
1973	91.74	12	2.08	0.48	0.42	0.06
1974	84.48	13	1.92	0.52	0.43	0.09
1975	95.92	10	2.50	0.40	0.40	0.00
1976	28.91	20	1.25	0.80	0.84	0.04
1977	49.68	17	1.47	0.68	0.70	0.02
1978	(10.01)	9	2.78	0.36	0.36	0.00
1979	(125.36)	6	4.17	0.24	0.24	0.00
1980	172.01	3	8.33	0.12	0.10	0.02
1981	100.54	8	3.13	0.32	0.36	0.04
1982	(5.42)	24	1.04	0.96	0.94	0.02
1983	(27.56)	21	1.19	0.84	0.84	0.00
1984	187.08	2	12.50	0.08	0.07	0.01
1985	138.31	4	6.25	0.16	0.19	0.03
1986	(40.06)	18	1.39	0.72	0.78	0.06
1987	128.87	5	5.00	0.20	0.22	0.02
1988	112.08	7	3.57	0.28	0.28	0.00

APPENDIX A-7

EXCEEDENCE PROBABILITY ANALYSIS FOR STATION 7

Water Year	Maximum Annual peak discharges (m^3/s)	Order Number	Recurrence Interval (T years)	Cumulative function $F(x) = \frac{1}{T}$	Theoritical distribution function f(x)	Fn(x) - f(x)
1	2	3	4	5	6	7
1965	(13.01	16	1.56	0.64	0.63	0.01
1966	15.93	13	1.92	0.52	0.53	0.01
1967	7.99	20	1.25	0.80	0.81	0.01
1968	40.31	1	25.00	0.04	0.56	0.00
1969	15.31	14	1.79	0.56	0.56	0.00
1970	11.92	17	1.47	0.68	0.67	0.01
1971	17.13	11	2.27	0.44	0.45	0.01
1972	1.49	23	1.09	0.92	0.95	0.03
1973	22.61	8	3.13	0.32	0.32	0.00
1974	16.28	12	2.08	0.48	0.52	0.04
1975	11.43	18	1.39	0.72	0.69	0.03
1976	14.29	15	1.67	0.60	0.58	0.02
1977	6.59	22	1.14	0.88	0.85	0.03
1978	19.51	10	2.50	0.40	0.42	0.02
1979	26.82	6	4.17	0.24	0.23	0.01
1980	(33.20)	3	8.33	0.12	0.13	0.01
1981	22.47	9	2.78	0.36	0.38	0.02
1982	(1.31)	24	1.04	0.96	0.95	0.01
1983	(6.64)	21	1.19	0.84	0.85	0.01
1984	(38.20)	2	12.50	0.08	0.07	0.01
1985	29.46	4	6.25	0.16	0.18	0.02
1986	(9.28)	19	1.32	0.76	0.78	0.02
1987	28.40	5	5.00	0.20	0.19	0.01
1988	24.21	7	3.57	0.28	0.28	0.00

APPENDIX A-8

EXCEEDENCE PROBABILITY ANALYSIS FOR STATION 8

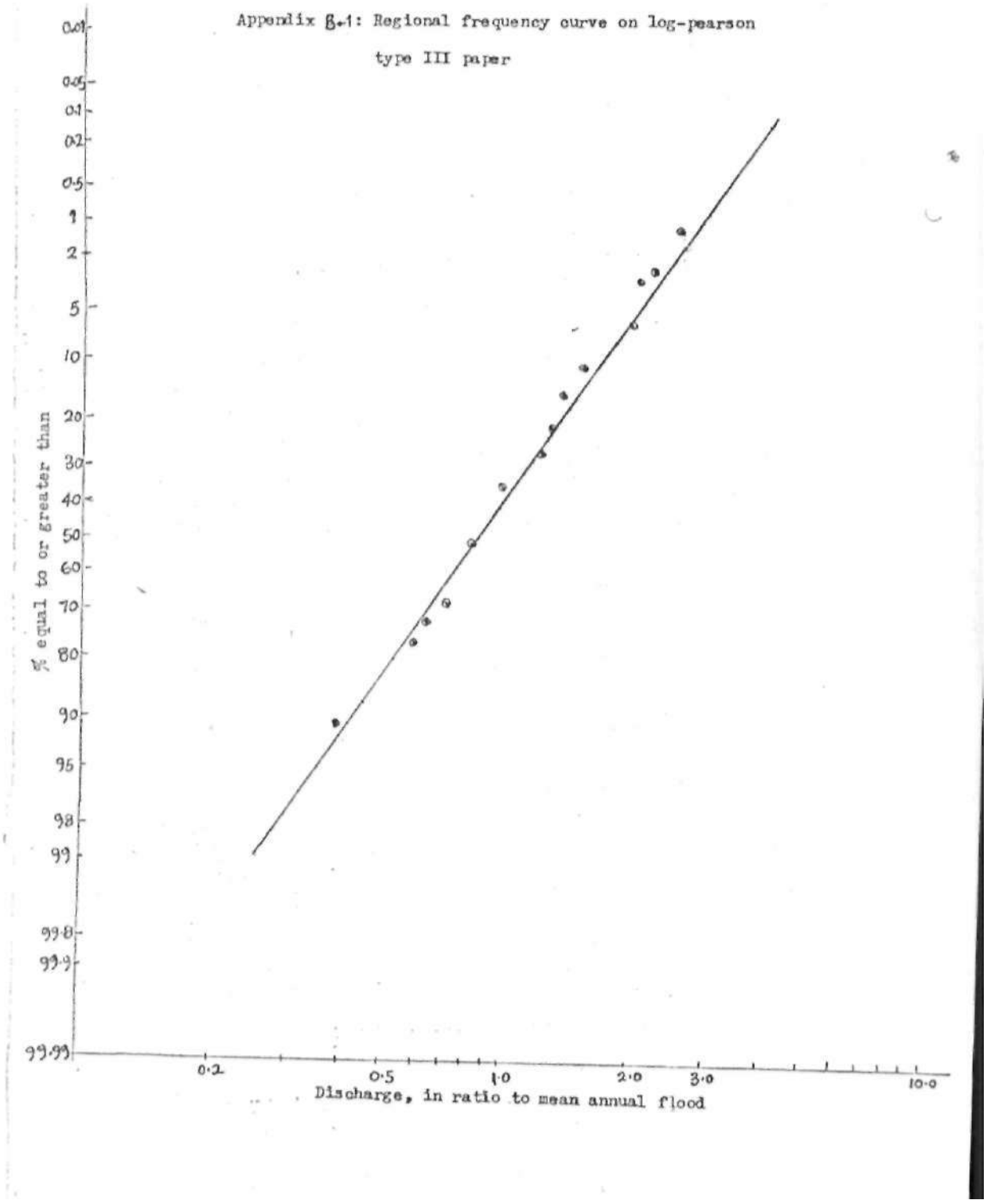
Water Year	Maximum Annual peak discharges (m^3/s)	Order Number	Recurrence Interval (T years)	Cumulative function $F(x) = \frac{1}{T}$	Theoritical distribution function f(x)	Fn(x) - f(x)
1	2	3	4	5	6	7
1965	(88.60)	14	1.79	0.56	0.61	0.05
1966	142.89	9	2.78	0.36	0.37	0.01
1967	69.83	20	1.25	0.80	0.75	0.05
1968	305.49	1	25.00	0.04	0.05	0.01
1969	140.12	10	2.50	0.40	0.39	0.01
1970	75.54	17	1.47	0.68	0.70	0.02
1971	75.72	16	1.56	0.64	0.71	0.07
1972	82.98	15	1.67	0.60	0.67	0.07
1973	114.59	12	2.08	0.48	0.51	0.03
1974	112.62	13	1.92	0.52	0.53	0.01
1975	66.06	22	1.14	0.88	0.77	0.11
1976	72,38	19	1.32	0.76	0.73	0.03
1977	66.68	21	1.19	0.84	0.76	0.08
1978	123.09	11	2.27	0.44	0.44	0.00
1979	183.11	6	4.17	0.24	0.24	0.00
1980	232.06	3	8.33	0.12	0.13	0.01
1981	153.45	8	3.13	0.32	0.32	0.00
1982	14.57	24	1.04	0.96	0.93	0.03
1983	58.21	23	1.09	0.92	0.80	0.12
1984	246.44	2	12.50	0.08	0.10	0.02
1985	220.09	4	6.25	0.16	0.15	0.01
1986	73.35	18	1.39	0.72	0.72	0.00
1987	203.87	5	5.00	0.20	0.18	0.02
1988	166.47	7	3.57	0.28	0.28	0.00

APPENDIX A-9

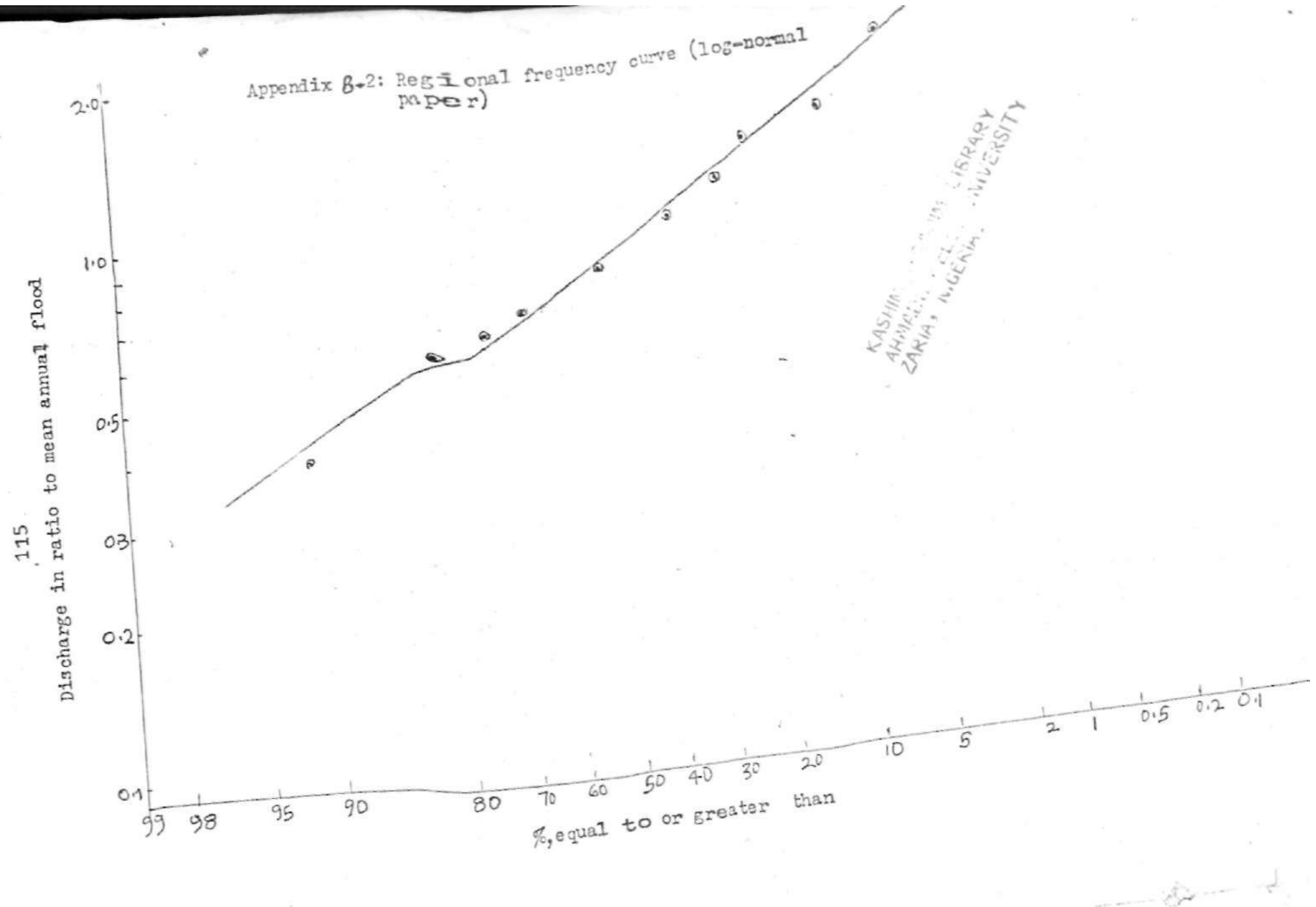
EXCEEDENCE PROBABILITY ANALYSIS FOR STATION 9

Water Year	Maximum Annual peak discharges (m^3/s)	Order Number	Recurrence Interval (T years)	Cumulative function $F(x) = \frac{1}{T}$	Theoritical distribution function f(x)	Fn(x) - f(x
1	2	3	4	5	6	7
1965	164.10	13	1.92	0.52	0.54	0.02
1966	199.94	12	2.08	0.48	0.49	0.01
1967	162.06	16	1.56	0.64	0.70	0.08
1968	313.31	2	12.50	0.08	0.08	0.00
1969	163.11	15	1.67	0.60	0.70	0.10
1970	164.10	14	1.79	0.56	0.68	0.12
1971	157.85	17	1.47	0.68	0.75	0.07
1972	33.67	24	1.04	0.96	0.94	0.02
1973	202.57	11	2.27	0.44	0.44	0.00
1974	153.11	18	1.39	.072	0.80	0.08
1975	203.36	10	2.50	0.40	0.46	0.06
1976	147.32	19	1.32	0.76	0.76	0.00
1977	144.69	21	1.19	0.84	0.84	0.00
1978	214.67	9	2.78	0.36	0.35	0.01
1979	244.33	6	4.17	0.24	0.24	0.00
1980	283.05	3	8.33	0.12	0.12	0.00
1981	228.65	7	3.57	0.28	0.29	0.01
1982	70.77	23	1.09	0.92	0.91	0.01
1983	128.91	22	1.14	0.88	0.85	0.03
1984	351.28	1	25.00	0.04	0.04	0.00
1985	254.10	4	2.75	0.16	0.20	0.04
1986	144.69	20	1.25	0.80	0.80	0.00
1987	253.26	5	5.00	0.20	0.20	0.00
1988	228.65	8	3.13	0.32	0.31	0.01

Appendix 8.1: Regional frequency curve on log-pearson type III paper



Appendix B-2: Regional frequency curve (log-normal paper)



KASHIF ... LIBRARY
AHMADU ... UNIVERSITY
ZARIA, NIGERIA.