ESTIMATION OF RUN-OFF AND TIME SERIES ANALYSIS FOR 1K3 AND 1K3A

A Dissertation
submitted in partial fulfilment of the
requirements for the award of the
MASTER'S DEGREE
in
HYDROLOGY

By

A. S. B. LUHUMBIKA
Ministry of Water Development and Power
Tanzania

UNESCO SPONSORED
INTERNATIONAL HYDROLOGY COURSE
UNIVERSITY OF ROORKEE
ROORKEE (U.P.)
INDIA
April, 1975

C E R T I F I C A T E

This is to certify that the dissertation entitled "ESTIMATION OF RUN-OFF AND TIME SERIES ANALYSIS FOR 1K3 and 1K3A GAUGING STATIONS" which is being submitted by Sri ALPHAXAD SAKUMI BUTINGO LUHUMBIKA in partial fulfilment of the requirements for the award of the MASTER'S DEGREE in HYDROLOGY of the University of Roorkee, Roorkee is a record of the candidate's own work carried out by him under our supervision and guidance. The material embodied in this dissertation has not been submitted for this award of any other Degree or Diploma.

This is further to certify that he has worked for a period of six months (from 1.10.1974 to 31.3.1975) for the preparation of this dissertation at this University.

B S Mathur

(B.S.Mathur)
Reader, Hydrology
University of Roorkee
Roorkee, U.P. India

Satish Chandra

(Satish Chandra)
Professor Hydrology
University of Roorkee
Roorkee , U.P. India.

Dated
April , 1975.

ACKNOWLEDGEMENTS

The author would like to express his deep sense of gratitude to Dr. Satish Chandra, Professor, Hydrology University of Roorkee, Roorkee and Dr. B.S. Mathur, Reader Hydrology , University of Roorkee, Roorkee for their keen interest, guidance , inspiration and encouragement, not only during the course of the present study, but also throughout my period of study at this University.

The author is very grateful for the kindness of the Principal Secretary, Ministry of Water Development and Power , Government of Tanzania, for making available some of the data from the Catchment.

Also , the prompt cooperation and guidance of the officers and staff of the Computer Centre, of S.E.R.C. Roorkee and to all persons, even whose names are not mentioned here who in one way or another contributed to the completion of this work, I wish to express my sincere gratitude.

A.S.B.Lubumbika
Tanzania

April, 1975.

TABLE OF CONTENTS

--

LIST OF TABLES

LIST OF FIGURES

SYNOPSIS

The peak discharge to be adopted for the design of hydraulic structures and time analysis of the sequences important in the planning of water Resources Projects. To arrive at the values of the peak discharge, a study of the available data has to be conducted and analysed, to arrive at a desired accuracy.

In this s-tudy, based on the available data situated at 1K3 the various methods for estimation of peak flow have been discussed bringing out their suitability and limitations. For this purpose frequency analysis has been carried out with probability distribution functions. Fourprobability distribution functions, viz., Normal, log-normal 2, Gamma 2 and Gumbel have been used to find out the best fit distribution for 1K3 gauging station and the best fit arrived at. Determination of the flood values with a return period of 100 years, has been computed using these fit distribution functions.

A mathematical model has been formulated, for time series of monthly runoff values for the 1K3A gauging station, in order to generate synthetic stream flow for use in the analysis of water resources system. The monthly time series are first analysed for trend by using the least-square method. The periodicities are detected by constructing correlogram and a Fourier Series model

with eight parameters has been fitted to the cyclic component. The correlogram of the residual series, after removal of cyclic component indicated that the first lag is significant. Therefore, first order Morkov model is fitted on the stochastic component. The model so formulated, considering both the deterministic and stochastic models, can be used to generate synthetic monthly stream flows for 1K3A gauging station.

CHAPTER I

INTRODUCTION

1.1 SIGNIFICANCE OF STUDY

For the last few years, theoretical approaches have been used to a great extent in solving many hydrological problems. However, the task is not an easy one. In developing countries, like Tanzania, the problem is much more complicated due to non-availability of hydrologic data of sufficient length.

In the hydrologic design the magnitude of peak flood, is of great importance to the designer. Also in Water Resources assessment, pollution, control mass curve analyses for hydro-power schemes and formulation of control rule curve for reservoirs, it is necessary to predict the characteristics and quantity of stream flows to arrive at critical flow sequences of their associated return intervals. Existing stream flow records are not sufficiently extensive to provide estimate of many statistical parameters. Thus investigations aimed at obtaining the solution to these problems are hempered due to lack of long term data records.

Therefore, working hydrologists and engineers in the Ministry of Water Development and Power of Tanzania are in search of workable and satisfactory procedures that would serve a guide in estimating the design flood and overcoming the problem of inadequate historical data.

In this study, a natural catchment situated in Tanzania has been studied. The main hydrometeorological features of the basin have been discussed in subsequent para.

1.2. PHYSIOGRAPHICAL AND HYDROMETEOROLOGICAL FEATURES OF THE CATCHMENT

The two gauging stations 1K3 and 1K3A under consideration are situated on the catchment area of 158200 sq.km and 15800 sq.km respectively.

The climate ranges from the tropical humid heat of the coastal regions to temperate conditions of the southern Highlands and the high mountain ranges. Mean rainfall vary widely with both location and height. In some places minimum rainfall is below 250 mm per annum while in some places it is over 1750 mm on the higher parts of the southern highlands. The rainfall in these catchments is about 1700 mm. Over the whole Rufiji Basin, winds are generally easterly and changing to south easterly. From March to October these winds change to southerly direction.

1.3 DATA

The observed data, used in this study, for the two stations, is for the period of 12 years, from 1961-1972

Yearly maximum peak floods for 18 years have been used for flood frequency analysis for 1K3 and monthly stream flows for 12 years have been used for time series analysis for 1K3A. Since observed stream flow for 1K3A were too short, a correlation between 1K3 and 1K3A has been due to extrapolated the data as per Fig. 1. and Tables 1-3-1 & 2 for the period 1961-1972 only. Also rating curves and best fit lines by least square method have been developed as per Fig. .

For the purpose of estimation of design flood, it is necessary to study only the peak flows. The hydrologic design is related to the frequency with which the flows of a given magnitude will be equalled or exceeded. Information concerning probable extremes which proposed structures may be required to withstand and many other hydrologic problems can be solved by frequncy analysis, using past records of flood peak.

Having obtained the frequency of the floods the magnitude the economically accepted, for the greater the discharge, the higher the construction cost. Usually, the maximum recorded floods in the past are the most significant measures of estimating the design flood of rivers. The discharge or water level that provides the maximum excess benefit, is the most preferable design flood, however this approach also has many disadvantages, even though it is considered the most favourable one.

TABLE 1-3-1 MONTHLY DISCHARGES IN 10^2 Cumecs FOR 1K3 (

Month	1954	1955	1956	1957	1958	1959	1960	1961	1962	1963
November	340.0	410.0	400.0	420.0	480.0	470.0	650.0	650.0	730.0	670.0
December	360.0	360.0	380.0	380.0	355.0	-	785.0	785.0	4100.0	710.0
January	-	38.0	1580.0	480.0	410.0	480.0	600.0	-	1800.0	1760.0
February		1390.0	2720.0	880.0	700.0	780.0	1880.0	740.0	4399.0	2760.0
March		1520.0	2220.0	880.0	1410.0	1400.0	1900.0	1150.0	4400.0	4280.0
April		2380.0	3780.0	1920.0	2360.0	1300.0	1900.0	1600.0	2500.0	4180.0
May		2440.0	2220.0	2280.0	1780.0	1100.0	4520.0	1758.0	2510.0	2940.0
June		1400.0	1170.0	1000.0	660.0	510.0	1220.0	1540.0	840.0	1110.0
July		840.0	660.0	520.0	420.0	390.0	620.0	750.0	430.0	680.0
August		630.0	420.0	400.0	320.0	320.0	360.0	530.0	340.0	-
September		530.0	360.0	320.0	430.0	380.0	410.0	360.0	460.0	420.0
October		450.0	360.0	380.0	360.0	420.0	540.0	360.0	1600.0	600.0

TABLE 1-3-1 (Contd..)

Month	1964	1965	1966	1967	1968	1969	1970	1971	1972
November	630.0	190.0	340.0	440.0	500.0	310.0	190.0	-	440.0
December	1210.0	420.0	390.0	2250.0	-	310.0	300.0	400.0	1220.0
January	4200.0	520.0	730.0	590.0	-	600.0	1820.0	560.0	730.0
February	4530.0	560.0	1050.0	600.0	1940.0	1000.0	2420.0	1110.0	1140.0
March	4220.0	930.0	1850.0	730.0	-	1320.0	2290.0	1450.0	2180.0
April	3810.0	1890.0	2270.0	1760.0	4200.0	1700.0	1950.0	2880.0	2730.0
May	3820.0	1130.0	1340.0	2010.0	-	1510.0	1020.0	1910.0	2030.0
June	540.0	440.0	700.0	1140.0	-	760.0	500.0	780.0	1550.0
July	340.0	300.0	480.0	730.0	570.0	560.0	590.0	590.0	690.0
August	390.0	270.0	390.0	590.0	-	420.0	360.0	490.0	560.0
September	530.0	200.0	360.0	490.0	-	340.0	400.0	410.0	510.0
October	690.0	190.0	350.0	420.0	490.0	290.0	310.0	540.0	405.0

Month	1960	1961	1962	1963	1964	1965	1966	1967	1968	1969
Nov mber	- 0	450.0	-	500.0	340.00	300.0	310.0	350.0	440.0	300.0
December	-	1580.0	320.0	985.0	330.0	550.0	410.0	2480.0	480.0	320.0
January	-	-	4470.0	1330.0	2480.0	700.0	660.0	550.0	4439.0	-
February	760.0	600.0	3770.0	2600.0	2480.0	680.0	980.0	560.0	2840.0	1000.0
March	-	540.0	4450.0	3685.0	5000.0	320.0	13300	680.0	4740.0	1080.0
April	-	880.0	3780.0	4970.0	3780.0	2830.0	2800.0	1660.0	4495.0	1620.0
May	-	1420.0	2200.0	2600.0	2040.0	1580.0	1270.0	1840.0	3800.0	2140.0
June	-	660.0	875.0	880.0	810.0	600.0	610.0	950.0	1330.0	710.0
July	-	440.0	630.0	660.0	600.0	440.0	480.0	540.0	500.0	440.0
August	-	365.0	-	500.0	510.0	400.0	400.0	420.0	610.0	380.0
September	360.0	300.0	450.0	420.0	460.0	380.0	340.0	380.0	490.0	330.0
October	-	285.0	385.0	380.0	390.0	320.0	310.0	320.0	410.0	300.0

TABLE 1-3-2

Month	1970	1971	1972	1973	1974
November	270.0	-	-	270.0	-
December	370.0	280.0	350.0	290.0	-
January	1010.0	330.0	310.0	540.0	315.0
Fembruary	2740.0	380.0	390.0	-	330.0
March	3100.0	450.0	570.0	660.0	360.0
April	2940.0	640.0	690.0	-	560.0
May	1240.0	580.0	590.0	-	-
June	590.0	788.0	400.0	380.0	-
July	460.0	770.0	631.0	-	-
August	400.0	486.0	612.0	-	-
Sept.	360.0	500.0	480.0	-	-
Oct.	310.0	507.0	485.0	-	-

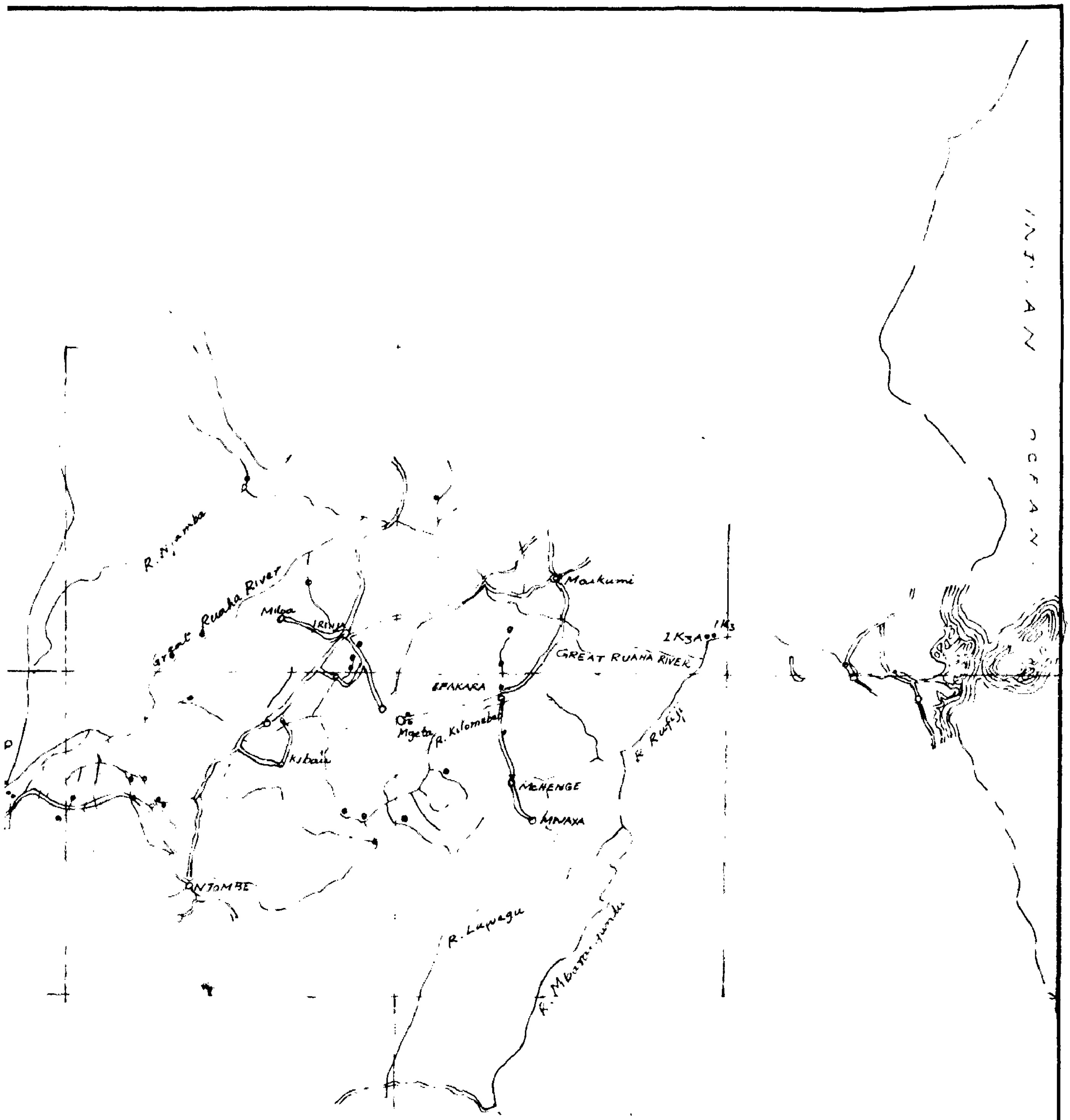

FIG. 1-1 MAP SHOWING 1K3 & 1K3A GAUGING STATIONS.

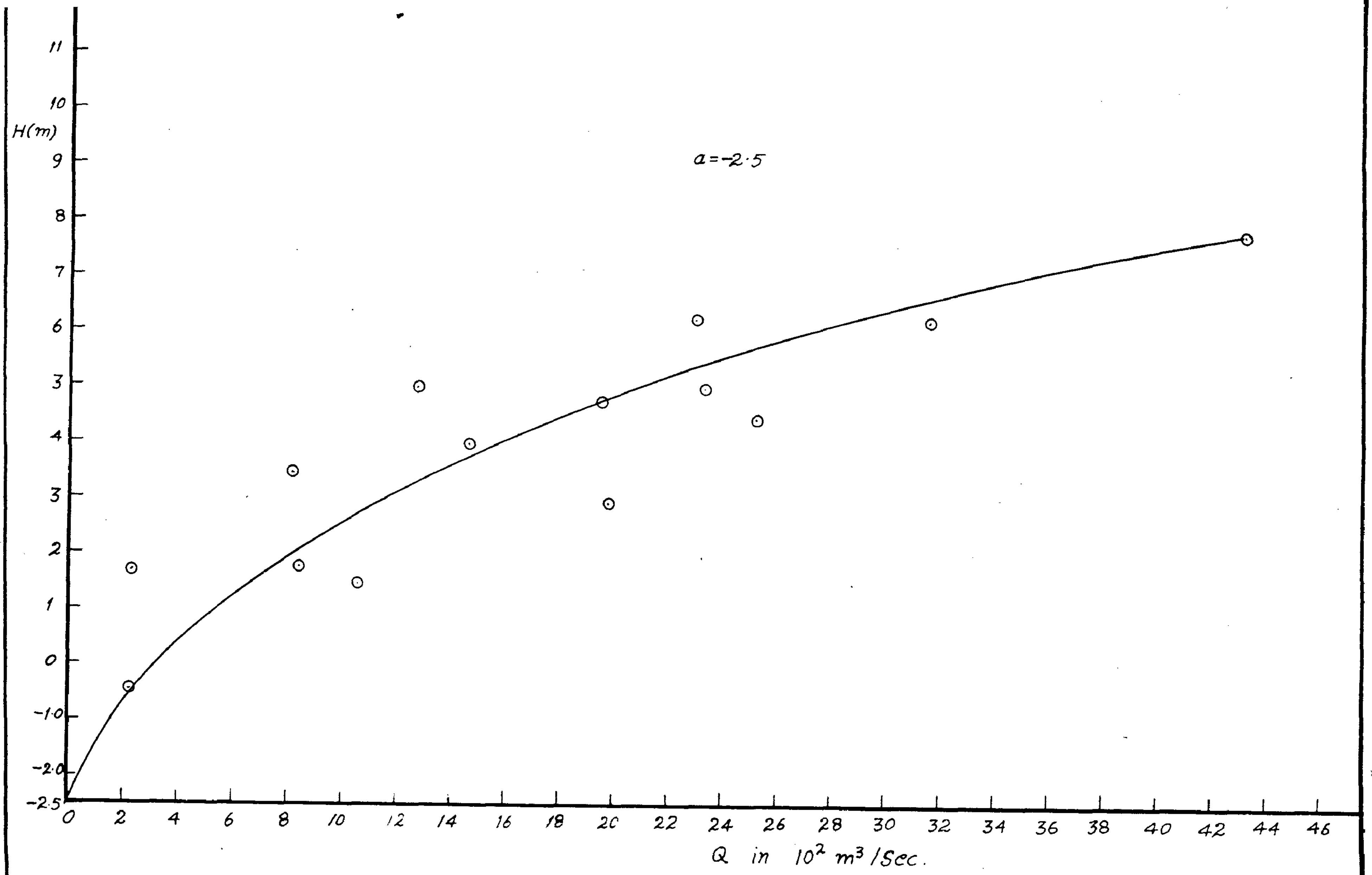

FIG. 1-3-1 RATING CURVE FOR 1K3 FROM 1961 - 1972

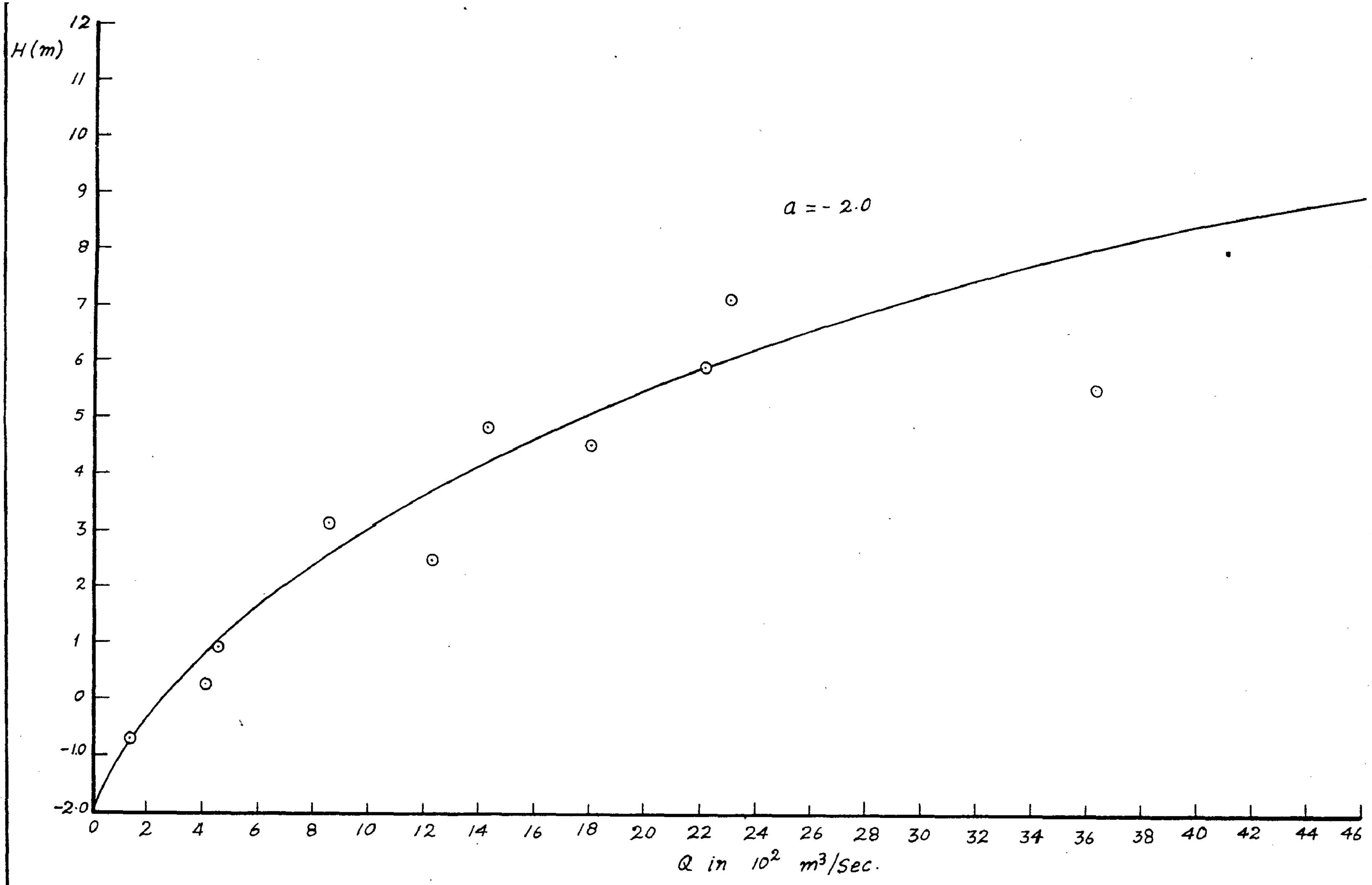

FIG. 1-3-2 RATING CURVE FOR 1K3A FROM 1961-1972

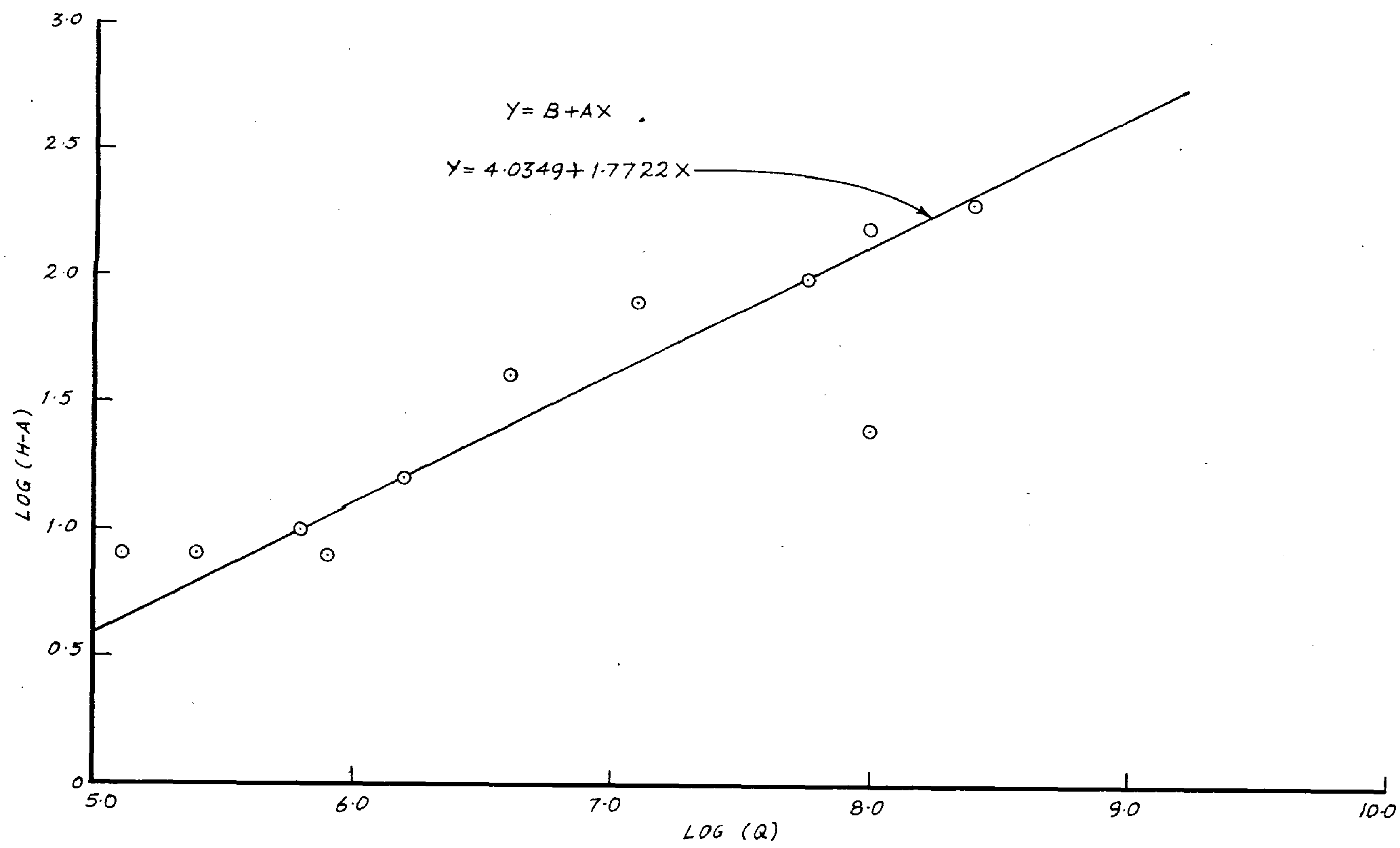

FIG. I-3-3 CURVE FITTING FOR IK 3 GAUGING STATION

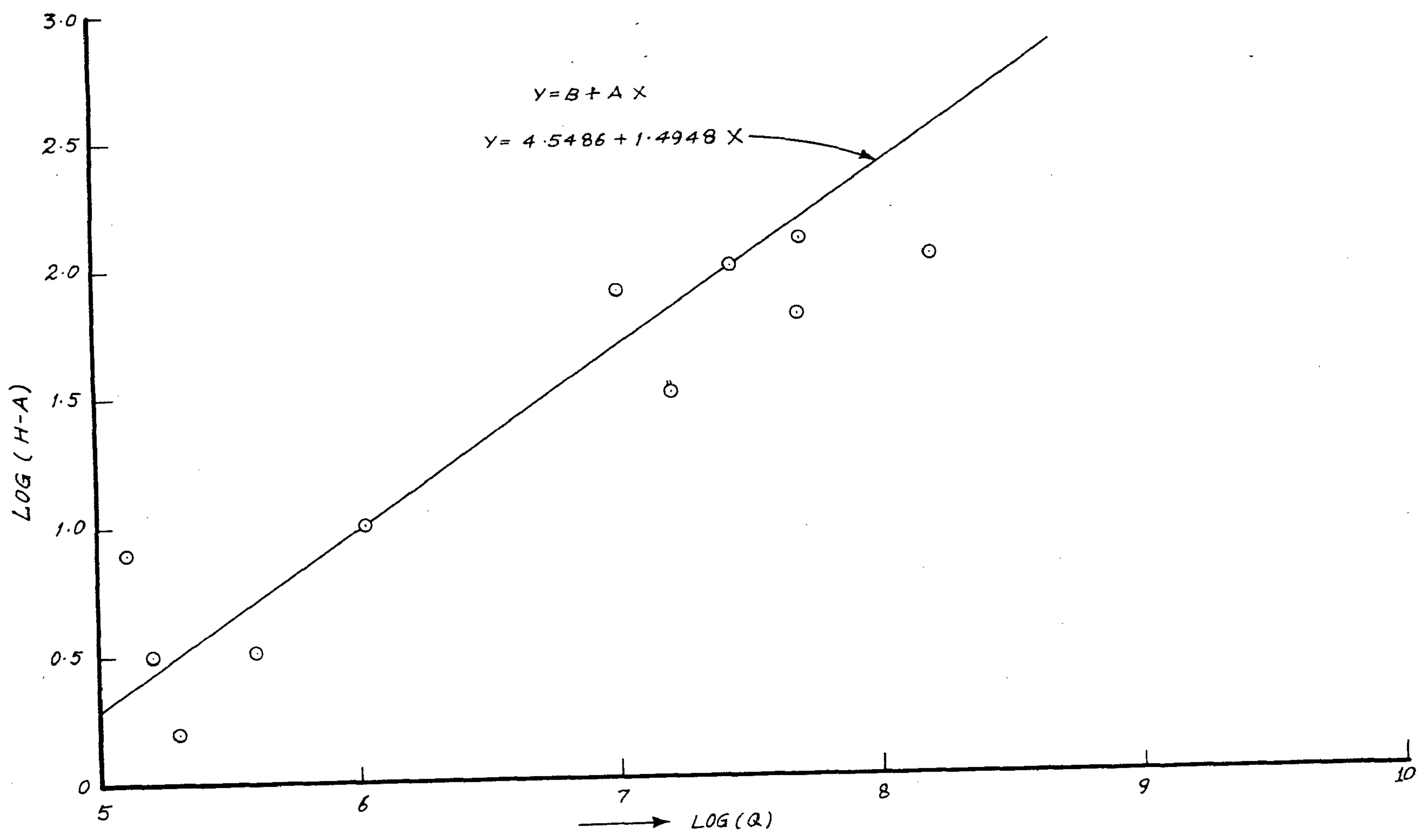

FIG. 1-3-4 CURVE FITTING FOR IK3A GAUGING STATION

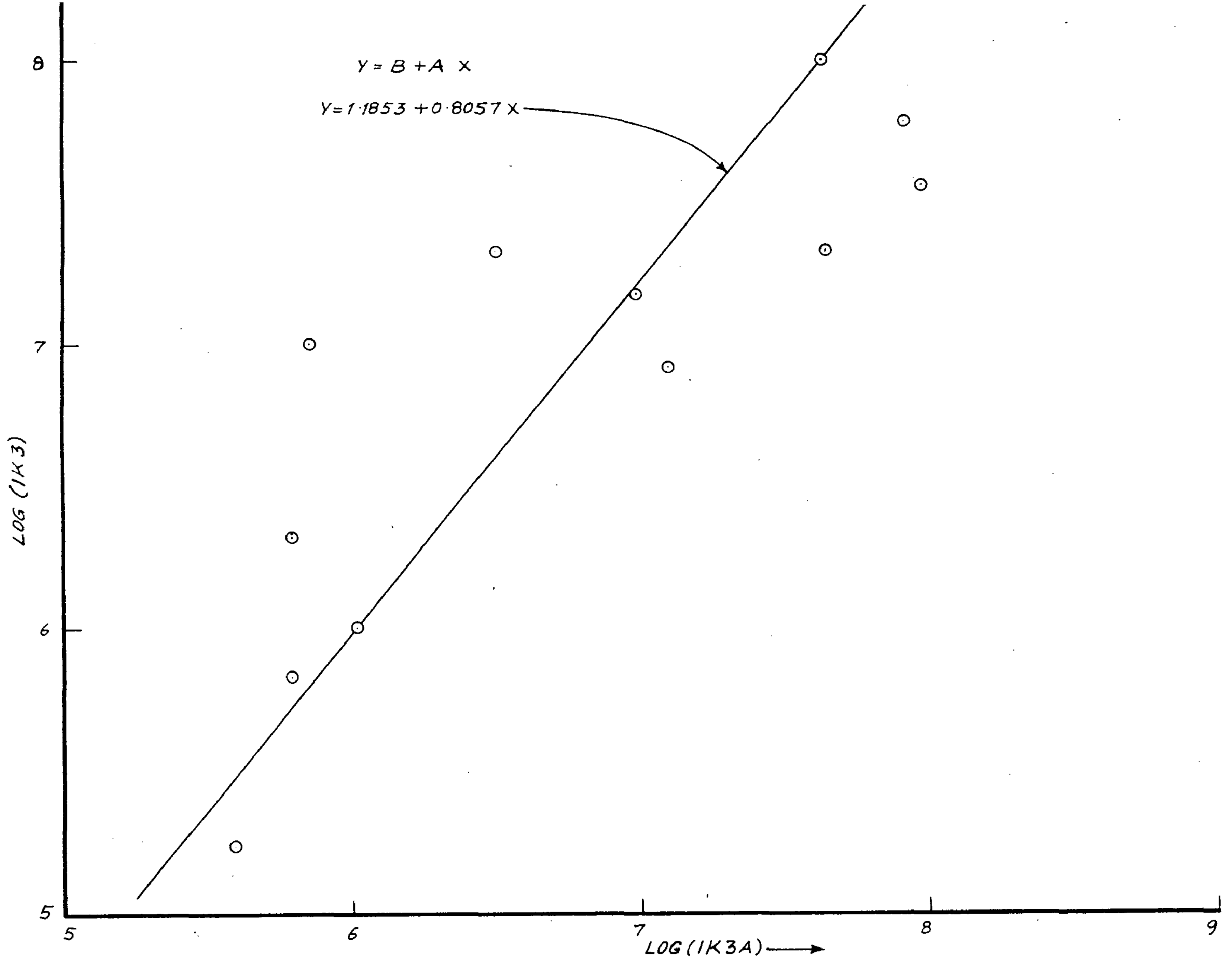

FIG. I-3-5 CORRELATION CURVE BETWEEN IK3 & IK3A

In order to overcome the problem of inadequate historical data, the approach, preferred is to investigate the process by which natural observed stream flow sequences are generated by mathematical simulation of the phenomena which provides a means of predicting future values. Such generated stream flows are extremely useful for planning water Resources Systems and thus a fairly balanced design can be evolved by subjecting the system to equally likely sets of inflows.

1.4 OBJECTIVES OF THE PRESENT STUDY

Most models evolved for generation of data, so far, evolve large number of parameters requiring a lot of computation work. In the present study an attempt has been made to evolve a model, with reduced number of parameters which can be used to generate s-tream flow by modelling of Time series of monthly stream flow sequences for 1K3A gauging station.

The main objectives of the present study are:-

(1) To estimate design flood by conventional methods.

(2) To conduct frequency analysis by probability distribution functions.

(3) To fit the best probability distribution function for 1K3 stream flows.

(4) To investigate the structure of the time series of monthly runoff for 1K3A.

(5) To condense the information contained in the time series of monthly values by formulating a mathematical model, for all the components, which can be used to generate the synthetic sequences.

However, synthesised data cannot be more precise than the original data. In all analysis based on statistical studies, there may be many inherent short-comings in the data used. These are mainly due to observational error, sampling errors and non-homogeneities.

1.5 STATEMENT OF THE PROBLEM

In this dissertation , an attempt is made towards estimation of run-off for 1K3 and construction of a mathematical model for generating of stream flow sequences for 1K3A. The specific problems treated in this dissertation can be, therefore stated as follows :

(1) Discussion of various conventional methods available for estimation of run-off and to arrive at a suitable methods. In this connection frequency analysis, using various methods, has been discussed. Flood magnitudes of various return period have been estimated using Gumbel Method and log-normal (Chow Factor Method) method.

(2) Flood frequency analysis by probability distribution functions, is conducted and best fit distribution

arrived at . In this case, log-normal 2 distribution has been found out to fit best to the annual series for 1K3 gauging station.

(3) The structure of the time series of monthly run-off for 1K3A gauging station is investigated. The information contained so as to be able to formulate a mathematical model, which will be used to generate the synthetic sequences.

CHAPTER II

FLOOD FREQUENCY STUDIES FOR 1K3 RUN- OFF DATA

2.1 GENERAL

Frequency studies interpretent a past record of events to predict the future probabilities of occurrence If stream flow records are of sufficient length and reliability, a satisfactory estimate can be achieved. However, in most cases, the records are of short length of time. Such records when analysed are likely to lead to inconsistent or incorrect results as they are not representative of long term trend. In addition to this, for the estimation of flood flows of large return periods, it is always necessary to extrapolate the magnitudes outside the observed range. Obviously, the accuracy of estimates reduces with the degree of extrapolation.

In applying statistical analysis methods, it is assumed that occurrences are individual events i.e. independent of each other, the factors influencing the character of each even remain unaltering and the measurement technique and the site of observation are identical. As a preliminary step, the basic data should be screened and adjusted to remove , as far as possible, any non-confirmities that may exist, In such, the following are the more important considerations:

(a) Effect of man-made changes in the regime of flow should be investigated and adjustment made as required.

(b) For small catchment areas a distinction should be made between daily maximum instantaneous or momentary flood peaks.

(c) Changes in the stage discharge relation render stage records non-homogeneous and unsuitable for frequency studies. It is therefore preferable to work with discharges and if stage frequencies are required, refer the results to the most recent rating.

(d) Any useful information contained in the data publications and manuscripts should be made use of after proper scrutiny.

The annual series (which is a convenient for the purpose of statistical analysis), commonly used, is a selection of the maximum event of a particular year even though this may be higher than the maximum of some other year.

2.2 PROBABILITY DISTRIBUTIONS

There are many probability distributions that have been found to be useful for hydrologic frequency analysis. The most commonly used are :

2.2.1 Normal Distribution

This is a symmetrical bell-shaped, continuous distribution theoretically representing the distribution of accidental errors about their mean or so called Gaussian law of errors. The probability density is

$$P(x) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-(X-\mu)^2/2\sigma^2} \qquad \ldots(2.1)$$

where X is the variate, μ is the mean value of the variate and σ is the standard deviation. In this dis-tribution the mean , mode and median are the same. The total area under the distribution is equal to 1.0 . The cumulative probability of a value being equal to or less than X is

$$P(X \quad X) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{X} e^{-(X-\mu)^2/2\sigma^2}\, dx \qquad \ldots(2.2)$$

This represents the area under the curve between the variate of $-\infty$ and X . Areas for various values of X have been calculated by statisticians, and tables for such areas are available in many text books and handbooks on statis-tics.

2.2.2 Poisson Distribution

If N is large and P is very small so that pN = m is a positive number, then,

$$P(X) = \frac{m^x e^{-m}}{x!} \qquad \ldots(2.3)$$

gives a close approximation to binomial probabilities when m is small. A distribution with this probability density is called the Poisson distribution and is generally referred to as the law of small numbers. It is most useful when neither N nor p is known but their product pN is given or can be estimated. The statistical parameters are: Mean = m , standard deviation = m and skewness = $1/\sqrt{m}$

2.2.3 Binomial Distribution

This is one of the most commonly used discrete distributions. It repres nts the distribution of probabilities in Binomial trials, say tossing a coin. The probability density is

$$P(X) = C_x N P^x q^{N-x} \qquad ...(2.4)$$

where P is the probability of occurrence of an event, for example, a success in tossing a coin, $C_x N$ is the number of combinations of N things taken X at a time q is the probability of failure or 1-p , N is the total number of trials, and X is the variate or the number of successful trials.

The statistical parameters are Mean = pN, standard deviation , $\sigma = \sqrt{pq\,N}$, and skewness $\alpha = \mu_3 / \sigma^3 = (q-p) / \sqrt{pqN}$, where μ_3 is the third moment about the mean. When p=q , the distribution is symmetrical.

In binomial distribution, the events or trials can be classified into only two categories: success and failure, yes or no, rainy and clear, etc. The probabilities P and q remain constant from one trial to another, i.e., the events are in-dependent to each other.

2.2.4 Gamma Distribution

The probability density of this distribution is

$$P(x) = \frac{x^a \, e^{-x/b}}{b^{a+1}\Gamma(a+1)}$$

with $b > 0$, $a > 1$ for $x = 0$

and $p(x) = 0$ for $x \leqslant 0$ ...(2.5)

where a and b are constants and $\Gamma(a+1) = a$ is a gamma function. The cumulative probability being equal to or less than $x(< \infty)$ is known as the incomplete gamma function. The statistical parameters are :

Mean $= b(a+1)$ and variance $= b^2(a+1)$.

2.2.5 Rectangular Distribution

The rectangular distribution is a uniform distribution of a continuous variable X between two constants a and b. The probability density of this distribution is

$P(x) = 0$ for $x < a$

$P(x) = \frac{1}{b-a}$ for $a \leq x \leqslant b$ ...(2.6)

and $P(x) = 0$ for $b < x$

The statistical parameters are Mean = (b+2)/2 and variance = $(b+a)^2/12$.

2.2.6 Extremal Distribution (Type I distribution)

This distribution results from any initial distribution of exponential type which converges to an exponential function as X increases. Example of such initial distributions are the normal, the chi-square, and the log-normal distributions. The probability density of this distribution is

$$P(x) = \frac{1}{c} e^{-(a+x)/C-e^{-(a+x)/c}} \quad \ldots(2.7)$$

with $-\infty < x < \infty$, where x is the variate, and a and c are parameters. The cumulative probability is

$$P(X \leqslant x) = e^{-e^{-(a+x)/C}} \quad \ldots(2.8)$$

By the method of moments, the parameters have been evaluated as

$$a = \gamma C - \mu \quad \ldots(2.9a)$$

$$C = \frac{\sqrt{6}}{\pi} \sigma \quad \ldots(2.9b)$$

where $\gamma = 0.57721$ - Euler's constant, μ is the mean, and σ is the standard deviation. The distribution has a constant coefficient of skewness equal to $C_s = 1.139$.

2.2.7 Logarithmically Transformed Distribution, the Log normal distribution.

This is transformed normal distribution in which the variate is replaced by its logarithmic value. This distribution represents the so called 'law of Galton' because it was first studied by Galton as early as 1875. Its probability density is

$$P(x) = \frac{1}{\sigma_y e^y \sqrt{2\pi}} e^{-(y-\mu_y)^2/2\sigma_y} \qquad \ldots(2.10)$$

where $y = \ln x$, x is a variate, μ_y is the mean of y and σ_y is the standard deviation of y. This is a skew distribution of unlimited range in both directions. Chow[ ] as derived the statistical parameters for x as

$$\mu = e^{\mu_y + \sigma_y^2/2} \qquad \ldots(2.10a)$$

$$\sigma = \mu\,(e^{\sigma_y^2} - 1)^{1/2} \qquad \ldots(2.10b)$$

$$\alpha = (e^{3\sigma_y^2} - 3\,C^{\sigma_y^2} + 2)\,C_v^3 \qquad \ldots(2.10c)$$

$$M = e^{\mu_y} \qquad \ldots(2.10d)$$

$$\frac{\mu}{M} = e^{\sigma_y^2/2} \qquad \ldots(2.10e)$$

$$C_v = (e^{\sigma_y^2} - 1)^{1/2} \qquad \ldots(2.10f)$$

$$C_s = 3\,C_v + C_v^3 \qquad \ldots(2.10g)$$

Where μ is the mean, σ is the standard deviation, C_s is the coefficient of skewness, M is the median and C_v is the coefficient of variation. Chow has also shown the type I extremal distribution is essentially a special case of the log normal distribution when $C_v = 0.364$ and $C_s = 1.139$.

2.2.8 Pearson Distribution

Karl Pearson has derived a series of probability functions of fit virtually any distribution . These functions have been used widely in practical statistical works to define the shape of many distribution curves, though they have only slight theoretical basis. The general and basic equation to define the probability density of a Peasson distribution is

$$P(x) = \int_{e^{-\infty}}^{x} (a+ \quad) / b_0 + b_1X + b_2X^2) dX \qquad ...(2.11)$$

where a, b_0, b_1 and b_2 are constants. The criteria for determining types of distribution are β_1, β_2 and K being defined as:

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} \qquad ...(2.12)$$

$$\beta_2 = \frac{\mu_4}{\mu_2^2} \qquad ...(2.13)$$

$$K = \frac{\beta_1 \; (\beta_2+3)^2}{4(4\;\beta_2 - 3\beta_1)(2\;\beta_2 - 3\beta_1 - 6)} \qquad ...(2.14)$$

Where μ_2, μ_3 and μ_4 are second third and fourth moment about the mean.

With $\beta_1 = 0$, $\beta_2 = 3$ and K = 0 the resulting Pearson distribution is identical with the normal distribution. Types I and III distributions are often used in hydrologic frequency analysis.

2.2.8.1 Type I Distribution - For type I, $K < 0$. This is a skew distribution with limited range in both directions usually bell-shaped but may be J-shaped or V-shaped. Its probability density is

$$P(X) = P_o\left(1 + \frac{X}{a_1}\right)^{m_1}\left(1 - \frac{X}{a_2}\right)^{m_2} \qquad \ldots(2.15)$$

with $m_1/\alpha_1 = m_2/X_2$ and the origin at the mode. The values of m_1 and m_2 are given by

$$m_1 \text{ or } m_2 = \frac{1}{2}\left[r - 2 \pm r(r+2)\frac{\sqrt{\mu_2\,\beta_1}}{2(a_1+a_2)}\right] \qquad \ldots(2.15a)$$

when μ_3 is positive, m_2 is the positive root and m_1 is the negative root and vice-versa in signs. The other values are ,

$$r = \frac{6(\beta_2 - \beta_1 - 1)}{6 + 3\,\beta_1 - 2\,\beta_2} \qquad \ldots(2.15b)$$

$$a_1+a_2 = \frac{1}{2}\sqrt{\mu_2\left[\beta_1\,(r+2)^2 + 16(r+1)\right]} \qquad \ldots(2.15c)$$

and

$$P_o = \frac{N}{a_1+a_2}\;\frac{m_1^{m_1}\;m_2^{m_2}}{(m_1+m_2)^{m_1+m_2}}\;\frac{\Gamma(m_1+m_2+2)}{\Gamma(m+1)\;\Gamma(m_2+1)}$$

$$\ldots(2.15d)$$

where N is the total frequency . The statistical parameters are :

Mean = mode = $(\mu_3/2\mu_2)\left[(r+2)/ r-2)\right]$ and

standard derivation = $\sqrt{\mu_2}$, and Pearson's s-kewness

$= (\beta_1/2)\left[(r+2)/(r-2)\right]$

2.2.8.2 Type II Distribution : For type II , K = ∞ or $2\beta_2 = 3\beta_1 + 6$. This is a skew distribution with limited in the left direction, usually bell-shaped but be J-shaped. Its probability density with the origin at the mode is

$$P(x) = P_o\left(1 + \frac{X}{a}\right)^c e^{-cx/a} \qquad \ldots(2.16)$$

where,

$$C = \frac{4}{\beta_1} - 1 \qquad \ldots(2.16a)$$

$$a = \frac{c}{2} \quad \frac{\mu_3}{\mu_2} \qquad \ldots(2.16b)$$

$$P_o = \frac{N}{a} \quad \frac{C^{c+1}}{e^c \, \Gamma(c+1)} \qquad \ldots(2.16c)$$

The statistical parameters are :

Mean = mode = $\mu_3 / 2\mu_2$,

Standard deviation = $\sqrt{\mu_2}$, and Pearson's skewness= $\sqrt{\beta_1}/2$

2.3 METHODS OF CURVE FITTING

The methods of frequency analysis are all based on the assumption that observed data follow the theoretical distribution to be fitted and will exhibit a straight line on the probability paper designed for the distribution. In as much as nature does not strictly obey the theoretical laws, the logical solution is to plot the observed data at determined plotting positions on a suitable probability paper and fit a best fit curve to the plotted points. Curve fitting may be done either mathematically or graphically. In general a mathematical curve fitting can be achieved by three methods: the method of least squares, the method of likelihood and the methods of moments.

2.3.1 Least Square Method

This method gives a better overall fit than the method of moments and involves relatively less computations and therefore is commonly adopted to avoid the subjective errors in graphical fitting. A brief outline of the principle of least squares and a procedure for fitting Gumbel's distribution using this principle are described here under.

In Figure 2.3.1 for a given value of x, say x_1 there will be a difference between the value y_1, and the corresponding value as determined from the curve. This difference (indicated as D in the figure) or the departure may be positive, negative or zero. A measure of the 'goodness of a fit of the curve to the given data is provided by the sum of the squares of departures, if this is small the fit is good and if large it is bad. The least square line approximating the set of points (X_1,Y_1) ,(X_2, Y_2) (X_3,Y_3) , ...(X_n, Y_n) has the equation Y = A+XB where the constant A and B are determined by solving simultaneously the equation s (2.14)

$$\sum Y = A.N. + B\sum X$$

and $$\sum XY = A\sum X + B\sum X^2$$

which are called normal equations for the least square line. From these equations the constant A and B can be found out as

$$A = \bar{Y} - B\bar{X} \quad \text{and}$$

$$B = (\sum XY - N \sum \bar{X}\bar{Y}) / (\sum X - N\sum \bar{X}^2)$$

where

$$\bar{X} = \frac{\sum X}{N} \quad \text{and} \quad \bar{Y} = \frac{\sum Y}{N}$$

2.3.2 Methods of Maximum Likelihood

This method gives the best estimates and by this method , the value of a parameter is determined to make the probability of obtaining the observed outcome as high as possible. Mathematically, $\partial \log P(X)/ \partial\mu = 0$,

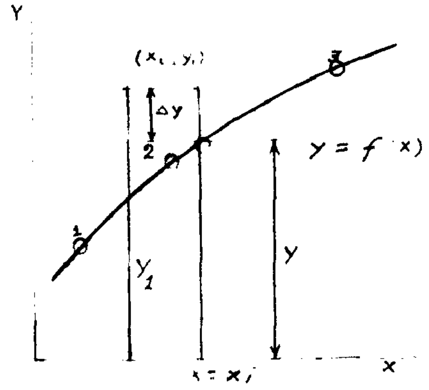

FIG. 2-3-1 SCHEMATIC REPRESENTATION OF SIMPLE REGRESSIO AND CORRELATION ANALYSIS.

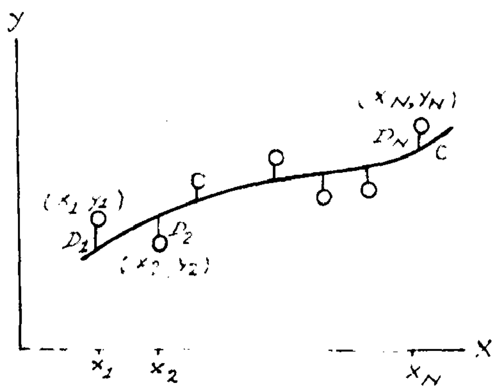

FIG 2-3-2 LEAST SQUARES PRINCIPAL

where p(X) is probability density and μ is a statistical parameter. This method provides the best estimate of the parameter but is usually very complicated for practical application Kimball has suggested this method for fitting extremal distributions and a practice procedure was later developed by Panchang and Aggarwal (4).

2.3.3. Methods of Moments

In this method, the statistical parameters or moments are computed from the data and then substituted in the probability function of the given distribution. This method gives a theoretically exact fitting but the accuracy can be substantially affected by any errors involved in the data at the tails of the distribution where the moments arms are long and the errors are thus magnified. The method originally proposed by Gumbel to fit type I extremal distribution is a method of moments. Liebleim modified this method by orders statistics and developed a procedure which maintains the original time order of the extreme value series, divides the values into subgroups, and then weights each observation according to its ordered rank in the subgroup which in turn is a function of the sample size. Hershfield made a comparison of the two procedures and concluded that the Gumbel procedure gives a better estimate beyond the range of data for the really independent data tests, but overestimates the longer recurrance-intervals in the dependent data tests.

2.3.4 Graphical Curve - Fitting Procedure

In a simple graphical curve-fitting procedure the observed floods are plotted on a probability paper and a best-fit curve drawn by eye through the points. Log-normal probability paper and extreme value probability paper are commonly used for the purpose. In the case of the former, the plotting positions of the individual floods of the annual series is found by the formula P = M/(N+1) where P is the exceedance probability M the order of magnitude of a given flood in an array of observed floods and N the number of years. If extreme values probabilit-y paper, also called Gumbel paper, is used, the plotting positions of the flood are found by the formula T = (N+1)/M , where T is the return period in years.

For determining the confidence bands firstly the wrong type of the theoretical distribution may have been used. The guide to this is the fit of the observed data. Secondly there may be errors due to sampling. It is therefore necessary to assign limits between which the estimated calue can be said to lie with a certain probability or confidence. The curves joining the equal confidence limits are drawn to show the confidence bands on both sides of the fitting curve. The reliability of any plotted point lying within the confidence band is thus indicated by the probability on which the confidence limits are based.

The regression and correlation analysis is one of the oldest statistical tools used in hydrology. Now its application has been broadened to cover the study of the relationship between two or more hydrologic variables and also the investigation of dependence between the successive values of a series of hydrologic data (26).

If two variables, given as a series with concurrent values (X_i, Y_i), show a concentration around an imaginary curve when plotted on a graph (Fig.2.3.2) then for a large series there will always be a distribution of y values for a given value of X_i or more precisely a distribution of y values for a given interval ΔX around X_i. The mean value y_o of all y values of this given interval ΔX around X_i is the expected value of y for the given $X = X_i$. A curve fitted to all mean values y_o, is called the regression line of y versus X. On the other hand, the curve fitted to all expected (mean) values, X_o, the given $y = y_i$, defines the regression line of X versus Y. These two lines do not coincide, but have different parameters, showing the regressional relationships between the variables.

A pure functional relationship between variables assumes that all points would follow a curve, without spread. In as much as the spread of points around the regression lines may actually be great or small, the degree of association of the availables involved is generally called correlation and is defined by the parameters of correlation. The correlation is greater when the points are closer to the lines.

Briefly, a regression problem considers the frequency distribution of one variable when another is held fixed at each of the several levels. A correlation problem considers the joint variation of two measurements, neither of which is restricted by the experimental or observer.

2.4 FREQUENCY ANALYSIS BY FREQUENCY FACTORS FOR ESTIMATION OF PEAK RUNOFF

These methods employ the general equation for hydrologic frequency analysis which may be expressed as $X = \bar{X} + K.S_x$ where X is the magnitude of flood of some given probability (P) or return period (T), $\bar{X}$ is the mean of floods of record, S_x is standard deviation and K is a freque cy factor. For the two distribution viz. log-normal and Gumbel, usually proposed for the purpose of analysis, the tables showing theoretical derived values of the factor K for selected values of probability or the recurrance interval) are furnished see Table 2.4.1 and 2.4.2 . In the case of other distributions the value of (K) should be known for determining the magnitude of flood.

It may be noted that in these methods it is not necessary to plot the observed data. Yet this may be done

TABLE 2-4-1 FR QUENCY FACTORS (K) FOR GUMBEL METHOD

Sample Size - N Return Period - Years

N	5	10	15	20	25	30	50	60	75	100	1000
1	2	3	4	5	6	7	8	9	10	11	12
15	0.967	1.703	2.117	2.410	2.632	2.823	3.321	3.501	3.721	4.005	6.265
20	0.919	1.625	2.023	2.302	2.302	2.690	3.179	3.352	3.563	3.836	6.006
25	8.888	1.575	1.963	2.235	2.444	2.614	3.088	3.257	3.463	3.729	5.948
30	0.866	1.541	1.922	2.188	2.393	2.560	3.026	3.191	3.393	3.653	5.227
35	0.0851	1.516	1.891	2.152	2.354	2.520	2.479	3.142	3.341	3.598	
40	0.838	1.495	1.866	2.126	2.326	2.489	2.943	3.104	3.301	3.554	5.576
45	0.829	1.478	1.847	2.104	2.303	2.464	2.913	3.078	3.268	3.520	
50	0.820	1.466	1.831	2.086	2.283	2.443	2.889	3.048	3.241	3.491	5.478
55	0.813	1.455	1.818	2.071	2.267	2.426	2.869	3.027	3.219	3.467	-
60	0.807	1.446	1.806	2.059	2.253	2.411	2.852	3.008	3.200	3.446	
65	0.801	1.437	1.796	2.048	2.241	2.398	2.837	2.992	3.183	3.429	
70	0.797	1.430	1.788	2.038	2.230	2.387	2.824	2.979	3.169	3.413	5.359
75	0.792	1.323	1.780	2.029	2.220	2.377	2.814	2.967	3.155	3.400	
80	0.788	1.417	1.773	2.020	2.212	2.368	2.802	2.956	3.145	3.387	
85	0.785	1.413	1.767	2.013	2.205	2.361	2.793	2.946	3.135	3.376	
90	0.782	1.409	1.762	2.007	2.698	2.353	2.785	2.938	3.125	3.367	
95	0.780	1.405	1.757	2.002	2.193	2.347	2.777	2.930	3.116	3.357	
100	0.779	1.401	1.752	1.998	2.187	2.341	2.770	2.922	3.109	3.349	5.261

TABLE 2-4-2 CHOW FREQUENCY FACTORS - THEORETICAL LOG PROBABILITY FREQUENCY FACTORS -K

Coeff. of skew C_s	Probablity at mean	Probability (in %) equal or greater than the given variate									Corresponding coeff. of variation C_v
		99	95	80	50	20	5	1	0.1	0.01	
1	2	3	4	5	6	7	8	9	10	11	12
0.0	50.0	2.33	1.55	0.84	0.0	0.84	1.64	2.33	3.09	3.72	0.000
0.1	49.3	2.25	1.62	0.085	0.02	0.84	1.67	2.40	3.22	3.95	0.033
0.2	48.7	2.18	1.59	0.85	0.04	0.83	1.70	2.47	3.39	4.18	0.067
0.3	48.0	2.11	1.56	0.85	0.06	0.82	1.72	2.55	3.56	4.42	0.100
0.4	47.3	2.04	1.53	0.85	0.37	0.81	1.75	2.62	3.72	4.70	0.136
0.5	46.7	1.98	1.49	0.85	0.09	0.80	1.77	2.70	3.88	4.96	0.166
0.6	46.1	1.91	1.46	0.85	0.10	0.79	1.79	2.77	4.05	5.24	0.197
0.7	45.5	1.85	1.43	0.85	0.11	0.78	1.81	2.04	4.21	5.52	0.230
0.8	44.9	1.79	1.40	0.84	0.13	0.77	1.82	2.90	4.37	5.81	0.262
0.9	44.2	1.74	1.37	0.84	0.14	0.76	1.84	2.97	4.55	6.11	0.292
1.0	43.7	1.68	1.34	0.84	0.15	0.75	1.85	3.03	4.72	6.40	0.324
1.1	43.2	1.63	1.31	0.83	0.16	0.73	1.86	3.09	4.87	6.71	0.357
1.2	42.7	1.58	1.29	0.82	0.17	0.72	1.87	3.15	5.04	7.02	0.381
1.3	42.2	1.54	1.26	0.82	0.18	0.71	1.88	3.21	5.19	7.31	0.409
1.4	41.7	1.49	1.23	0.81	0.19	0.69	1.88	3.26	5.35	7.62	0.436
1.5	41.3	1.45	1.21	0.81	0.20	0.68	1.89	3.31	5.51	7.92	0.462
1.6	40.8	1.41	1.18	0.80	0.21	0.67	1.89	3.36	5.66	8.26	0.490

TABLE 2-4-2 (Contd...)

1	2	3	4	5	6	7	8	9	10	11	12
1.7	40.4	1.38	1.16	0.79	0.22	0.65	1.89	3.40	5.80	8.58	0.517
1.8	40.0	1.34	1.14	0.78	0.22	0.64	1.89	3.44	5.96	8.88	0.544
1.9	39.6	1.31	1.12	0.78	0.23	0.63	1.89	3.48	6.10	9.20	0.570
2.0	39.2	1.28	1.10	0.77	0.24	0.61	1.89	3.52	6.25	9.51	0.596
2.1	38.8	1.25	1.08	0.76	0.24	0.60	1.89	3.55	6.39	9.79	0.620
2.2	38.4	1.22	1.06	0.76	0.25	0.59	1.89	3.59	6.51	10.12	0.643
2.3	38.1	1.20	1.04	0.75	0.25	0.58	1.88	3.62	6.55	10.43	0.667
2.4	37.7	1.17	1.02	0.74	0.26	0.57	1.88	3.65	6.77	10.72	0.691
2.5	37.4	1.15	1.00	0.74	0.26	0.56	1.88	367	6.90	10.95	0.713
2.6	37.1	1.12	0.99	0.73	0.26	0.55	1.87	3.70	7.02	11.25	0.744
2.7	36.8	1.10	0.97	0.72	0.27	0.54	1.87	3.72	7.83	11.55	0.755
2.8	36.6	1.08	0.96	0.72	0.27	0.53	1.86	3.74	7.25	11.80	0.776
2.9	36.3	1.06	0.95	0.71	0.27	0.52	1.86	3.76	7.36	12.10	0.796
3.0	36.0	1.04	0.93	0.71	0.28	0.51	1.85	3.78	7.47	12.36	0.818
3.2	35.5	1.01	0.90	0.69	0.28	0.49	1.84	3.81	7.65	12.85	0.857
3.4	35.1	0.98	0.88	0.68	0.29	0.47	1.83	3.84	7.84	13.36	0.895
3.6	34.7	0.95	0.86	0.67	0.29	0.46	1.84	3.87	8.00	13.83	0.9300
3.8	34.2	0.92	0.84	0.66	0.29	0.44	1.80	3.89	8.16	14.23	0.966
4.0	33.9	0.90	0.82	0.65	0.29	0.42	1.78	3.81	8.30	14.70	1.000
4.5	33.0	0.84	0.78	0.63	0.30	0.39	1.75	3.83	8.60	15.62	1.081
5.0	32.3	0.80	0.74	0.62	0.30	0.37	1.71	3.95	8.86	16.45	1.155

for comparison purposes i.e. to see how closely the estimated frequency line fits to the observed data.

2.4.1 The Log Normal Method Us-ing Chow's Frequency Factors

This method is based on the log-normal probability law and assumes that the flood are so distributed that their natural logs are normally distributed.

2.4.2. The Gumbel Method

Gumbel was the first to appreciate that the annual peak flood data (or the maximum storm rainfall and similar types of data) are nothing but the extreme values in different years observations and hence they should follow the extreme values distribution law. This form of distribution law with a bearing on the nature of the data is accepted as best suited for the frequency analysis.

2.5 BEST FIT DISTRIBUTION

2.5.1 Selection Criteria

According to properties of observed data, the theoretical distribution functions of best fit to observed distributions of annual precipitation and annual runoff should have the following characteristics: (1) The function is continuous and defined for all positive values of the observed variable K, (2) the lower tail is bounded by zero value or by a positive value K_o, (3) the upper tail is unbounded, (4) the density curve is asymptotic to the axis for large values of K, (5) the basic shape is one peak bell-shaped two tailed curve, with a large variety of skewness, and (6) the number of parameters which describe theoretical functions is limited to three .

2.5.2 Selected functions

Screening of the applicable functions with respect to the criteria required, their convenience for use in mass computation and the experience already obtained in applying them in hydrology leads to the selection of

(1) Normal density, function, or Normal

(2) Gumbel density function or Gumbel

(3) Log-normal density function with two parameters or log-normal 2

(4) Gamma dens-ity function with two parameters or Gamma 2

The expressions and parameters for these functions are:

(1) Normal with the classical form

$$f(K) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(K-\mu)^2}{2\sigma^2}}, \quad -\infty \leq K \leq +\infty \qquad ...(2.17)$$

with K the variable values, μ the population mean and σ = the population standard deviation.

(2) Gumbel with the form

$$f(K) = \frac{1}{C} e^{-\frac{(a+K)}{C}} \times \left[e^{-\frac{(a+k)}{c}} \right] \qquad ...(2.18)$$

with K the variate, a and c are parameters.

(3) Log-normal 2 with the form

$$f(K) = \frac{1}{K\sigma\sqrt{2\pi}} e^{-\left[\frac{\ln K - \ln\mu}{2\sigma^2}\right]^2}$$

$$0 \leqslant K \leqslant \infty \qquad \ldots(2.19)$$

With μ the population mean and σ the population standard deviation of the ln K values.

(4) Gamma 2 with the form

$$f(K) = \frac{1}{\beta^{\alpha}\,\Gamma(\alpha)} K^{\alpha-1} e^{-K/B} \qquad \ldots(2.20)$$

with α - the shape parameter

β - the scale parameter and $\Gamma(\alpha)$ the gamma function of α. It is skewed to the right for all values of parameters α and β.

According to R.A. Fisher, the Maximum likelihood method is based upon likelihood function L. This function is maximised by setting the first derivative of luk with respect to θ equal to zero and solving the resulting equation for θ

$$\frac{\partial(\ln L)}{\partial\theta} = \frac{\partial\left\{\sum_{i=1}^{n} \ln\left[f(K_i,\theta)\right]\right\}}{\partial\theta} = 0 \qquad \ldots(2.21)$$

where $K_i = \frac{Q_i}{\bar{Q}}$, $\bar{Q}$ denotes the sample mean and Q_i is the annual observed river flow. This yields a single equation for the solution of θ in terms of K's. For m parameters, m equations of equn. (2.21) gives m estimations or known parameters. Maximum likelihood estimators are consistent, asymptotically normal and asymptotically efficient under general conditions. The method is completely numerical , applicable to all selected functions and convenient for mass computation. The maximum likelihood method gives the following equations for parameter estimators.

Normal - Based on Eqn (2.21) and the concept of Eqn(2.25) the maximum likelihood function produces:

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^{n} K_i \qquad \ldots(2.22)$$

as estimator of the population mean, and

$$\hat{\sigma} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (K_i) - \hat{\mu})^2} \qquad \ldots(2.23)$$

as estimator of population standard deviation.

Gumbel: According to eqn (2.22) and the concept of Eqn (2.25) , the maximum likelihood function produces

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^{n} K_i \qquad \ldots(2.24)$$

as estimator of the population mean and

$$\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (K_i - \hat{\mu})^2} \qquad \ldots(2.25)$$

as the estimator of the population standard deviation.

Log-normal 2 According to eqn (2.18) and using the maximum likelihood equation, the maximum likelihood estimator of the population mean is :

$$\ln \hat{\mu} = \frac{1}{n} \sum_{i=1}^{n} \ln K_i \qquad ...(2.26)$$

and the estimator of the population standard deviation

$$\hat{\sigma} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (\ln K_i - \ln \hat{\mu})^2} \qquad ...(2.27)$$

Gamma 2 - According to Eqn (2.20) the maximum likelihood equation gives the two maximum likelihood partial differential equations for parameters α and β_1, and from them it follows

$$\ln \hat{\alpha} - \frac{\partial \left[\ln \Gamma(\hat{\alpha})\right]}{\partial \hat{\alpha}} = \ln \bar{K} - \frac{1}{n} \sum_{i=1}^{n} \ln K_i$$

with $\hat{\alpha}$ the estimator of α and

$$\hat{\beta} = \frac{1}{\hat{\alpha}} \frac{1}{\hat{n}} \sum_{i=1}^{n} K_i = \frac{1}{\hat{\alpha}} \bar{K} \qquad ...(2.28)$$

with $\hat{\beta}$ the estimator of β

2.6 TEST OF GOODNESS OF FIT

To test the theoreticalprobability distribution functions for goodness of fit to observed data, as in other frequency analysis, the distribution of a exclusive and exhaustive categories or class intervals. In classifying the observed data, it is necessary to decide upon the number and the length of class intervals.

Number and class intervals to be used has no satisfactory hard-and-fast rule, howver, if too many classes are used some of them would have few or no frequencies and the resulting frequency distribution would be irregular. Likewise, if there are too few classes, the observed data would be very compressed, a large proportion of the frequencies would fall in one or two classes, and much information would be lost.

The choice of the length of class intervals should be done in such a manner that the main characteristic features of the observed distribution are emphasised and chance variations are observed. Basically, there are two concepts of choice of the length of class intervals: (a) equal length, and (b) equal probabilities .

Equal probability of class intervals, which can be considered as special case of unequal length, has some advantages over the previous method. The arbitrary s-teps for equal lengths may be avoided by choosing intervals of equal probabilities instead of intervals of equal lengths. The required intervals are obtained from the probability integral transformation. The probabilities are uniformly distributed. Thus, the comparison of the observed distribution with any conditions theoretical distribution is reduced to the comparison of an observed with a theoretical uniform distribution. According to this method, and with the fact that the total value of the probability integral is unity the probability of each class interval is determined by

$$P_j = \frac{1}{K} \quad , \text{ with } j = 1,2,\ldots,K \qquad \ldots(2.29)$$

For this value of probability, the required length of any class interval can be obtained from the probability integral transformation.

The well known and frequently applied chi-square test is used here as a measure of goodness of fit of the theoretical probability distributions to observed ones. The basic concept of chi-square test can be summarized as follows The total range of sample observations is divided into K mutually exclusive and exhaustive class intervals, each having the observed class frequency O_j and corresponding expected class probability $E_j (j=1,2,....K)$ using the expected value E_j as the norm of any class interval, it is reasonable to choos the quantity $(O_j-E_j)^2$ as a measure of departure from the norm. A suitable measure is expressed by $(O_j-E_j)^2/E_j$ and the measure of total discrepency between observations and expectations, χ^2 becomes

$$\chi^2 = \sum_{j=1}^{K} \frac{(O_j - E_j)^2}{E_j} \qquad ...(2.30)$$

This statistic is distributed asymptotically as Chi-square (χ^2) with K-1 degrees of freedom, if the population parameters have not been estimated from the sample observations. For v parameters, the total number of degree of freedom is

$$f = K - 1 - v \qquad ...(2.31)$$

As total number of class intervals is 5 and probability of each interval is the same, for given sample size, the expected class probability of any interval should be the same and independent of the type of probability function i.e., it is dependent only on sample size, n or

$$E_j = P_{jn} = \frac{n}{K} \qquad \ldots(2.32)$$

Therefore the computation of expected class probabilities is simplified by choosing the constant number of class intervals of the same probability. The sample observations should be arranged in an array in increasing order. Then to determine how many observations will fall in each of the five chosen class intervals, four class intervals limits must be computed for each of four selected functions separately.

<u>Normal</u> - In this case the class interval limits K_j of the variables K_i are

$$K_j = \hat{\mu} + U_j \hat{\sigma} \qquad \ldots(2.33)$$

in which U_j are class intervals limits of the variable U_i of equation

$$F(U) - j\, P_j = \int_{-\infty}^{U_i} \frac{1}{\sqrt{2\pi}}\, e^{-\frac{U^2}{2}}\, du \qquad \ldots(2.34)$$

with $j = 1, 2, \ldots 5$ and with the lower integral limit $-\infty$ the mean zero and variate unity. This is a well known probability integral, the value of which is generally given in terms of U_j are determined and given in Table 2-6-2.

<u>Gumbel</u> - Similarly, the class interval limits of Gumbel are

$$K_j = \hat{\mu} + V_j \hat{\sigma} \qquad \ldots(2.35)$$

in which V_j are class interval limits of the variable V_i.

<u>Log - normal 2</u> - Similar to the previous case, the class interval limits of log-normal 2 are computed by using Eqn (2.19) which is fist transformed into a normal probability integral form. The class interval limits are then computed from the expression

$$K_j = \exp\left[\ln \hat{\mu} + U_j \hat{\sigma}\right] \qquad \ldots(2.36)$$

in which K_j are class inteval limits for the variable K_i, $\ln \hat{\mu}$ is the mean of $\ln K_i$ and $\hat{\sigma}$ is the standard deviation of $\ln K_i$ while U_j are class interval limits of the variable U_i from Eqn (2.38).

<u>Gamma 2</u> - The class interval limits of Gamma 2 are -

$$K_j = \frac{U_j}{\sqrt{\hat{\alpha}}} \qquad \ldots(2.37)$$

with selected for given value of $\hat{\alpha}$ from Table 2-6-2.

2.6.3 <u>Computation of Station Sample Chi-Squares</u>

The computational procedure is identical for the station samples. To each of them, four selected probability functions are fitted. Since five class intervals are already chosen, four class interval limits for each function and station sample are determined according to following equations for Normal function by eqn (2.33), log-normal 2 by Eq (2.36) and Gamma 2 by Eq (2.38).

Knowing the class interval limits, the corresponding observed class frequencies are determined, squared and summed and then station sample chi-square computed by equation,

$$\chi^2 = \frac{K}{n} \sum_{j=1}^{K} O_j^2 - n \qquad \text{...(2.38)}$$

with n sample size. Since four functions are fitted to annual observations, the station sample is represented by four Chi-square values. These two computed values normal and Gumbel are distributed as chi-square (χ^2) with two degrees of freedom (f = 2.d.f) , while log-normal 2 and Gamma 2 distributed as chi-square (χ^2) with one degree of freedom (f = 1.d.f). These four chi-square values for the station, one of each of the four probability density functions, give automatically the measure of goodness of fit of a particular theoretical function to observed data. Class interval limits, observed class interval frequencies and chi-square for all four functions and the station sample, are been computed, in the next section. For this purpose the chi-squares with one and two degrees of freedom and different level of significance are given in Table 9-1-7.

The computations covered are as follows.

TABLE 2-6-1

NORMAL AND GUMBEL DENSITY FUNCTION FOR COMPUTATION OF CLASS INTERVAL LIMIT VALUES

No. of class interval limit j	1	2	3	4
Probability F(u)	0.20	0.40	0.60	0.80
Abscissa U_j	-0.840	-0.255	0.255	0.840
Abscissa V_j	0.349	0.517	0.520	1.620

TABLE 2-6-3

CORRECTION FACTOR $\Delta\hat{\alpha}$ FOR COMPUTATION OF MAXIMUM LIKELIHOOD ESTIMATES OF THE SHAPE PARAMETERS OF GAMMA FUNCTION WITH 2 and 3 PARAMETERS

$\hat{\alpha}$	$\Delta\hat{\alpha}$	$\hat{\alpha}$	$\Delta\hat{\alpha}$
0.200	0.024	1.400	0.006
0.300	0.029	1.500	0.005
0.400	0.025	1.600	0.005
0.500	0.021	1.700	0.004
0.600	0.017	1.000	0.004
0.700	0.014	1.900	0.003
0.800	0.012	2.200	0.003
0.900	0.011	2.300	0.002
1.000	0.009	3.100	0.002
1.100	0.008	3.200	0.001
1.200	0.007	5.500	0.001
1.300	0.006	5.600	0.000

TABLE 2-6-2

INCOMPLETE GAMMA FUNCTION FOR COMPUTATION OF CLASS INTERVAL LIMIT VALUES

Interval J		1	2	3	4	5	6
$I(U,p) = \frac{\Gamma_u(p+1)}{\Gamma_\infty(p+1)}$		0.200	0.400	0.600	0.800	NOT CALCULATED in this case	
$P=\alpha-1$	α	U_1	U_2	U_3	U_4	U_5	U_6
1	2	3	4	5	6	7	8
-0.8	0.2	0.007	0.015	0.036	0.092	0.303	0.932
-0.6	0.4	0.021	0.060	0.147	0.335	0.675	1.381
-0.4	0.6	0.048	0.140	0.299	0.540	0.919	1.630
-0.2	0.8	0.094	0.240	0.434	0.708	1.103	1.806
0.0	1.0	0.153	0133 8	0.559	0.850	1.254	1.947
0.5	1.5	0.313	0.557	0.819	1.131	1.546	2.218
1.0	2.0	0.468	0.748	1.033	1.357	1.774	2.430
1.5	2.5	0.614	0.914	1.217	1.349	1.967	2.610
2.0	3	0.749	1.074	1.382	1.786	2.136	2.770
3.0	4	1.00	1.349	1.670	2.013	2.429	3.049
4.0	5	1.224	1.591	1.921	2.267	2.682	3.291
5	6	1.029	1.810	2.145	2.494	2.907	3.508
6	7	1.620	2.90	2.350	2.70	3.112	3.707
7	8	1.799	2.196	2.540	2.891	3.302	3.891
8	9	1.996	2.370	2.717	3.070	3.480	4.065
9	10	2.126	2.535	2.884	3.238	3.647	4.228
10	11	2.278	2.692	3.043	3.397	3.809	4.383
11	12	2.420	2.838	3.191	3.568	3.983	4.528
12	13	2.563	2.985	3.339	3.694	4.101	4.674
13.	14	2.690	3.120	3.476	3.831	4.238	4.808
14	15	2.828	3.215	3.612	3.986	4.374	4.942
15	16	2.952	3.382	3.740	4.096	4.502	5.067
16	17	3.076	3.508	3.867	4.223	4.624	5.192
17	18	3.194	3.627	3.987	4.344	4.748	5.310
18	19	3.311	3.740	4.107	4.464	4.868	5.429
19	20	3.422	3.859	4.220	4.578	4.981	5.541
20	21	3.532	3.072	4.334	4.681	5.094	5.653

(1) The main values of sample is converted into series of dimensionless quantities i.e.

$K_i = \frac{\sum Q_i}{Q}$, and arranging the series into descending order.

(2) The class interval limits for each distributed one calculated as follows :

(a) Normal Distribution

i) Mean = $(\hat{\mu}) = \frac{1}{n} \sum_{i=1}^{n} K_i$ ...(2.39)

(ii) Standard deviation

$$= \hat{\sigma} = \frac{1}{n} \sqrt{\sum_{i=1}^{n} (K_i - \hat{\mu})^2} \qquad ...(2.40)$$

(iii) Class Interval limits -

$$K_j = \hat{\mu} + U_j \hat{\sigma} \qquad ...(2.41)$$

(b) Gumbel Distribution

(i) Mean $(\hat{\mu}) = \frac{1}{n} \sum_{i=1}^{n} K_i$ ...(2.42)

(ii) Standard deviation

$$\hat{\sigma} = \frac{1}{n} \sqrt{\sum_{i=1}^{n} (K_i - \hat{\mu})^2} \qquad ...(2.43)$$

(iii) Class Interval

$$K_j = \hat{\mu} + V_j \hat{\sigma} \qquad ...(2.44)$$

(c) Log-normal 2

(i) Mean of $\ln K_i = \ln \hat{\mu} = \frac{1}{n} \sum_{i=1}^{n} \ln K_i$ ...(2.45)

(ii) Standard deviation

$$\hat{\sigma} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (\ln K_i - \ln \hat{\mu})^2} \qquad \ldots(2.46)$$

(iii) Class interval limits

$$K_j = \exp (\ln \hat{\mu} + U_j \hat{\sigma}) \qquad \ldots(2.47)$$

(d) Gamma 2

$$\text{(i)} \quad \hat{\alpha} = \frac{1 + \sqrt{1 + \frac{4}{3} (\ln\bar{K} - \frac{1}{n} \sum_{i=1}^{n} \ln K_i)}}{4 (\ln\bar{K} - \frac{1}{n} \sum_{i=1}^{n} \ln K_i)} - \Delta\hat{\alpha} \qquad \ldots(2.48)$$

$$\text{(ii)} \quad \hat{\beta} = \frac{1}{\hat{\alpha}} \frac{1}{n} \sum_{i=1}^{n} K_i = \frac{1}{\hat{\alpha}} \bar{K} \qquad \ldots(2.49)$$

(iii) Class interval limits

$$K_j = \frac{U_j}{\sqrt{\hat{\alpha}}} \qquad \ldots(2.50)$$

3.7 ESTIMATION OF PEAK FLOW ON THE BASIS OF COMPUTED RESULTS

After getting the best fit distribution for the station (1K3) the peak flow can be determined as following :

3.7.1 Normal Distribution

$$\frac{Q_T}{\bar{Q}} = 1 + K\sigma$$

or

$$Q_T = \bar{Q}(1 + K\sigma) \qquad \ldots(2.51)$$

where Q_T is the expected discharge in (T) years.

$\bar{Q}$ is the mean discharge

K is factor coefficient

σ is the standard deviation of the distribution.

3.7.2 Gumbel Distribution

$$Q_T = \bar{Q} \; (1 + K \sigma) \qquad \ldots(2.52)$$

3.7.3 Log Normal ~~gn~~ Distribution

$$\frac{Q_T}{\bar{Q}} = \text{Antilog} \left[\frac{\log \mu + K \log \sigma}{2.3} \right] \qquad \ldots(2.53)$$

3.7.4 Gamma 2 Distribution

$$Q_T = \bar{Q} \; (\sqrt{\beta \cdot U_j} \;) \qquad \ldots(2.54)$$

CHAPTER III

STREAM FLOW SEQUENTIAL MODEL FOR 1K3A GAUGING STATION

3.1 GENERAL

Existing stream flow records are normally not sufficiently extensive provide reliable estimates of many important statistics. This results in reduced precision of estimated future stream flows as it does not give indication of the long term sequences of flows to which the system would be subjected. Besides this the existing data is being subjected to changes resulting from continuous natural or man-made causes that are taking place. To overcome this difficulty modelling of stream flows process is done to generate data which preserves the statistical properties of past records.

Hydrologic processes being stochastic in nature, their modelling is based on the concepts of statistics and probability. Thes hydrologic processes can be essentially classified into the following types of stochastic processes (12) (Kissiel 1965).

(i) Processes characterised by first and second moment function (Time series Model).

(ii) Counting processes

(iii) Probabilistic processes

(iv) Transition type of processes.

The first process is based on emperical investigation of the first and second moment of actual time series. Time s ries involve the estimation and reconstruction of the properties of the underlying process from the sample. The larger the Historical time series the better the estimates of its parameters, assuming stationarity (12,Kissiel, 1965).

The techniques for analysis of time series can be divided into two broad categories, first cateogry is in frequency domain. Power spectrum and cross spectrum are specific techniques of interest for analysis of time series in frequency domain. Practical applications of these techniques are given in studies of Box and Jenkins (1970) Yevdjevich (25) (1971, 1972). The second category , is suitable for analysis of time series in the time domain. Analysis by surplus, deficit, range, and auto-correlation analysis come under this category. Details of these techniques are given in studies by Box and Jenkins (1970) Yevdjevich (25)(1965, 1971, 1972,)Quimpo (18) (1967, 1973) and Kissiel (12) (1962).

Many hydrologists have devised models of flow generation but there is not yet a conclusive model of stream flows generation, let alone one that is capable of predicting future stream flows.

The second process is the one that counts the occurrences of simple events of a specified type. The methods of analysis of these processes include the application of queueing theory and Markov-chains. Queueing theory was first applied to reservoir by Moran (1954). He developed a model of predicting probability distribution function of water in a reservoir. This was later extended by Langbein(13) (1956).

The third process is in which the chances of occurrence of purely random variable is assumed to follow definite probability distribution are termed probabilistic processes such processes are time invariant i.e., the future of the process is independent of the past and present. Study of flow duration curves, frequency graphs, probability of exceedence, recurrence interval etc., are all examples of probabilistic processes.

The fourth process, includes processes that develope in time as a series of transitions of a system from state to state. The process is specified by the probabilities of transition from one state to another and by the degree of dependence upon its past history. Application to hydrology includes extension of rainfall records by Chow and Ramasechan (6) (1965), augumentation of stream flow records by Julian (10) (1960), Brittan (3) (1961), Thomass and Fiering (22) (1962).

3.2 MATHEMATICAL MODELS FOR STREAM FLOWS SYNTHESIS

Many investigators have analysed the time series structure of daily, weekly, monthly and annual runoff series and they have incorporated the corresponding mathematical models for generating equally likely stream flow sequences into the use of water resources system. The concept of stream flow synthesis has been used by several hydraulic engineers by simulation of models based on historical records or generation of synthetic data by means of stochastic models.

Subsequently numerous types of mathematical models have been used for stream flow synthesis. These depending upon the nature of mathematical formulation can be broadly classified as under :

(i) The auto-regressive model.

(ii) Multiple Regression Models .

(iii) Time Series Model.

3.2.1 The Auto-regressive Model

Models of these type represent a regression between recent values of s-tream flow and its past occurrences. The general form of this model may be expressed as

$$X_t = f\,(X_{t-1}, X_{t-2}, \ldots X_{t-k} + \epsilon_t \ldots) \qquad ..(3.1)$$

where,

K is an integer

ϵ_t is a random variable.

(a) High Order Auto-regressive Model

In this model the value of an event is assumed to depend on the value of several past events. The expression given by Prassad (1967) is :

$$X_t = \sum_{k=1}^{m} r_k X_{t-k} + \varepsilon_t \qquad ...(3.2)$$

where,

X_t	is the magnitude of the even under consideration
r_k	serial correlation between X_t and X_{t-k}
X_{t-k}	Magnitude of the event at time (t-k)
m	Limit to which dependence is significant
ε_t	Independent normal variable.

(b) First Order Autoregressive Model

This model assumes that the value of an event at certain time t is only dependent on the value of the event immediately proceeding it i.e., at (t-1) time. When K = 1, the equation (3.2) becomes first order autoregressive model, or also called first order Morkov Model. The above equation them becomes

$$X_t = r_1 X_{t-1} + \varepsilon_t \qquad ...(3.3)$$

The drawback in this model, is that the means and variance of the recorded sequences are not preserved.

Brittan (3) (1961) used a model for generating stationery sequence of annual stream flows which is of the type

$$Q_t = r\, Q_{t-1} + (1-r)\bar{Q} + S(1-r^2)^{1/2} \varepsilon_t \qquad ...(3.4)$$

Here $\bar{Q}$ and S are mean and standard deviation respectively of the recorded sequences of annual stream flows. However this model cannot be used for generation of seasonal stream flows.

The above equations was latter modified by Thomas and Fiering (2) (1962) and developed a model for generation of monthly stream flows by serial correlation of monthly flows.

$$Q_{i+1} = \bar{Q}_{j+1} + b_j(Q_i - \bar{Q}_j) + S_{j+1}(1-r_j^2)^{1/2} \; t_i \qquad \ldots(3.5)$$

where Q_i and Q_{i+1} - are the discharges during the ith and (i+1) th month.

$\bar{Q}_j$ and $\bar{Q}_{j+1}$ are the mean monthly discharges during the j th and (j+1) th month, within a repetitive annual cycle of 12 months.

b_j is the regression coefficient for estimating volume of discharge in (j+1) th month from the jth month.

S_{j+1} is the standard deviation of discharge in the (j+1) month.

r_j is the correlation coefficient between flows in the jth and (j+1) th month.

t_i is the random normal deviate with zero mean and unit variance.

3.2.2. Multiple Regression Model

A model for stochastic stream flow simulation by multiple regression analysis utilising precipitation data, was developed by Bonne(2) (1971) . The generating equation is basically a Markovian model as it includes the former state of the watershed in terms of the proceeding flow and precipitation. The model is in the form of :-

$$X_t = A + B X_{t-1} + CP_t + D P_{t-1} + E \sum_{i=t-j}^{t-1} P_i + S_e \epsilon \qquad \ldots(3.6)$$

where,

X_t	Current monthly flow
X_{t-1}	Previous month flow
P_t	Current month precipitation
j	= 1,2,3,...12, water year month counter.
ΣP_i	accumulative precipitation since the beginning of snow peak.
A,B,C,D,E	Multiple regression coefficient
S_e	Standard error of estimates of the flows.
ϵ	Random deviate with zero mean and unit variance.

3.2.3 Time Series Model

Most of the hydrologic record constitutes a time series denoted by X_t, t, ϵ and T, where X_t is the hydrologic variable attributed to the time interval t, and T is the length of hydrological record. The general model is described as,

$$X_t = T_t + C_t + S_t + R_t \qquad \ldots(3.7)$$

where,

X_t	observed monthly river flow sequence
T_t	Trend component
C_t	Cyclic component
S_t	Stochastic component
R_t	Random component.

As it has been mentioned in the earlier chapter, that the objective of this study is to analyse the structure of the time series of monthly runoff for 1K3A gauging station, and formulate a mathematical model, which will be used to generate the synthetic sequences, the detailed analysis of time series is discussed in the following section.

3.3 TIME SERIES ANALYSIS

Time series of river flows, is a sequence of values arranged in order of their occurrence and can be characterised by statistical properties, a sequence of a variable as a function of another independent variable, usually time, represented by :

$X(t_1), X_{(t2)}, \quad X(t_3) \quad \ldots\ldots$ where $t_1 < t_2 < t_3 \; < \ldots t_m$

In the typical time series there are discernible three main features which seem to be independent of one another and attributable to distinct causes:

(a) a broad long-term movement, called the TREND.

(b) an oscillation about the trend which may be a seasonal effect with fairly regular period or a rather long-period, irregular oscillation, often called a cycle.

(c) A-n irregular, unsystematic or random component, sometimes called the Residual.

However, not all time series exhibit all three of these features.

The hydrologic time series of runoff is a continuous record of flows and for analytical purposes should be transformed into a descrete time series. The choice of a suitable time interval is a necessary first step.

It is, generally, possible to classify time series as being either of two types: stationery or non-stationery. In stationery time series, the general structure and the statistical parameters representing the same, like the mean, do not vary from one segment of series to another. Non-stationery time series, the different segments are dissimilar in one or more aspects. However, in nonstationery time series, it is necessary to consider absolute time since the series cannot be assumed to have begun prior to the time of the initial observation (Chow).

3.3.1 General Model

It has been assumed that a time series X_t of monthly flow sequences of River Rufiji at Pangani Falls (1K3A) can be adequately represented by a linear additive model

$$X_t = T_t + C_t + S_t + R_t \qquad \ldots(3.8)$$

where,

X_t Observed monthly river flow sequences.

T_t Trend component

C_t Cyclic component

S_t Stochastic component

R_t Random component.

3.3.2 Trend Analysis

Trend represents a smooth motion of the series over a long period of time. It always reveals the general tendency of increase or decrease of the hydrological variable with time. Analysis of trend can be done by either the moving average method, which eliminates the minor fluctuation to show up the long term trends, if any more clearly, or by fitting a mathematical trend to the data, the advantage of which lies in extrapolation and interpolation. The draw back in moving average method is that though it tends to smooth out the data it may introduce an oscillatory movement into the random element which may not be present in the original data and this does not preserve the main feature of the time series.

So as to remove the trend, it is necessary to smooth out irregularities in the time series. Assume that the observations x_1 , x_2 ....x_N are taken at equal intervals of time the methods of moving average consists of determining overlapping means of m successive weighted values, for m = 3 .

$$Y_2 = (b_1x_1 + b_2x_2 + b_3x_3) / 3 \qquad ...(3.9)$$

$$Y_3 = (b_1x_2 + b_2x_3 + b_3x_4) / 3 \qquad ...(3.10)$$

$$Y_4 = (b_1X_{N-2} + b_2 X_{N-1} + b_3 X_N)/3 \qquad ...(3.11)$$

The weights of the moving average b_1 , b_2 and b_3 are such that that sum equal to 3. In general, for moving averages of m ,

$$\sum_{i=1}^{m} b_i = m . \qquad ..(3.12)$$

The weights may be either positive or positive and negative.

In the present case, the least square method has been adopted to a mathematical model. The only advantages being that :-

(a) the method expresses trend in the form of a mathematical formula which may be easily interpreted.

(b) Results obtained under the method are definite and independent of any subjective estimate on the part of the statistician.

(c) The resulting equation is in convenient form for extrapolation (extension into future or past).

The only disadvantages are that the technique used is mathematical and the method is based on the assumption that the data follows a trend that can be expressed by a mathematical equation.

If a straight line trend is assumed, the line of the trend will have a formula of the type ,

$$Y = a + b X \qquad ..(3.13)$$

In this formula the values of a and b must be determined The formula, however, will describe any one of an infinite number of lines. It is necessary, therefore to decide which line best describes the data.

The principle of least squares aids in determining the line that best describes the trend of the data. The principle s-tates that a line of best fit to a series of values is a line the sum of the squares of the deviations (the differences between the line and the actual value) about which will be a minimum. There can be only one line having this qualification . (12).

By taking the sum of squares of the residuals as minimum, the normal equation, obtained are

$$\sum Y = nA + B\sum X \qquad ...(3.14)$$

$$\sum XY = A\sum X + B\sum X^2 \qquad ...(3.15)$$

These are solved as simultaneous equations and the values of constants A and B can be found out by -

$$A = \frac{Y\sum X^2 - X\sum XY}{n\sum X^2 - (\sum X)^2} \qquad ...(3.16)$$

$$B = \frac{n\sum XY - X\sum Y}{n\sum X^2 - (\sum X)^2} \qquad ...(3.17)$$

The summation $\sum Y$ denotes the sum of discharges for n number of months of observation.

After the values of the constants A and B are calculated, the relation trend curve can be fitted. Deducting this trend values from the stochastic hydrologic process, the time series will be left with, the period and residual components which may be taken as measure of deviations from the trend line of the time series.

3.3.3 Periodic Component

When significant long term fluctuations in the series of recorded river flows are removed, then removal of periodicities is a pre-requisite to the analysis of stochastic component. The periodic component represents a regular or oscillatory form of variations such as diurnal, seasonal and secular changes, that exist frequently in the hydrological phenomenon. Such variations are of nearly constant length and may be reasonably be assumed sinuosidal with varying frequencies.

The monthly time series $X_1, X_2, X_3, \ldots X_n$ with a fundamental period of length $T = n . t$, where n is the total number of observations equal spaced by t in the period T from t and (t+T), may be expanded into a Fourier Series according to the following formula-

$$X_t = A_{N0/2} + \sum_{n=1}^{M} \left(A_{XN} \cos \frac{2\pi n t}{T} + B_{XN} \sin \frac{2\pi nt}{T}\right) \quad ..(3.18)$$

where,

X_t = flow at month t,

$A_{NO/2}$ mean of series X_t

m number of significant harmonics.

A_{XN}, B_{XN} Fourier coefficient

T Basic period of series for monthly data equal to 12.

ϵ_t — Stochastic component for months t represented by an autogressive scheme and an independent uncorrelated random number.

The harmonic coefficient are defined as :

$$A_{Xn} = (\frac{2}{N}) \sum_{t=1}^{n} X_t \cos (2\pi nt/T) \qquad ..(3.19)$$

$$B_{Xn} = (2/N) \sum_{t=1}^{n} X_t \sin (2\pi n_t/T) \qquad ...(3.20)$$

where

N is the number of the data points

n is number of years

T the basic period of series.

Also Equation(3.18) can be expressed in a different form as

$$X_t = A_{No/2} + \sum_{n=1}^{n} C_{Xn} \cos (2\pi n_t/T - \theta_{Xn}) \qquad ...(3.21)$$

Here

$$C_{Xn} = (A_{Xn}^2 + B_{Xn}^2)^{1/2} \qquad ...(3.22)$$

and

$$\theta_{Xn} = \tan^{-1}(\frac{B_{Xn}}{A_{Xn}}) \qquad ...(3.23)$$

θ_{Xn} is defined as the phase angle .

Equation (3.21) states that complex periodic data consists of a stationery mean value component, $A_{No/2}$ and an infinite number of sinusoidal component (harmonics) that have amplitude C_{Xn} and phase θ_{Xn} .

If S^2 is the total variance of the time series X_t, the part of the variance accounted for by the Kth harmonic is

$$C_{Xn}^2 / 2S^2 \text{ or } (A_{Xn}^2 + B_{Xn}^2) / 2S^2 \qquad ...(3.24)$$

If $Y_{n1}, Y_{n2}, \ldots, Y_{nn}$ be the magnitude of n -harmonic and if n harmonics are present in the series, then the value of the periodic component is given by :

$$C_t = (Y_{n1} + Y_{n2} + Y_{n3} + \ldots + Y_{nn}) \quad \ldots(3.25)$$

3.3.3.1 Correlogram and Auto-correlation Analysis

The auto-correlation analysis is used to find the inter-dependence of successive values of a time series. A measure of this dependence is given by the auto-correlation coefficient.

For a discrete time series, it is defined as :

$$\rho_k = \frac{E(X_t \cdot X_{t+t}) - E(X_t)\, E(X_{t+t})}{(E)(X_t^2) - (E\, X_t)^2} \quad \ldots(3.26)$$

where, $K = 1, 2, 3, \ldots, m$ and $m < N$

If ρ_k is plotted as ordinates against their respective lag values K as abscissa and the plotted pointed are joined each to the next by a straight line, the resulting plot is a correlogram. The auto-correlation coefficient of the continuous series is commonly known as serial correlation coefficient. If the correlation between the two is referred to as the kth order serial correlation and is given by :

$$r_k = \frac{Cov\,(X_t \cdot X_{t+k})}{\left[Var\,(X_t)\; Var\,(X_{t+k})\right]^{1/2}} \quad \ldots(3.27)$$

where

Cov (X_t, X_{t+k}) is the sample autocovariance and $Var(X_t)$ and Var (X_{t+k}), the sample variance.

Further,

$$Cov(X_t \cdot X_{t+k}) = \frac{1}{N-K} \sum_{t=1}^{N-K} X_t \cdot X_{t+k} - \frac{1}{(N-K)^2} \left(\sum_{t=1}^{N-K} X_t \right) \left(\sum_{t=1}^{N-K} X_{t+k} \right)^2$$

$$Var(X_t) = \frac{1}{N-K} \sum_{t=1}^{N-K} X_t^2 - \frac{1}{(N+K)^2} \left(\sum_{t=1}^{N-K} X_t \right)^2$$

$$Var(X_{t+k}) = \frac{1}{N-K} \sum_{t=1}^{N-K} X_{t+k}^2 - \frac{1}{(N-K)^2} \left(\sum_{t=1}^{N-K} X_{t+k} \right)^2$$

Then,

$$r_t = \frac{(N-K) \sum_{t=1}^{N-K} X_t X_{t+K} - \left(\sum_{t=1}^{N-K} X_t \right) \left(\sum_{t=1}^{N-K} X_{t+k} \right)}{\left[(N-K) \sum_{t=1}^{N-K} X_t^2 - \left(\sum_{t=1}^{N-K} X_t \right)^2 \right]^{1/2} \left[(N-K) \sum_{t=1}^{N-K} X^2{}_{t+k} - \sum_{t=1}^{N-K} X_{t+K} \right]^{1/2}} \quad \ldots(3.28)$$

Where N = length of samples. For stricktly mean random sequencies the correlogram will have a value of 1 at K = 0 and an expressed value of zero at all other points.

<u>Confidence Bands</u>

The correlogram can be tested with confidence limits at a given level of significance. Anderson(1947) has formulated the formula for confidence limits L_α for the correlogram as:

$$L_{\alpha} = \frac{-1 \pm n_{\alpha} \sqrt{N-L-2}}{N-L-1}$$

where

N is the number of observed values in the time series X_t

L Lags used

n_{α} Normal standard deviate from the standard normal distribution for two tail test at a significance level.

Common value of α and the corresponding values of the n_{α} are

α	n_{α}
80%	1.28
90%	1.64
95%	1.96

3.3.5 Power Spectrum Analysis

A power spectrum is the distribution of the variance of X_t on a frequency scale. The mathematical development of spectrum indicates the relation between the auto correlation function and the spectrum. Thus the ordinate of the spectrum represents the variance density as a function of the angular frequency ($\omega = 2\pi f$), given by $S(\omega)$, then we get,

$$S_x(\omega) = \frac{1}{\pi} \int_0^{\infty} \rho_\tau \, e^{-2\pi f \tau i} \, d\tau \qquad \ldots(3.29)$$

$$S_x(\omega) = \frac{1}{\pi} \int_{\tau=0}^{\infty} \rho_\tau \, e^{-2\pi f \tau i} \, dz \qquad \ldots(3.30)$$

$$\text{Also } \rho_\tau = \frac{1}{\pi} \int_0^\infty SX(\omega) e^{-2\pi f \tau i} \, d\omega \qquad ..(3.31)$$

$$\rho_\tau = \frac{1}{\pi} \sum_{\omega=0}^{\infty} SX(\omega) e^{-2\pi f \tau i} \, d\omega \qquad ...(3.32)$$

When τ is the time lag in case of continuous time series.

The one sided power spectrum is given by

$$G_x(f) = 2\, S_x(f), \quad 0 \leq f \leq \infty \qquad ...(3.33)$$

$G_x(f)$ is a continuous power spectra over the +ve frequency range $(0, \infty)$

For real value process, the above equations are simplified . The real valued two sided power spectrum is obtained from equation (3.29) by making the imaginary part equal to zero

$$\therefore \quad S_x(f) = \frac{1}{\pi} \int_0^\infty \rho_\tau \cos 2\pi f \tau \, d\tau \qquad ...(3.34)$$

or

$$S_x(f) = \frac{1}{\pi} \sum_{\tau=0}^{\infty} \rho_\tau \cos 2\pi f \tau \qquad ...(3.35)$$

and for discrete series

$$S_x(f) = \frac{1}{\pi} \sum_{k_{min}}^{k_{max}} r_k \cos 2\pi f t \qquad ...(3.36)$$

$S_x(f)$ is known as normalised spectral density functio . Since it is the Fourier transform of the auto-correlation function, and also if $G_x(f)$ is Fourier transform of auto covariance function , then

$$S_x(f) = \frac{G_x(f)}{\sigma_x^2} \qquad ...(3.37)$$

$$S_x(f) = C_x(f) + i\, Q_x(f) \qquad ...(3.38)$$

where $C_x(f)$ and $Q_x(f)$ are the co-spectrum and quadrature spectrum respectively.

The euqation (3.36) is used for the calculation of spectral density function at various frequencies, giving the normalised power spectrum.

The peaks in the spectral density functionsuggests a possible periodicity in the stochastic process. The spectral analysis gives a powerful tool to recognise not only the presence of periodic oscillations but also the relative importance between the several harmonics and also in identifying mathematical models and to simulate hydrologic time series. Limitations in applying the spectral density of theory to hydrologic series are :

1. The series for analysis is always a finite in size.
2. Time scale is discrete rather than continuous, owing either to the instantaneous observations or to averaging of matural processes over time interval.

Therefore in an actual case the fundamental assumption of a continuous spectrum corresponding to all frequencies from 0 to ∞ is unattainable.

The normalised power spectrum $S_x(f)$ is calculated by assuming some value of time lag interval and the values of correlation coefficient with the same time lag interval are found out by the help of computer programme (No.2) Appendix.(For) the monthly river flows data K has been taken as 6 months. K_{min} and K_{max} has been taken as 6 and

120 and the value of $S_x(f)$ is found out corresponding to the different values of frequencies.

3.3.6 The Stochastic Component

The residual series, Z_t , after the removal of trend and periodic component from the original time series, consists of stochastic component S_t, and an uncorrelated random component, Rt.

$$Z_t = X_t - T_t - P_t = (St + Rt) \qquad ...(3.39)$$

The residual series belongs to a class of non-deterministic processes which include auto-regressive moving averages and other schemes of linear regression.

The type of model to be fitted to the stochastic component can be ascertained from the correlogram analysis of the residual series . First or second order auto-regressive model can be fitted depending upon the shape of the correlogram.

The measure of best of fit of the auto-regressive model is simplified by the determinationof coefficient, R_i^2 , i = 1,2,3,...m. Since $R_m^2 > R_3^2 > R_2^2 > R_1^2$ a certain criteria can be developed as when a model of a given order should be selected . The determination of the coefficient for the first, second or third auto-regressive model in terms of rk are given by-

$$\hat{R}_1^2 = r_1^2 \qquad ...(3.40)$$

$$R_2^2 = (r_1^2 + r_2^2 - 2r_1^2\, r_2) \;/\; (1 - r_1^2) \qquad \ldots(3.41)$$

$$R_3^2 = \frac{r_1^2+r_2^2+r_3^2+2r_1^3r_3+2r_1^2r_2^2+2r_1r_2^2r_3-2r_1^2r_2-4r_1r_2r_3-r_1^4-r_2^4-r_1^2r_3^2}{(1-2r_1^2-r_2^2+2r_1^2r_2)}$$

$$\ldots(3.42)$$

The first order model is selected if

$$R_2^2-R_1^2 \leqslant 0.01 \quad \text{and} \quad R_3^2 - R_2^2 \leqslant 0.02 \qquad \ldots(3.42a)$$

The second order model is selected, if

$$R_2^2 - R_1^2 > 0.01 \quad \text{and}$$

$$R_3^2 - R_2^2 \leqslant 0.01 \qquad \ldots(3.42b)$$

The third order model is selected, if

$$R_2^2 - R_1^2 > 0.01 \quad \text{and}$$

$$R_3^2 - R_2^2 > 0.01 \qquad \ldots(3.42c)$$

Once the order of the model is ascertained according to the above criteria, the auto-regressive coefficient can be calculated for the 1st , 2nd and 3rd order as follows:

<u>1st Order</u>

$m = 1$

$a_1 = r_1$

<u>2nd Order</u>

$m = 2$

$$a_1 = \frac{r_1 - r_1 r_2}{1 - r_1^2}\,, \quad a_2 = \frac{r_2 - r_1^2}{1 - r_1^2}$$

<u>3rd order</u>

$m = 3$

$$a_1 = \frac{(1-r_1^2)\ (r_1-r_3) - (1-r_2)\ (r_1 r_2 - r_3)}{(1-r_2)\ (1-2r_1^2 + r_2)}$$

$$a_2 = \frac{(1-r_2)\ (r_2+r_2^2 - r_1^2 - r_1 r_3)}{(1-r_2)\ (1-2r_1^2 + r_2)}$$

$$a_3 = \frac{(r_1-r_3)\ (r_1^2-r_2) - (1-r_2)\ (r_1 r_2 - r_3)}{(1-r_2)\ (1-2r_1^2 + r_2)} \qquad \ldots(3.43)$$

So as to decide upon the order of the model to be used the values of R_1^2 , R_2^2 and R_3^2 have to be calculated.

These values have been found out to be

$R_1^2 = 0.371$

$R_2^2 = 0.372$

$R_3^2 = 0.386,$

Therefore the first order Markov Model was fitted to the stochastic component, as equation (3.42c) is satisfied.

$$Z_t = r_1\ Z_{t-1} + R_t \qquad \ldots(3.44)$$

3.3.7 Random Component

The Random component is obtained by removing from the river run off series, the trend, periodicity and the stochastic component

$$R_t = X_t - T_t - C_t - S_t \qquad \ldots(3.45)$$

The random function may be defined simply as one which cannot be formulated in a manner to provide precise prediction of values of function. Although the function is concerned as being the wet effect of certain physical causes the number of caus s is viewed as being very large, with each cause producing a small effect.

The random component in the absence of any trend is obtained by removing the cyclic and stochastic component from the original time series as given in the above equation.

3.4 GENERATION OF RANDOM NUMBERS

The most acceptable methods for generating random number-s are :

(a) Unifommly distributed.

(b) Statistically independent

(c) Reproducible.

(d) Non-repeating for any desired length.

(e) Capable of generating random numbers at high rates of speed.

(f) Require minimum computor memory.

One of the simplest methods of generating random numbers is the method of sampling cards. The cards are first numbered, one for each value . The cards are then shuffled and a card is drawn at random, its value is taken as first generated random numbers.

However, this method have been simplified by the use of random number tables, which have been constructed by more efficient methods. These tables have been subject to the standard statistical tests of randomness and are considered accpptable for general sampling use. Generally, standard programs and subroutines, for generating normally distributed random number, and rectangularly distributed random numbers, respectively, are available in IBM and in most mathematics Libraries , s tatistics Laboratories and offices. The s-tandard Computer program for this purpose, is furnished as per appendix (X).

CHAPTER IV

ANALYSIS OF RESULTS AND CONCLUSIONS

4.1 ANALYSIS OF RESULTS

4.1.1 Application of Curve Fittings to Rufiji River at Stiegler's Gorge (1K3).

A simple procedure making use of the generalized frequency equation by the application of Gumbel method curve fitting is given below for the above mentioned river gauging station:

i. List the annual (or seasonal) floods

ii. Compute $\bar{X}$ and SX by using equations

$$\bar{X} = \frac{\sum X}{N} \quad \text{and}$$

$$SX = \sqrt{\sum (X-X)^2/N-1} = \sqrt{\frac{N}{N-1}(\bar{X}^2 - \bar{X}^2)}$$

iii. Prepare a computation form with column heading from left to right, as follows:

T , $\bar{X}$, SX , KSX , X = $\bar{X}$ KSX . The table is convenient method of computing X-values from given (T) values by formula X = $\bar{X}$+KSX

iv. 8 From Table 2-4-1 showing (K) factors for the Gumbel method, list in the list (T) column of the computation form a representative selection of return periods for which there are columns in Table 2-4-1.

v. Entér the computed values for ($\bar{X}$) and (Sx) on the computation form in the appropriate columns. The same values apply for all T-values.

vi. For each of the selected T-values, extract the K-factor from Table 2-4-1 and these in the computation form. Note that the values of (N), which is the number of floods of records, is used in extracting the K-factors and that interpolation may be necessary.

vii. Compute values for (KSx) and (X) for each T-value and enter these values in the computation form. The X-values are the flood magnitudes for the return period (T). They are used for constructing the frequency curve.

viii. Using the extreme probability paper plot the X-values (or ordinates) from the computation form and join them with a straight line to obtain the required frequency curve.

ix. Note that it is necessary to plot the entire frequency curve if the (T) value for a given (X) value, or the (X) value for a given (T) value is required. After carrying out step (i),(ii) from Table 2-4-1, formula $X = \bar{X} + KSX$can be used in conjunction with (K) factors to derive the required value of either (T) or (X), as the case may be.

x. To judge the goodness of fit the observed data are also plotted on the extreme value probability paper depending on Table 4-1-4.

The constructed frequncy curve computations for 1K3, by the Gumbel method is shown in Table 4-1-2 Table 4-1-5 shows the computations for fitting Gumbel's law (as adopted by Ven Techow) by least square method. The law is expressed as

$$Y = A + B \log_{10} \log_{10} \frac{T}{T-1}$$

where (Y) is the flood with a return period T. The step by step procedure is as given below:

(i) Rank the observed floods (Y) of the annual series in decreasing order.

(ii) Compute T-values for each of Y-values by using $T = \frac{N+1}{M}$

(iii) Compute X-values where $X = \log_{10} \log_{10} \frac{T}{T-1}$ for all the items.

(iv) Compute the product (XY) and X^2 for all the items.

(v) Find out summations $\sum X$, $\sum Y$, $\sum X^2$ and $\sum XY$, and substitute these values in the normal equations to obtain parameters (A) and (B) of the least square line.

(vi) Plot of the fitted equation of line on extreme value probability paper after computing a few values of (Y) for selected (T) values. This is the required frequency line.

vii. To judge the goodness of fit, the observed data are also plotted on the same paper depending on Table 4-1-5.

Figure(4-1-1) shows the best fit line and the observed flood plotted on an extreme value probability paper for 1K3. Table 4-1-5 shows the comput tions for fitting Gumbel law by method of least squares.

For determining the confidence bands , a simple procedure to compute the confidence limits for the Gumbel frequency curve for 1K3 station is as follows:

(i) Compute $Sx / \sqrt{N}$

(ii) For the given return period (T) and (N) compute τ (a factor derived from Gumbel K-factors) using relation $\tau = 1+1.4K + 1.1K^2$. The values of (τ) are given in Table 4-1-2 for the appropriate values of (N) and (T).

(iii) Compute the factor $\sigma_H = \tau\ Sx / \sqrt{N}$

(iv) Select the desired confidence limit and the corresponding value of t from Table 4-1-1.

(v) Compute product $t.\ \sigma_H$.

(vi) Compute values for $X- t\ \sigma_H$ and $X+t\ \sigma_H$. These values represent the upper and lower limits of the X- values for the selected confidence limit at the given return period T. Plot the results at the appropriate (T) abscissa on the frequency curve.

vii. Repeat the operation for one or more other values of T.

A straight line joining the plotted limiting points will provide the acquired bands for selected confidence limits. The procedure is illustrated in an example in Table 4-1-6 and the confidence bands are shown in Figure (4-1-1).

In the case of curve fitting by log normal method and using chow's frequency factors, the frequency curve is derived for the station in question. The computation procedure is as follows:

(i) Compute $\bar{X}$ and S_x from the annual series of floods as shown in the Table 4-1-5.

(ii) Compute Cv_1 i.e. $Sx/\bar{X}$ where (Cv) is the coefficient of variation.

(iii) Compute Cs from $Cs = 3Cv + C_x^3$ where Cs is the coefficient of skewness, or extract (Cs) from the subsidiary columns in Table 2-4-1 , the Chow Table for (K) factors for the computed value of Cv , using interpolation if necessary.

(iv) Set up a computation form as shown in Table 4-1-3

(v) Enter representative (P) value from columns in Table 2-4-1 . Enter $\bar{X}$ and Sx values in the computation form.

(vi) From Table 2-4-1 of Chow Frequency factor select (K) factor for selected (P) value entering in the table on the line for the computed (Cv) and (Cs). Enterpolation may be necessary. Enter the (K) factors in the computation form.

TABLE 4-1-1 VALUES OF STANDARD NORMAL VARIATE FOR VARIOUS PROBABILITIES

Probability	0.500	0.683	0.800	0.900	0.950	0.980	0.990
t	0.674	1.000	1.282	1.645	1.960	2.326	2.576

TABLE 4-1-2 VALUES OF τ for use IN COMPUTING CONFIDENCE LIMITS FOR GUMBEL CURVE

n	T=10	20	25	30	50	75	100
15	2.476	3.233	3.409	3.604	4.113	4.525	4.818
20	2.400	3.075	3.292	3.468	3.968	4.362	4.843
25	2.350	3.007	3.228	3.391	3.874	4.259	4.533
30	2.137	2.960	3.166	3.356	3.811	4.187	4.455
40	2.272	2.898	3.099	3.264	3.725	4.093	4.353
50	2.244	2.857	3.056	3.217	3.671	4.031	4.288
60	2.224	2.830	3.025	3.185	3.633	3.489	4.242
75	2.201	2.800	2.976	3.150	3.592	3.943	4.194
100	2.181	2.769	2.959	3.114	3.549	3.896	4.142

TABLE 4-1-2

CONSTRUCTION OF FREQUENCY CURVE FOR 1K3 BY GUMBEL METHOD

Return period(T)	$\bar{X}$	SX	K from Table III	K.SX Col.3x4	Flood flow in cumsecs $X=\bar{X}+KSx$ Col.2+5
1	2	3	4	5	6
5	36.42	12.68	0.950	12.05	48.47
10	36.42	12.68	1.672	21.20	57.62
20	36.42	12.68	2.366	30.00	66.42
50	36.42	12.68	3.264	41.39	77.81
100	36.42	12.68	3.947	50.05	86.47

TABLE 4-1-3

CONSTRUCTION OF FREQUENCY CURVE FOR 1K3 BY LOG NORMAL METHOD

Cv = Coefficient of Variation = $Sx/\bar{X} = \frac{12.68}{36.42} = 0.348$

Cv = 0.348

Cs = Coefficient of Skew = $3Cv+Cv^3 = 1.044+0.042=1.086$

Probability(P) in %	$\bar{X}$	Sx	K From Table (	K.Sx	Flow in cumsec $X=\bar{X}+KSx$
1	2	3	4	5	6
95	36.42	12.68	-1.22	-15.47	20.95
50	36.42	12.68	-0.20	-2.54	33.88
20	36.42	12.68	+0.688	8.62	45.04
5	36.42	12.68	+1.89	23.96	60.38
1	36.42	12.68	+3.29	41.72	78.14

TABLE 4-1-4

FLOOD FREQUENCY ANALYSIS BY USING FREQUENCY FACTOR FOR RUFIJI RIVER AT STIEGLERS GORGE (1K3) FOR THE PERIOD 1955-1972

Year	Annual Peak(X) discharges in (100)cumecs in descending order	Order (M)	Plotting Positions		X^2
			Return Period $T=\frac{N+1}{M}$ (in yrs)	Probability $P(X \geq X)=\frac{M}{N+1}$ in percent	
1	2	3	4	5	6
1956	54.20	1	19.00	5.26	2937.64
1960	44.80	2	9.50	6.13	2007.04
1962	44.00	3	6.33	6.69	1936.00
1968	43.00	4	4.75	10.25	1849.00
1963	42.80	5	3.80	12.82	1831.84
1961	41.90	6	3.17	15.38	1655.61
1964	41.00	7	2.71	20.05	1681.00
1970	40.80	8	2.35	20.50	1664.64
1955	40.00	9	2.11	23.09	1600.00
1958	36.40	10.	1.90	25.63	1324.96
1972	35.60	11	1.73	28.19	1267.36
1971	34.60	12	1.58	30.74	1197.16
1957	32.60	13	1.46	32.32	1062.76
1965	27.80	14	1.36	35.88	772.84
1967	27.80	15	1.27	38.45	772.84
1966	26.00	16	1.18	41.00	676.00
1969	24.40	17	1.12	45.67	595.36
1959	18.00	18.00	1.06	46.13	324.00
Total	$\sum$ 655.70				$\sum X^2 = 23616.05$

Table 4-1-4 (Contd..)

$\bar{X}$ = 36.42 cumecs

SX = 12.68 Cumecs

$$\text{Mean } \bar{X} = \frac{\sum X}{N} = \frac{655.70}{18} = 36.42$$

$$\text{Squared Mean } (\bar{X})^2 = (36.42)^2 = 1326.42$$

$$\text{Mean of squares } \overline{X^2} = \frac{\sum X^2}{N}$$

$$= \frac{23616.05}{18}$$

$$= 1303.00$$

$$\text{Standard Deviation } S_x = \sqrt{\frac{N}{N-1} \left(\overline{X^2} - \bar{X}^2 \right)}$$

$$= \sqrt{\frac{18}{17} (1403.11) - 1326.42}$$

$$= 12.68.$$

TABLE 4 - 1 -5

FITTING OF GUMBEL'S LAW $Y = A + B \log_{10} \log_{10} \frac{T}{T-1}$

TO ANNUAL FLOODS OF RUFIJI RIVER AT STIEGLER'S GORGE(1K3)

Year	Annual Peak Discharges (Y) in 100 cumecs	Order M	$T = \frac{N+1}{M}$ (Years)	$\frac{T}{T-1}$	$X = \log_{10} \log_{10} \frac{T}{T-1}$	XY	X^2
1	2	3	4	5	6	7	8
1956	54.20	1	19.00	1.055	-1.6326	-1 722	2.965
1960	44.80	2	9.50	1.11	-1.134	-1.260	1.587
1962	44.00	3	6.33	1.19	-1.1221	-1.335	1.782
1968	43.00	4	4.75	1.27	-0.9839	-1.250	1.563
1963	42.80	5	3.80	1.39	-0.8446	-1.174	1.378
1961	41.90	6	3.17	1.46	-0.784	-1.145	1.311
1964	41.00	7	2.71	1.58	-0.7007	-1.007	1.225
1970	40.80	8	2.35	1.740	-0.6189	-1.077	1.160
1955	40.00	9	2.11	1.90	-0,5547	-1.054	1.111
1958	36.40	10	1.90	2.111	-0.4890	-1.032	1.065
1972	35.60	11	1.73	2.369	-0.4264	-1.010	1.020
1971	34.60	12	1.58	2.724	-0.3581	-0.975	0.950
1957	32.60	13	1.46	3.173	-0.2999	-0.951	0.840
1965	27.80	14	1.36	3.711	-0.2393	-0.904	0.817
1967	27.80	15	1.27	4.703	-0.1724	-0.811	0.658
1966	26.00	16	1.18	6.555	-0.0980	-0.642	0.412
1969	24.40	17	1.12	9.333	-0.0132	-0.123	0.015
1959	18.00	18	1.06	17.666	00.0969	1.712	2.931

$\sum Y = 655.70$

$\sum X = -10.3750$ $\sum XY = -15.860$ $\sum X^2 = 22.854$

Table 4-1-5 (Contd..)

$\bar{Y} = \frac{\sum Y}{N} = 36.42$, $B = \frac{\sum XY - N\bar{X}\bar{Y}}{\sum X^2 - N\bar{X}^2}$

$\bar{X} = \frac{\sum X}{N} = -0.576$

$B = -21.25$

$\bar{X}^2 = 0.321$

$A = \bar{Y} - B\bar{X} = 24.18$

Line of best Fit, $Y = A + B \log_{10} \log_{10} \frac{T}{T-1}$

$= 24.18 - 21.25 \log_{10} \log_{10} \frac{T}{T-1}$

Return period (T) in years	$X = \log_{10} \log_{10} \frac{T}{T-1}$	Estimated Flood flow(Y) in 100 cumecs
10	-1.3439	52.76
25	-1.7696	61.78
100	-2.3665	74.47

TABLE 4 - 1 - 6

COMPUTATION OF CONFIDENCE LIMITS FOR GUMBEL FREQUENCY CURVE FOR RUFIJI RIVER AT STIEGLER'S GORGE

(i) The standard normal variate for 95% probability

$t = 1.96$ from Table (4-1-1)

(ii) $\frac{Sx}{N} = \frac{Sx}{\sqrt{18}} = \frac{12.68}{\sqrt{18}} = \frac{12.68}{4.24} = 2.99$

(iii) τ - a factor derived from Gumbel (K) factor using relation , $\tau = 1 + 1.4K + 1.1\,K^2$

Return period (T) in years	Estimated flood flow (X) in 100 cumecs from Table 4-1-2	τ from table 4-1-2	$\sigma_H = \tau \frac{SX}{\sqrt{N}}$ $= \tau \times 2.99$	$t\,\sigma_H$	Confidence limits (100 cumecs) Upper $X + t\sigma_H$	Lower $X - t\sigma_H$
1	2	3	4	5	6	7
10	59.63	2.445	7.31	14.33	73.96	45.30
20	70.38	3.156	9.44	18.50	88.88	51.88
50	84.50	4.055	12.12	23.75	107.80	60.30
100	94.65	4.748	14.20	27.83	122.48	66.82

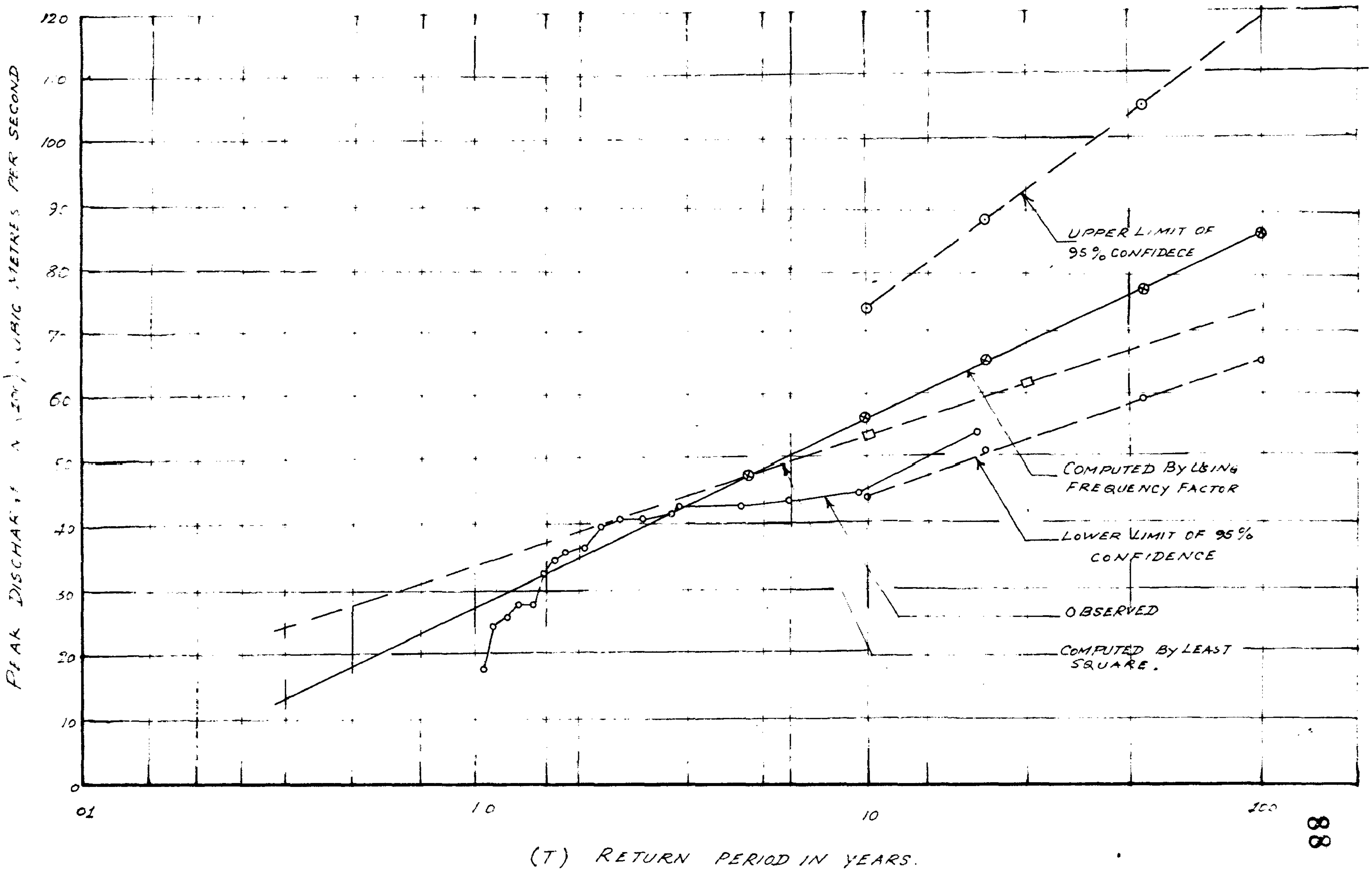

FIG. 4-1-1 CONSTRUCTION OF ANNUAL FLOOD FREQUENCY BY GUMBELS METHUD FOR IK3 GAUGING STATION (1961 - 1972)

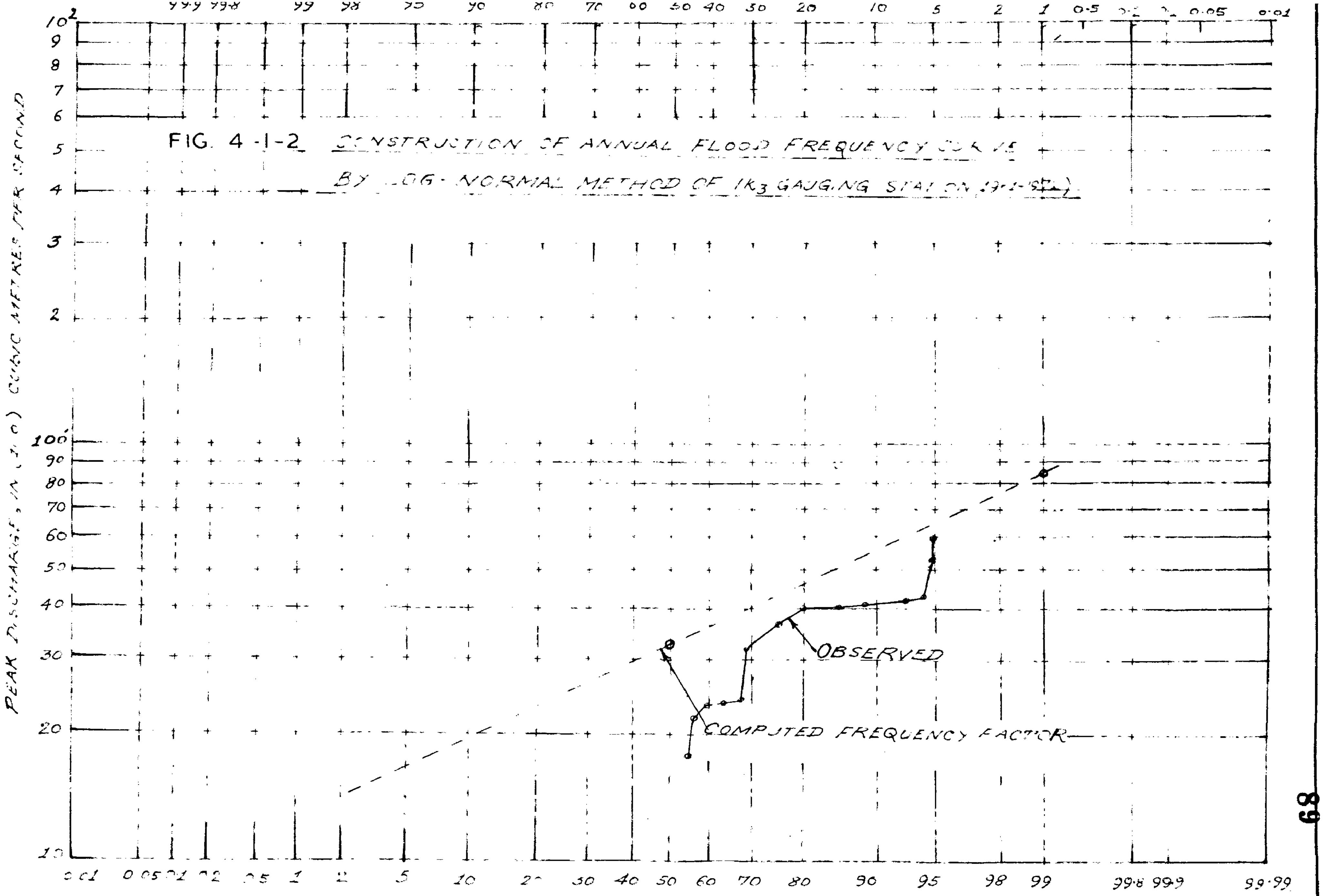

FIG. 4-1-2 CONSTRUCTION OF ANNUAL FLOOD FREQUENCY CURVES BY LOG-NORMAL METHOD OF IK3 GAUGING STATION

(vii) Compute values for (KSx) . Compute (X) values for each of the (P) values by the formula $X = \bar{X} + KSx$

(viii) Plot the (X) values at the appropriate (P) value abscissa on log-probability paper. Draw a straight line through the plotted points to produce the required frequency curve.

(ix) To judge the goodness of fit the observed data are also plotted on the same probability paper depending on the Table 4-1-5. Figure (4-1-2) shows the best fit line and the observed flood plotted on any log-normal probability paper. Also table (4-1-3) shows the computations for the construction of frequency curve for the station, by the log-normal method.

4.1.2 Maximum Likelihood Estimates

Before the maximum likelihood estimates are calculated, the observed data for 1K3, gauging station is tabulated as given in Table 4-1-9. Using the sum in column 3 the sample mean of the actual observed data is calculated, which comes out to be 36420 cumecs. With this the, observed annual flows are transformed into dimensionless form, in terms of sample mean by Eqn $K_i = \frac{Q_i}{\bar{Q}}$ and given in column (4). where K_i is the modular coefficient, $\bar{Q}$ is the sample mean, Q_i is the annual observed river flow. The modular coefficients are arranged in an array, in order to simplify or facilitate the further computation.

After this, procedure the maximum likelihood estimates are computed as follows:

(1) Normal Function - Equation (2.39) and numerical data in column (4) Table 4-1-9,

$$\hat{\mu} = \frac{1}{18} \times 18 = 1.0$$

and Equation (2.40) with column (6) Table 4-1-9

$$\hat{\sigma} = \sqrt{\frac{1}{18} \times 1.034} = 0.075$$

(2) Log Normal 2 Function - Applying Eqn.(2.45) and column (7), Table 4-1-9

$$\ln \hat{\mu} = (-0.5086) \frac{1}{18}$$

$$= -0.028$$

With Eqn (2.45) and Column (9)

$$\hat{\sigma} = \sqrt{\frac{1}{18} \times 1.109}$$

$$= 0.248$$

(3) Gumbel Function - The procedure is the same as Normal Function.

(4) Gamma 2 Function - Using Eqn (2.48) and Column (7), Table 4-1-9

$$\hat{\alpha} = \frac{1 + \sqrt{1 + \frac{4}{3}\left[0 - \frac{1}{18}(-0.5086)\right]}}{4\left[0 - \frac{1}{18}(-0.5086)\right]} - \Delta\hat{\alpha}$$

$$= 5.05 - \Delta\hat{\alpha}$$

$$= 5.05 - 0.001 = 5.049.$$

Since the correction factor $\Delta \hat{\alpha}$ is 0.001 for $\hat{\alpha} = 5.05$ according to Table 2-6-3 and Table 2-6-2

$$\beta = \frac{1}{\alpha} \times \frac{1}{18} \times 18 = \frac{1}{5.049} \times \frac{1}{18} \times 18$$

$$= 0.200$$

4.1.3 Class Interval Limits and Observed Class Frequencies

(1) Normal For five class intervals, four class interval limits one computed by equation (2-4-1) and Table (4-1-9) and observed class frequencies, O_j are determined and squared as follows :

	O_j	O_j^2
	6	36
$K_1 = 1.000 - 0.840 \times 0.075 = 0.949$	2	4
$K_2 = 1.000 - 0.255 \times 0.075 = 0.980$	1	1
$K_3 = 1.000 + 0.255 \times 0.075 = 1.02$	1	1
$K_4 = 1.000 + 0.840 \times 0.075 = 1.06$	8	64
	18	106

(2) Log-Normal 2 - According to Eqn (2.47) and parameter estimates, computed previously, the class intervals limits are :

	O_j	O_j^2
	5	25
$K_1 = \exp(-0.028 - 0.840 \times 0.248) = 0.819$	1	1
$K_2 = \exp(-0.028 - 0.255 \times 0.248) = 0.905$	7	49

$K_3 = \exp(-0.028 + 0.253 \times 0.248) = 1.152$

$K_4 = \exp(-0.028 + 0.840 \times 0.248) = 1.197$

2	4
3	9
18	88

(3) Gamma 2 - Equation (2.50) with corresponding values of U_j from table 2-6-2 gives.

$$K_1 = \frac{1}{\sqrt{5.049}} \times 1.244 = 0.560$$

$K_2 = 0.45 \times 1.613 = 0.725$

$K_3 = 0.45 \times 1.943 = 0.874$

$K_4 = 0.45 \times 2.289 = 1.030$

O_j	O_j^2
1	1
2	4
3	9
3	9
9	81
18	104

4.1.4 Computation of Sample Chi-Squares

The sample Chi-squares are computed by Eqn (2.38) for each selected function separately and then converted into corresponding probability by the help of table 4-1-7.

(1) Normal : for f = 2 degrees of freedom.

$$\chi^2 = \frac{5}{18} \times 106 - 18 = 12 \text{ and } P(\chi^2) = >0.999$$

(2) Log Normal 2 - For f = 1 degree of freedom

$$\chi^2 = \frac{5}{18} \times 88 - 18 = 6.4 \text{ and } P(\chi^2) = 0.940$$

TABLE 4 - 17

CHI - SQUARE DISTRIBUTION

χ^2 for

$F(\chi^2)$	f=1 d.f	f=2 d.f	$F(\chi^2)$	f=1 d.f f	f = 2 d.f
0.001	0.000	0.020	0.700	2.706	4.605
0.005	0.001	0.040	0.750	3.841	5.991
0.010	0.004	0.103	0.800	5.020	7.380
0.020	0.016	0.211	0.900	5.412	7.824
0.025	0.064	0.446	0.950	6.635	9.210
0.050	0.102	0.575	0.980	7.800	10.607
0.100	0.148	0.713	0.990	10.827	13.815
0.200	0.455	1.386			
0.250	1.074	2.408			
0.300	1.320	2.770			
0.500	1.642	3.219			

TABLE 4 - 1 - 9

DATA FOR 1K3 GAUGING STATION

S. No.	YEAR	Annual River flow(Q_i) cumecs	$K_i=\frac{Q_i}{\bar{Q}}$ (inarray)	$K_i-\hat{\mu}$ $\hat{\mu}=1.000$	$(K_i-\hat{\mu})^2$	ln K_i	lnK_i - ln$\hat{\mu}$ ln$\hat{\mu}$ =0.028	(lnK_i -ln$\hat{\mu}$)^2
1	2	3	4	5	6	7	8	9
1	1956	5420.0	1.4881	0.488	0.238	0.3927	0.3647	0.133
2	1960	4480.0	1.2300	0.230	0.053	0.2852	0.2572	0.066
3	1962	4400.0	1.2081	0.208	0.043	0.1887	0.1607	0.026
4	1968	4300.0	1.1806	0.181	0.033	0.1664	0.1384	0.019
5	1963	4280.0	1.1751	0.175	0.031	0.1614	0.1334	0.018
6	1961	4190.0	1.1504	0.150	0.023	0.1398	0.1118	0.013
7	1964	4100.0	1.1257	0.126	0.016	0.1185	0.0905	0.008
8	1970	4080.0	1.1202	0.120	0.014	0.1133	0.0853	0.007
9	1955	4000.0	1.0982	0.098	0.010	0.0938	0.0658	0.004
10	1958	3640.0	0.9994	-0.001	0.000001	-0.0006	-.0274	0.00073
11	1972	3560.0	0.9774	-0.023	0.000529	-0.0229	-.0051	0.00026
12	1971	3460.0	0.9500	-0.050	0.002500	-0.0513	-.0233	0.00054
13	1957	3260.0	0.8951	-0.105	0.011025	-0.1108	-.0828	0.0068
14	1965	2780.0	0.7633	-0.237	0.056169	-0.2701	-.2521	0.0586
15	1967	2780-0	0.7633	-0.237	0.056169	-0.2701	-.2421	0.0586
16	1966	2600.0	0.7138	-0.286	0.081796	-0.3372	-.3092	0.0456
17	1969	2440.0	0.6694	-0.33 0	0.108900	-0.4006	-.3726	0.1388
18	1959	1800.0	0.4942	-0.506	0.256036	-0.7048	-.6768	0.4580
TOTAL		65570.0	18.0		1.034	-0.5086		1.109

$$\bar{Q} = \frac{65570.0}{18} = 3642.0 \text{ cumecs}$$

TABLE 4 - 1 - 10

ESTIMATION OF FLOOD FOR RUFIJI RIVER AT STIEGLER'S GORGE 1K3 BY NORMAL, GUMBEL AND LOG NORMAL 2

Formula $Q_{100} = \bar{Q}\,(1+K\sigma)$

Station	$\bar{Q}$ cumecs	K From Chow Book	σ	$(1+K\sigma)$	Q_{100} cumecs	Remarks
	NORMAL DISTRIBUTION					
Rufiji River at Stiegler's Gorge (1K3)	3642	2.33	0.075	1.175	7940	
	GUMBEL DISTRIBUTION					
(1K3)	3642	3.349	0.075	1.251	8190	

LOG NORMAL DISTRIBUTION $Q_{100} = Q\,\text{antilog}\left[\frac{\log\mu + K\log\sigma}{2.3}\right]$

Station	$\bar{Q}$ cumecs	K	$\log_e\mu$	$\log_e\sigma$	$K\log_e\sigma$	Antilog $\frac{\log_e\mu + K\log_e\sigma}{2.3}$	Q_{100} cumecs	Remarks
Rufiji River (1K3)	3642	2.33	-0.028	0.248	0.578	1.173	7900	

(3) Gamma 2 - for f = 1 degree of freedom

$$\chi^2 = \frac{5}{18} \times 104 \times 18 = 10.8 \quad \text{and } P(\chi^2) = 0.991$$

It can be observed from the results, that log-Normal 2 distribution is applicable to this river gauging station (1k3) , since it has the probability of Chi-square less than the commonly used level of significance (0.95). Hence the statistical test of this distribution is non-significant, but for Normal and Gamma 2 distributions, they are significant . Also, the smaller probability of Chi-square, the better fitting to observed data.

4.1.5 Statistical Analysis

The various statistical terms used in this study are as follows:

Mean ($\bar{X}_t$) $= \sum X_t/N$

Variance (S^2) $= (X_t^2 - N\,\bar{X}_t^{\,2}) \times N/N-1$

Standard of Variation "S" = $\sqrt{\text{Variance}}$

Coefficient of Variation (Cv) $= \dfrac{S}{\bar{X}_t}$

First serial auto-correlation coefficient

$$r_1 = \frac{\overline{X_t \cdot X_{t+1}} - \bar{X}_t \bar{X}_{t+1}}{\left[\bar{X}_t^2 - (X_t)^2\right]^{1/2}\left[\bar{X}_{t+1}^2 - (\bar{X}_{t+1})^2\right]^{1/2}}$$

Skewness coefficient

$$C_s = \frac{\sum X_t^3 - 3\bar{X}_t \sum X_t^2 + 2N X_t^3}{N\left(\frac{1}{N}\sum X_t^2 - \bar{X}_t^2\right)^{3/2}}$$

Range 'R' $= S_{max} - S_{min}$

where

$S_{max} = Max\ (X_t - \bar{X}_t)$

$S_{min} = Min\ (X_t - \bar{X}_t)$

Where

X_t is the observed time series

N is the length of the time series.

The above statistical properties have been computed by running the computer program given in Appendix (VIII).

4.1.6 Analysis of Historical Data

For the analysis of time series of the available hydrologic data, all computations have been done by computor, IBM 1620 Model. The mathematical treatment for the required analysis has been discussed earlier. First, the trend component is removed from the composite time series, and then periodic components are removed, thus leaving the random component alone. For the detection and isolation of periodic component correlogram analysis, periodogram analysis and spectral analysis have been used. The procedure for the detection of each component from one another is analysed as below.

4.1.6 Trend Component Analysis

The method of least squares already discussed in Chapter III is used for development of trend. The trend is represented by a straight line. The equation for the trend line is assumed as given by $Y = A+BX$ where the trend constants A and B are given by equations 3.16 and 3.17 respectively. The computation method has been given in the computer programme as per Appendix (iii) the trend line has been shown in Figure 4-1-3.

(a) CorrelationAnalysis

The seria-l dependence of the hydrologic data is found out with the help of this analysis. The serial correlation coefficients are calculated from Eqn 3-28.

The procedure for calculating serial correlation coefficient (r_k) values has been given in the computer programme, appendix (iV) . The correlograms are constructed by plotting r_k values against time lag. The correlogram for Rufiji river at Pangain Falls , has been drawn as per Fig. 4-1-5 to 8.

The brief procedure for this analysis is as follows:

1. Remove the trend component and get YR (programme Appendix () iii)).
2. Find out correlogram with YR with the help of program Appendix (iV).

3. Remove the cycle indicated by the correlogram and then find new values of YR with the help of Computer programme appendix (V).

4. Repeat the operations 2 and 3 for various other harmonics which are required to be removed, till no cycle is indicated by the correlogram.

The results of the above analysis is shown in Fig. 4.1.5 to 8 . The values of YR after removing 12 months, 6 month, 4 month, 3 month and 2 month periods are given in Appendix (IVb).

(b) Periodogram Analysis

The Schuster's periodogram is developed by using the harmonic analysis and magnitude of the squared amplitudes is calculated with the equation 3.22 . These values are plotted against the frequencies for which the amplitude values have been calculated. The periodogram will indicate the periods present in the hydrologic data and the reactive magnitude of various harmonics will be shown by the periodogram as per Fig. 4.1.9.

(c) Power Spectrum

The variance spectrum is developed by calculating the normalised spectral density function. This will give two sided normalised power spectrum. The power spectrum will show peaks with different frequencies and it will give the possible periodicities precisely at its frequency.

The normalised spectral density is given by Equation 3.36.

This is calculated by assuming some values of time lag interval and values of correlation coefficients with the same time lag interval as shown in Appendix (VII) Here the value values of K have been taken as 6 and 120 months . Thus by knowing all the terms, values of SX(f) is found out corresponding to the different values of frequencies. The computor program, for this purpose has been developed as per appendix (VII). Fig. 4.1.11 shows the power spectrum, which shows clearly the presence of 12, 6, 3 months periods.

4.1.6.2 Stochastic Component

The residual series Z_t is subject correlogram and variance spectrum analysis is ascertained that it is free from the presence of any significant sub-harmonics. First order Markov Model has been fitted as earlier discussed, the value r_1 for lag one being significant (r_1=0.6089), the model then fitted becomes

$$Z_t = 0.6089\ Z_{t-1} + R_t$$

Based on this equation the stochastic component has been computed by the help of computor program, appendix (IX).

4.1.6.3 Random Component

The random in the absence of any trend, is obtained, by removing the cyclic and stochastic component from th e original time series as given in Equation. The random

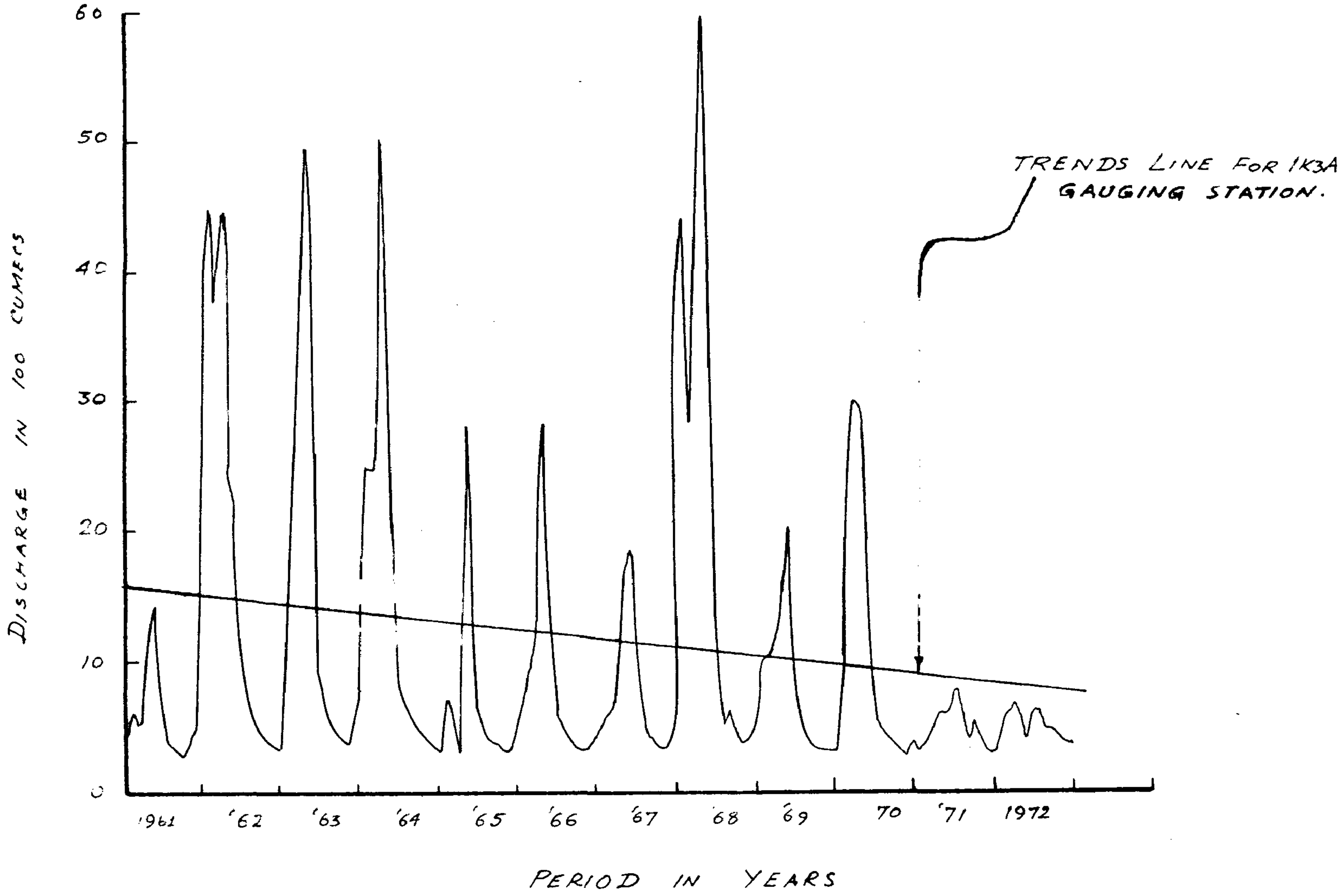

FIG 4-1-3 TREND COMPONENT SEQUENCE OF MONTHLY FLOW FOR IK3A GAUGING STATION

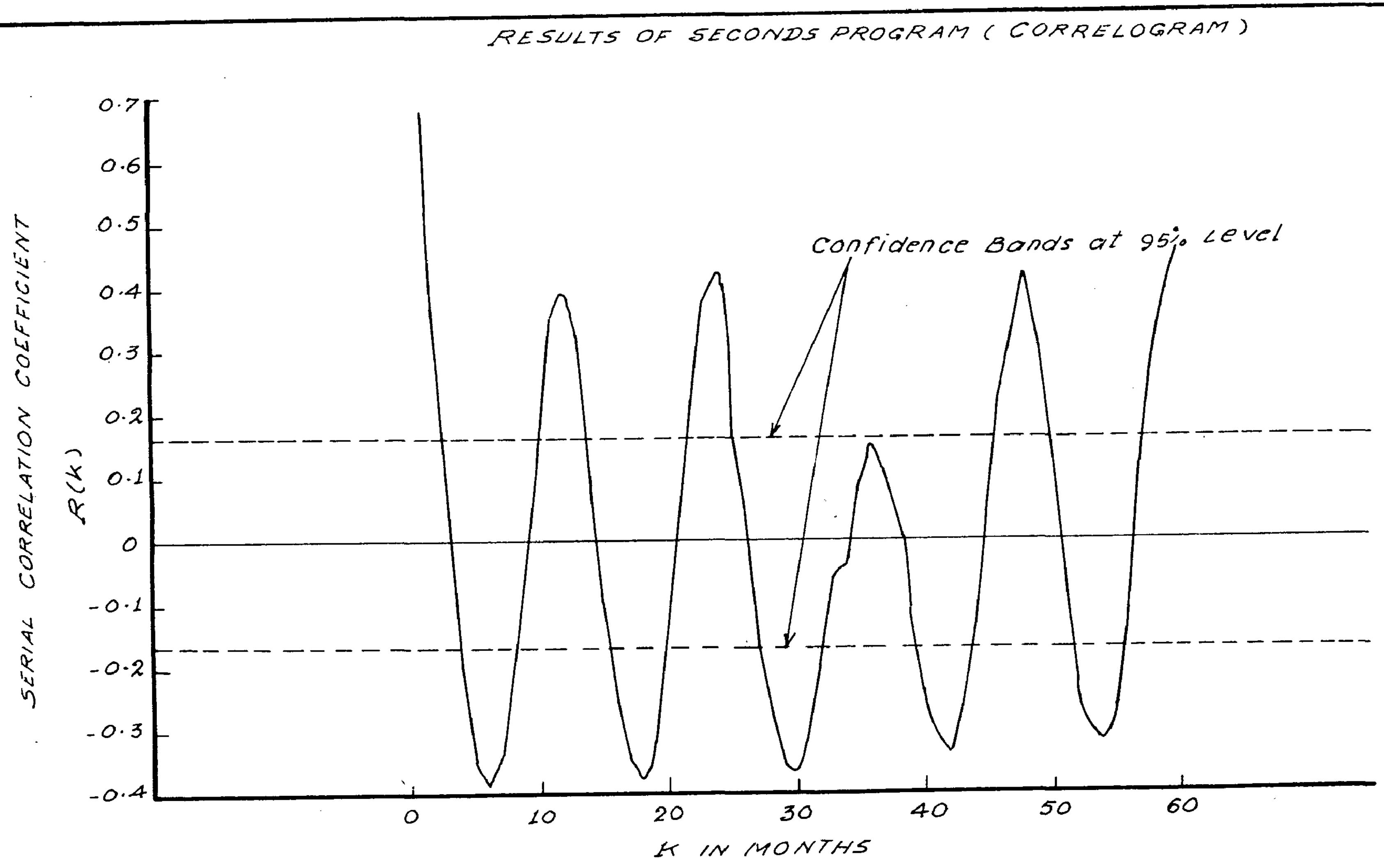

FIG. 4-1-4 CORRELOGRAM CURVE FOR ALL PERIODS PRESENT

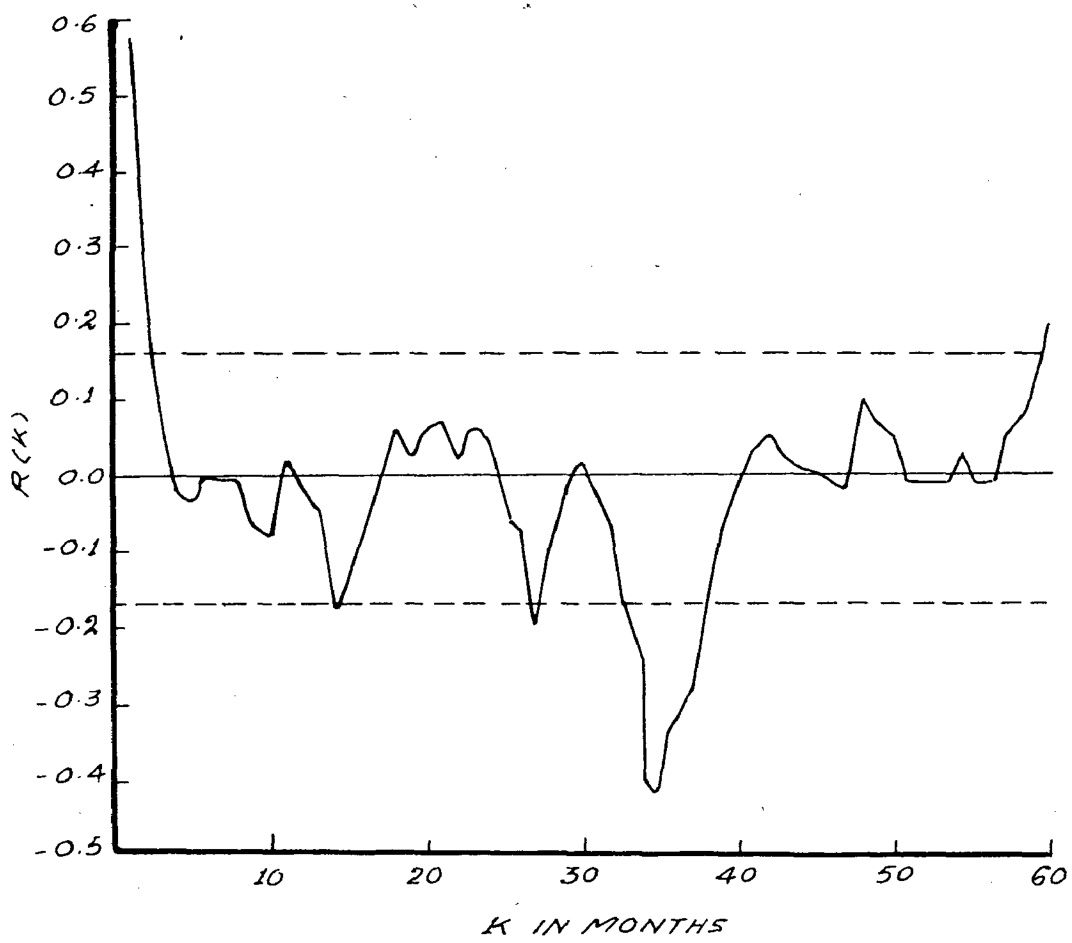

FIG. 4-1-5 CORRELOGRAM AFTER 12 MONTHS REMOVAL

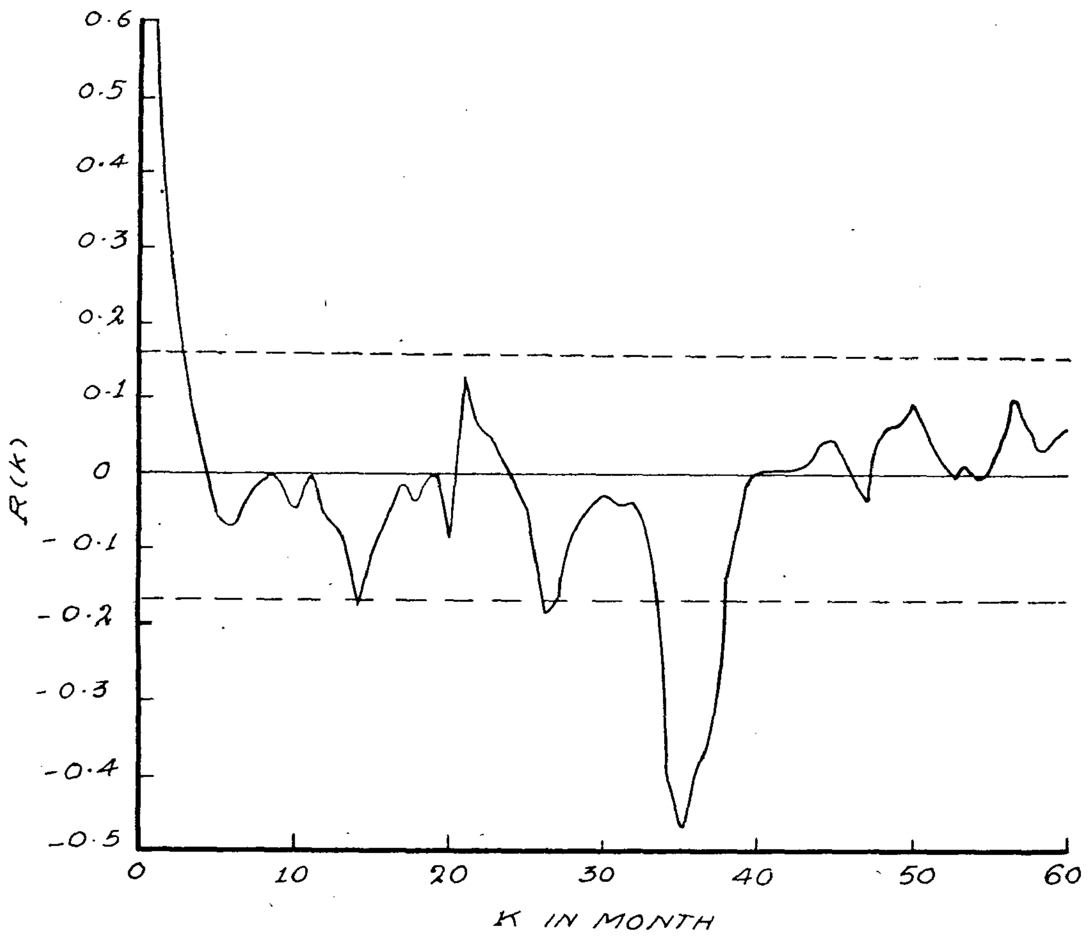

FIG. 4-1-6 CORRELOGRAM AFTER 6 MONTHS REMOVAL

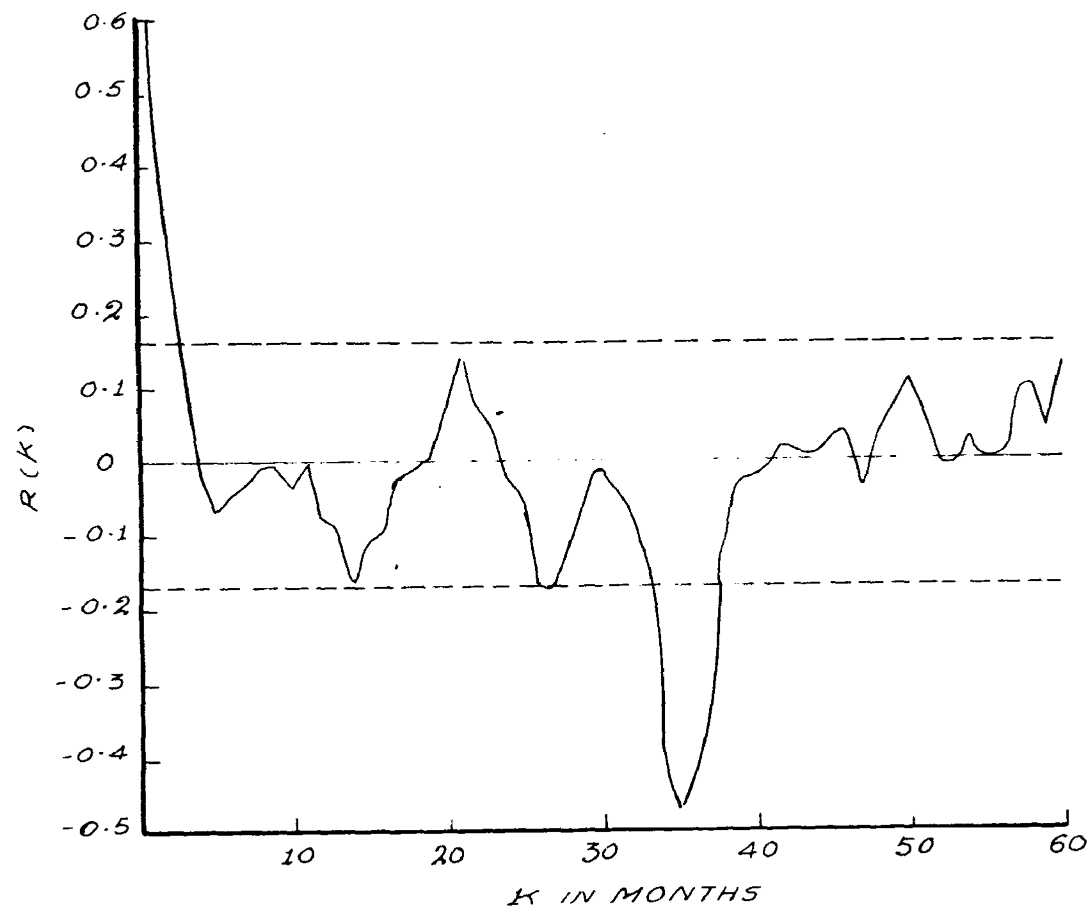

FIG. 4-1-7 CORRELOGRAM AFTER 4 MONTHS REMOVAL

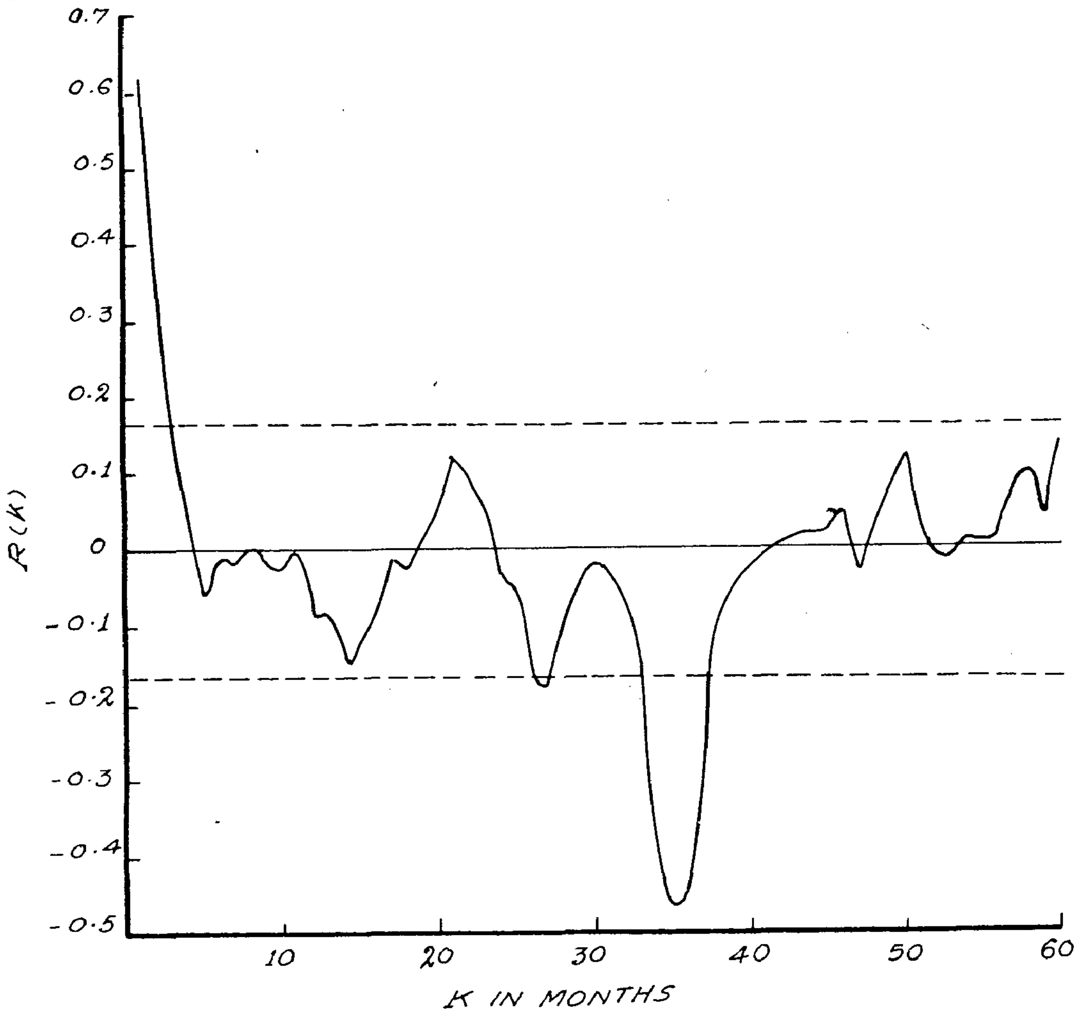

FIG. 4-1-8 CORRELOGRAM CURVE AFTER 3 MONTHS REMOVAL

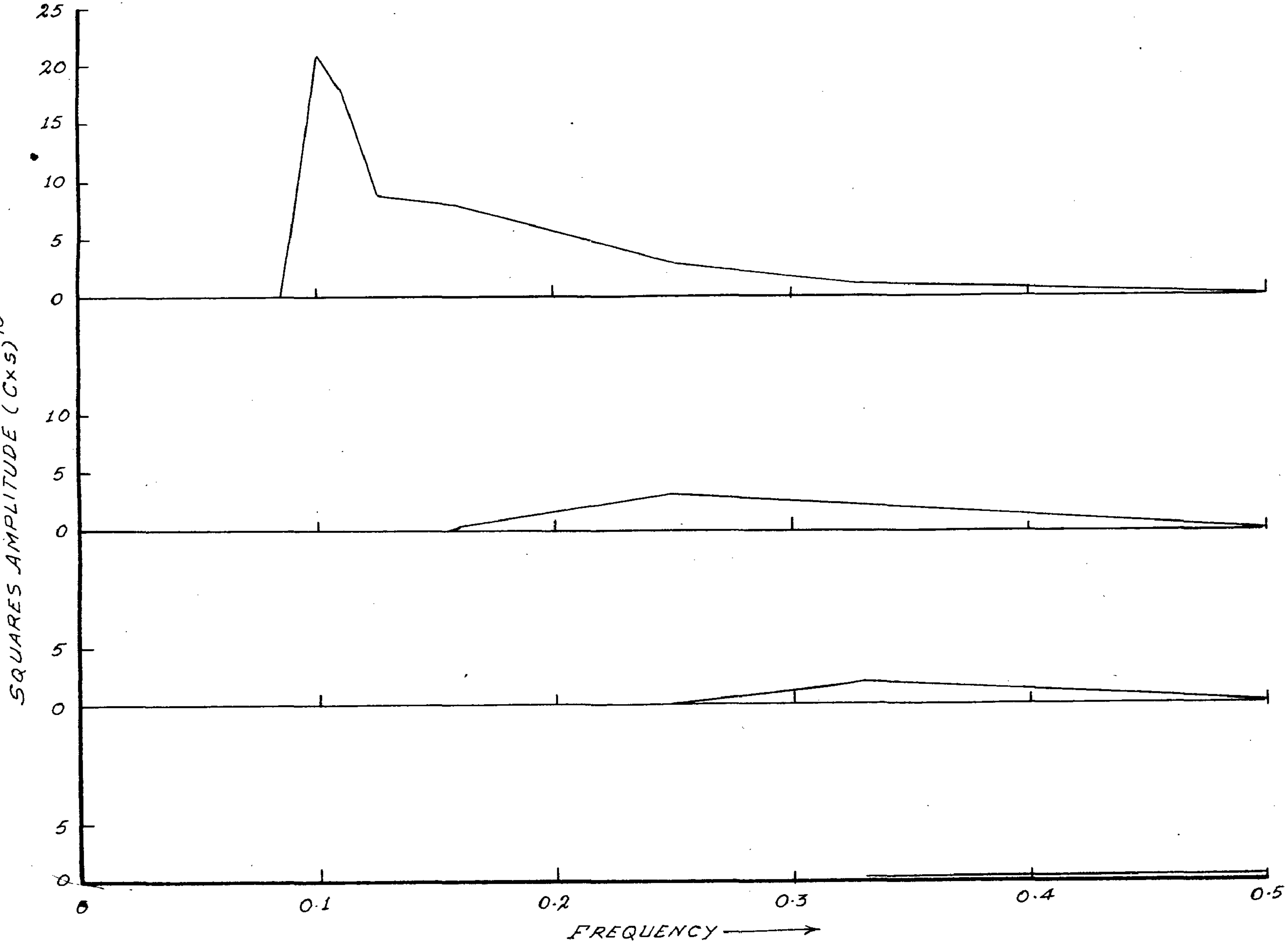

FIG. 4-1-10 PERIODOGRAM ANALYSIS REMOVAL OF 12 6 4 & 3 MONTHS

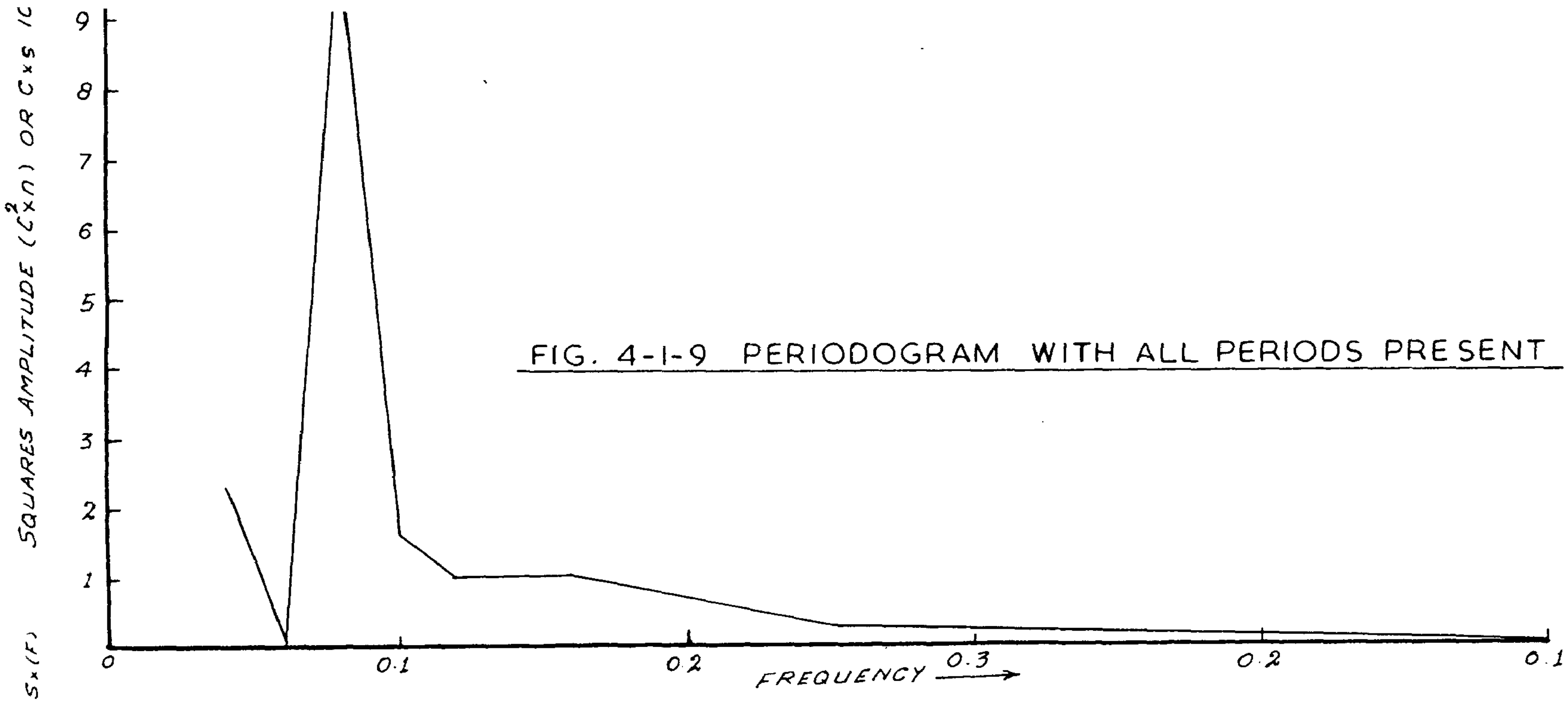

FIG. 4-1-9 PERIODOGRAM WITH ALL PERIODS PRESENT

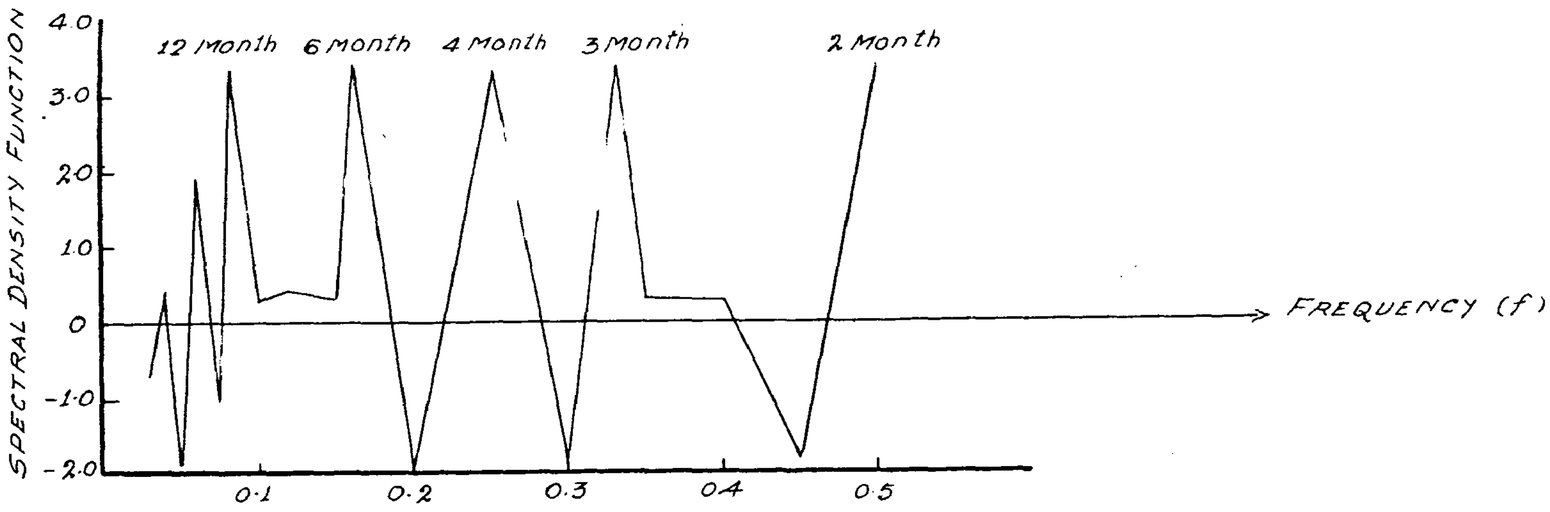

FIG. 4-1-11 NORMALISED POWER SPECTRUM ANALYSIS

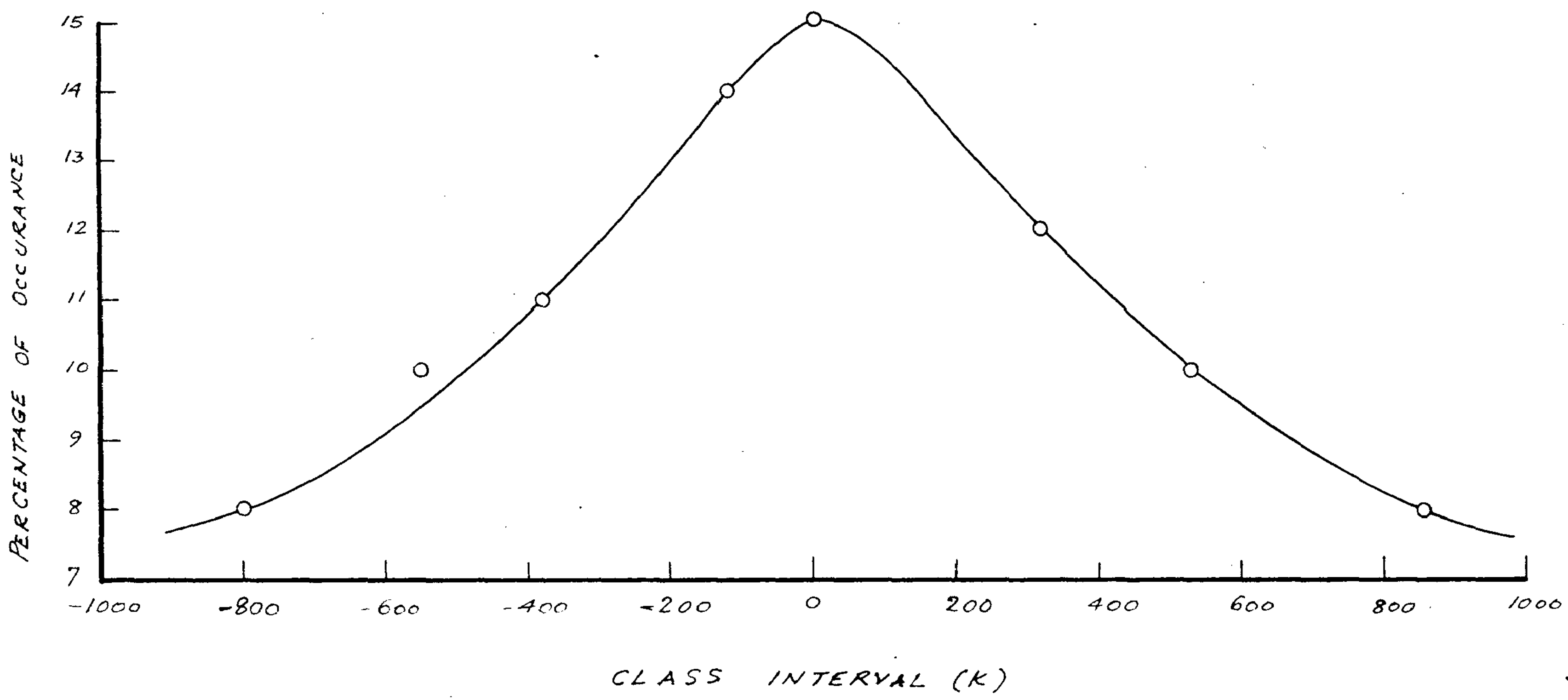

FIG. 4-1-12 NORMALLY DISTRIBUTED RANDOM COMPONENT.

component is obtained (Appendix IXg) by the help of computor program. Computor programme Appendix (IX) has been developed to compute the mean, variance and standard deviation. The series was broken into 8 samples of 30 values and was plotted as given in Fig. 4.1.12. This shows that the random component is normally distributed.

Mathematical Model Adopted for Data Generation

The general model adopted for the generation of monthly stream flow sequences for 1K3A gauging station as discussed earlier is as follows:

$$X_t = 156088 - 6.11t + A_{XN1}\cos\frac{2\pi t}{12} + B_{XN1}\sin\frac{2\pi t}{12}$$
$$+ A_{XN2}\cos\frac{4\pi t}{12} + B_{XN2}\sin\frac{4\pi t}{12}$$
$$+ A_{XN3}\cos\frac{6\pi t}{12} + B_{XN3}\sin\frac{6\pi t}{12}$$
$$+ A_{XN4}\cos\frac{8\pi t}{12} + B_{XN4}\sin\frac{8\pi t}{12}$$
$$+ 0.6089 Z_{t+1} + R_t$$

Values of $A_{XN1} \ldots A_{XN4}$ and $B_{XN1} \ldots B_{XN4}$ are given in Appendix (VId) and Random numbers (R_t) can be generated by programme Appendix (X).

4.2 CONCLUSION

In Chapter II various methods of estimating peak run-off have been discussed, based onthe available data. In

In this connection 1K3 Gauging s-tation has been considered to arrive at the best fit distribution to the observed annual series. In Chapter 3, time series analysis for 1K3A monthly flows has been discussed. As whole, based on this conclusion, the following inferences can be stated:

(1) Normal distribution is not fitting to the observed annual flows as the deviation of the observed series is too short.

(2) Also Gamma 2 distribution does not fit well close to too short annual series.

(3) Log-normal 2 and Gumbel distributions are fitting well to the observed annual series. This is verified as in Figures 4-2-1 and 4-2-2-.

(4) The trend, periodic stochastic and random components, all constituting the composite time series have been isolated.

(5) The trend indicated in the hydrologic series is decreasing at a rate of 5-95 cumecs/month from 1961-1972 , Fig. 4-1-3.

(6) The periodic component in monthly run off series has been described by Fourier series with 12 months fundamental cycle and its harmonics as 6,4,3 and2 mohths. The 3 and 2 months are less predominant comparatively to 12 and 6 months Fig. 4-1-10.

(7) The correlogram for monthsly run-off has been constructed which tends to reach its maximum values at lags, Figs. 4.1.4 to 8.

(8) Power spectrum shows clear peaks at 12,6,4 and 3 mohths at its frequencies. Fig. 4.1.11 . The presence of peaks at 6,4,3 months periods are more easily seen in the spectrum, than on the correlogram. The presence of these sub-harmonics is indicative of the effects of oscillatory seasonal effects in the record.

(9) Both the correlogram and power spectrum are useful and should be used simulatneously for analysis of hydrologic data. The spectral analysis complements the correlogram analysis is detecting periodicity in the hydrologic time series.

4.3 SCOPE OF FURTHER STUDY

(1) The flood frequency analysis of annual series can be conducted using other probability distribution function, like log-normal 3 and Gamma 3. The study also can be conducted to other rivers, in different weather regimes.

(2) The model, and its efficiency can be tested by comparing its performance with results obtained by other time series models.

(3) The model has been developed for monthly river flows. This also can be applied to time series of 10 daily weekly and yearly seriesl.

REFERENCES

1. Blackman, R.B. and Takey J.W. "The Measurement of Power Spectra" Dover, New York, 1958.

2. Bonne, J. "Stochastic simulation of monthly stream flow by Multiple regression model utilising precipitation data". Jnl. Hydrology Vol.12, No.4, March,1971.

3. Brittan, M.R., " Probability Analysis applied to the development of synthetic hydrology for Colorado River", Bureau of Economic Research, University of Colorodo, Boulder, Colorado, October, 1961.

4. Central Water and Power Commission,"Estimation of Design Flood Recommended Procedures"- New Delhi, September, 1972.

5. Chow, V.T., " Handbook of Applied Hydrology". McGraw-Hill Book Company, New York, 1964.

6. Chow V.T. and Ramaseshan, S. " Sequential generation of Rainfall and Runoff data". Jnl. Hydraulic Div. ASCE 91, New York-4, 1965.

7. Hamnan, E.E." The statistical analysis of Hydrological time Series" Proc. National Symp. Water Resources USC, Management, Australian Acad. Science, Canberra, Australia, Sept. 1963.

8. Herbert A. and Raymod R.C. "Statistical Methods", Barues and Noble College Qutline Series No.27, 1969.

9. Jenkins, G.N. and Wates, D.G. "Spectral Analysis and its Applications" Holden Day 1968.

10. Jullian P.R. " Variance Spectrum Analysis" American Geophysical Union 1966.

11. Kenny and Keeping "Mathematics Statistics" Van Nostrand, East West Press, 1964.

12. Kisiel, C.C. " Time Series Analysis of Hydrologic Data" in Advances in Hydrosciences, edited by Chow, V.T. Vol. 5, 1969.

13. Langbein W.B. "Queueing Theory and water storage" Jnl Hydrology, water Resources Research 3(4) 1967.

14. Leo, R.B. "Statistical Methods in Hydrology" USAE District Sacramento, California, January 1962.

15. Linsley, Kohler and Paulhus,"Hydrology for Engineers" McGraw Hill Book Company, New York, 1958.

16. Linsley, Kohler and Paulhus, " Applied Hydrology" McGraw Hill Book Company Inc, N.Y. (1949).

17. Nemer, J. "Engineering Hydrology" Tata McGraw Hill Publ. Company Ltd, New Delhi, 1973.

18. Quimpo, R.G. "Stochastic model of daily river flow Sequencies" Hydrology Paper No.18, Colorado State University, Fort Collins, Colorado, Feb, 1967.

19. Radmilo, D. Markovie " Probability Functions of Bes t Fit to Distributions on Annual Precipitation and Runoff", Hydrelogy Paper No.8, Colorado State Univ. Fort Collins Colorado, August, 1965.

20. Roesuer, L.A. and Yevjevich, V.M. "Mathematical Models for Time series of monthly precipiation and monthly runoff". Hydrology Paper No. 15 Colorado State University Fort Collins, Colorado, October, 1966.

21. "Statistical Methods in Hydrology" Proc. of Hydrology Symp. No.5 held at McGraw University.

22. Thomas and Fiering." Mathematical Synthesis of Stream Flow sequencies for the analysis of River Basins by Simulation" Havard University Press. Cambridge Massachussetts, 1962.

23. Todorovic, P.& Yevjevich V.M. "Stochastic Process of Precipitation" Hydrology Paper No.35, Colorado Station University, Fort Collins, Colorado Sept, 1969.

24. Varshney, R.S. " Engineering Hydrology" 1974.

25. Yevjevich, V.M. "Stochastic in Geophysical and Hydrological Times Series" Nordic Hydrology. Vol.2, No.4, 1971.

26. Yevjevich, V.M. " Probability and Statistics in Hydrology" Water Resources Publications Fort Collins, Colorado USA,1972.

27. Wisler G.O. and Brater E.F. " Hydrology" Wiley International Edition John Wiley and Sons Inc. (1959).

APPENDIX I

```
C  C  LUHUMBIKA ,  HYDROLOGY DESSERTATION WORK
C     COEFFICIENT OF CORRELATION R AND A,B
      DIMENSION Q(150),H(150),QL(150),HL(150)
  100 READ 10,N,AS
   10 FORMAT(I4,F10.4)
      AN=N
      DO 20 I=1,N
      READ 30,Q(I),H(I)
   20 CONTINUE
      SUMQ=0.0
      SUMH=0.0
      DO 40 I=1,N
      QT=Q(I)
      HT=H(I)-AS
      QL(I)=LOGF(QT)
      HL(I)=LOGF(HT)
      SUMQ=SUMQ+QL(I)
      SUMH=SUMH+HL(I)
   40 CONTINUE
      QBAR=SUMQ/AN
      HBAR=SUMH/AN
      PUNCH55,QBAR,HBAR
   55 FORMAT(14HVALUE OF QBAR=,F10.4,4X,14HVALUE OF HBAR=,F10.4)
      SUMXY=0.0
      SUMXX=0.0
      SUMYY=0.0
      DO 50 I=1,N
      X=HL(I)-HBAR
      Y=QL(I)-QBAR
      SUMXY=SUMXY+X*Y
      SUMXX=SUMXX+X*X
      SUMYY=SUMYY+Y*Y
   50 CONTINUE
      A=SUMXY/SUMXX
      B=QBAR-A*HBAR
      R=SQRTF(SUMXX*SUMYY)
      R=SUMXY/R
      PUNCH 60,A,B,R
   60 FORMAT(12HVALUE OF A =,F10.4,4X,12HVALUE OF B =,F10.4,4X,
     1 12HVALUE OF R =,F10.4)
      PUNCH 70,A,B,R
   70 FORMAT(3E20.8)
      PUNCH 80
   80 FORMAT(17HVALUE OF LOG(H-A))
      PUNCH 30,(HL(I),I=1,N)
   30 FORMAT(8F10.4)
      PUNCH 90
   90 FORMAT(15HVALUE OF LOG(Q))
      PUNCH 30,(QL(I),I=1,N)
      GO TO 100
      END
```

```
C  LUHUMBIKA, HYDROLOGY DESSERTATION WORK                    APPENDIX II
   COEFFICIENT OF CORRELATION R AND A,B,IK3ABAR AND IK3BAR
   DIMENSION QIK3A(300),QIK3(300),QAL(300),QL(300)
00 READ 10,N
10 FORMAT(I4,2F10.4)
   AN=N
   DO 20 I=1,N
20 READ 10,J,QIK3A(J),QIK3(J)
   SUMQA=0.0
   SUMQ=0.0
   DO 40 I=1,N
   QAT=QIK3A(I)
   QT=QIK3(I)
   QAL(I)=LOGF(QAT)
   QL(I)=LOGF(QT)
   SUMQA=SUMQA+QAL(I)
40 SUMQ=SUMQ+QL(I)
   QABAR=SUMQA/AN
   QBAR=SUMQ/AN
   PUNCH 30,QABAR,QBAR
30 FORMAT(9HIK3ABAR =,F10.4,5X,9HIK3 BAR =,F10.4)
   SUMXY=0.0
   SUMXX=0.0
   SUMYY=0.0
   DO 50 I=1,N
   Y=QAL(I)-QABAR
   X=QL(I)-QBAR
   SUMXY=SUMXY+X*Y
   SUMXX=SUMXX+X*X
50 SUMYY=SUMYY+Y*Y
   A=SUMXY/SUMXX
   B=QABAR-A*QBAR
   R=SQRTF(SUMYY*SUMXX)
   R=SUMXY/R
   PUNCH 60,A,B,R
60 FORMAT(3HA =,F10.4,3HB =,F10.4,3HR =,F10.4)
   PUNCH 70,A,B,R
70 FORMAT(3E20.8)
   PUNCH 80
80 FORMAT(3HNO.2X9HLOG(IK3A)2X8HLOG(IK3))
   DO 90 I=1,N
90 PUNCH 10,I,QAL(I),QL(I)
   GO TO 100
   END
```

APPENDIX III

```
**
C   C LUHUMBIKA DISSERTATION WORK HYDROLOGY
C       PROGRAMME FOR SEPERATION OF TREND COMPONENT
      DIMENSION X(400),Y(400),YT(400),YR(400)
       READ 50,N
50     FORMAT(I5)
       READ 60,(X(I),I=1,N)
60     FORMAT(12F6.2)
       READ 60,(Y(I),I=1,N)
       SUMX=0.
       SUMY=0.
       SUMXX=0.
       SUMXY=0.
       DO 25 I=1,N
       SUMX=SUMX+X(I)
       SUMY=SUMY+Y(I)
       XX=X(I)*X(I)
       SUMXX=SUMXX+XX
       XY=X(I)*Y(I)
       SUMXY=SUMXY+XY
25     CONTINUE
       AN=N
       DENOM=AN*SUMXX-SUMX*SUMX
       A=(SUMY*SUMXX-SUMX*SUMXY)/DENOM
       B=(AN*SUMXY-SUMX*SUMY)/DENOM
       X1=0.
       DO 29 I=1,N
       X1=X1+1.
       YT(I)=A+B*X1
       YR(I)=Y(I)-YT(I)
29     CONTINUE
       PUNCH 11
11     FORMAT(15X,50HRESULTS OF VALUES TREND COMPONENT YT AND YR VALUES)
       PUNCH 12
12     FORMAT(21X,38HVALUES OF TREND COMPONENT YT SEPERATED)
       PUNCH 10,(YT(I),I=1,N)
10     FORMAT(6F12.4)
       PUNCH 14
14     FORMAT(2X,53HVALUES YR AFTER REMOVING TREND COMPONENT YT FROM Y(I)
      1)
       PUNCH 10,(YR(I),I=1,N)
       PUNCH 30,A
30     FORMAT(10X,11HVALUE OF A=,F15.5)
       PUNCH 31,B
31     FORMAT(10X,11HVALUE OF B=,F15.5)
       STOP
       END
```

(iv) APPI

```
**
C  C  LUHUMBIKA /HYDROLOGY/DISSERTATION WORK
C     PROGRAMME FOR CORRELOGRAM
C     CALCULATION OF K AND R(K)
     DIMENSION X(400),R(400),KR(400)
      READ 5,N,K1,K2
5     FORMAT(14I5)
      READ 900,(X(L),L=1,N)
900   FORMAT (6F12.4)
      IK=0
      DO 10 K=K1,K2,K1
      XZ=0.
      X1=0.
      X2=0.
      X3=0.
      X4=0.
      NK=N-K
      DO 20 I=1,NK
      J=I+K
      XZ=XZ+X(I)
      X1=X1+X(J)
      X2=X2+X(I)*X(I)
      X3=X3+X(J)*X(J)
      X4=X4+X(I)*X(J)
20    CONTINUE
      ANK=NK
      A=SQRTF(ANK*X2-XZ*XZ)
      B=SQRTF(ANK*X3-X1*X1)
      R(K)=(ANK*X4-XZ*X1)/(A*B)
      IK=IK+1
      KR(IK)=K
10    CONTINUE
      PUNCH 5,(KR(I),I=1,IK)
      PUNCH 900,(R(K),K=K1,K2,K1)
      END
```

```
:  LUHUMBIKA /HYDROLOGY/DISSEI TATION WORK                APPENDIX (V)
   PROGRAMME FOR ISOLATION OF PERIODIC COMPONENT
   THAT IS PROGRAMME FOR ESTIMATION OF SINNSOIDAL COMPONENT
   DIMENSION YR(400)
   READ50,N
   FORMAT(I5)
   READ 11,(YR(I),I=1,N)
   FORMAT(6F12.4)
   READ 12,M
   FORMAT(I5)
   PUNCH29,M
   FORMAT(6X,12HVALUES OF M=,,5)
   PIE=3.1416
   AN=M
   AXN=0.
   BXN=0.
   DO 13 I=1,N
   AI=I
   X=2.0*PIE*AI/AN
   N RIFFERED BELOW IS SMALL I. IN 2*PIE*(SMALL)N(SMALL)T/T
   FOR TIME PERIOD T=12
   N=1 FOR REMOVING 12 MONTHS PERIOD
   N=2 FOR REMOVING 6 MONTHS PERIOD
   N=3 FOR REMOVING 4 MONTHS PERIOD
   N=4 FOR REMOVING 3 MONTHS PERIOD
   SO HERE FOR T=180,N=15,30,45,60 RESPECTIVELY
   HENCE N=T/(SMALL)N=12,6,4,3 RESPECTIVELY
   BN=N
   CN=BN/2.
   AXN=AXN+YR(I)*COSF(X)/CN
   BXN=BXN+YR(I)*SINF(X)/CN
   CONTINUE
   CXN=SQRTF(AXN*AXN+BXN*BXN)
   THETA=ATANF(AXN/BXN)
   PUNCH 21
   FORMAT(5X,3HAXN,10X,3HBXN,10X,3HCXN,9X,5HTHETA)
   PUNCH 20,AXN,BXN,CXN,THETA
   FORMAT(4F13.2)
   DO 100 I=1,N
   AI=I
   X=2.0*PIE*AI/AN
   YR(I)=YR(I)-(AXN*COSF(X)+BXN*SINF(X))
10 CONTINUE
   PUNCH  61
   FORMAT(18X,43HYR VALUES AFTER REMOVAL OF     MONTHS PERIOD)
   PUNCH 60,(YR(I),I=1,N)
   FORMAT(6F12.4)
   GO TO 15
   END
```

**

```
C   C   LUHUMBIKA /HYDROLOGY/DISSERTATION WORK                    APPENDIX (VI)
C       PROGRAMME FOR PERODOGRAM ANALYSIS
C       CALCULATION OF AXN,3XN,CXS,FREQUENCY(F)
        DIMENSION YR(200)
        READ 50,N
50      FORMAT(I5)
        READ11,(YR(I),I=1,N)
11      FORMAT(6F12.4)
15      READ12,M
12      FORMAT(I5)
        PIE=3.1416
        AN=M
        AXN=0.
        BXN=0.
        DO13 I=1,N
        AI=I
        X=2.0*PIE*AI/AN
        BN=N
        CN=BN/2.
        AXN=AXN+YR(I)*COSF(X)/CN
        BXN=BXN+YR(I)*SINF(X)/CN
13      CONTINUE
        F1=M
        F=1./F1
        CXS=AXN*AXN+BXN*BXN
        PUNCH7
7       FORMAT(6X,3HAXN,12X,3HBXN,12X,3HCXS,14X,1HF)
        PUNCH6,AXN,BXN,CXS,F
6       FORMAT(4F15.4)
        GOTO 15
        END
```

APPENDIX VII

```
C  LUHUMBIKA, HYDROLOGY DESSERTATION WORK
   PROGRAMME FOR ANALYSIS OF POWER SPECTRUM
   DIMENSION R(30),VD(30),F(30)
   READ 1,N,K1,M
   FORMAT(3I5)
   READ2,(R(I),I=1,N)
   READ2,(F(K),K=1,M)
   DO19K=1,M
   VD(K)=0.
   II=0
   DO20I=1,N
   II=II+K1
   AK=II
   X1=R(I)*COSF(2.*3.1416*F(K)*AK)/3.1416
   VD(K)=VD(K)+X1
)  CONTINUE
)  CONTINUE
   PUNCH 6
   FORMAT(39X,11HVALUES OF F)
   PUNCH2,(F(K),K=1,M)
   PUNCH 7
   FORMAT(39X,12HVALUES OF VD)
   DO 50 I=1,M
50 VD(I)=VD(I)/10.0**8
   PUNCH3,(VD(K),K=1,M)
   FORMAT(6F10.5)
 3 FORMAT(4F20.8)
   END
```

```
C   C  LUHUNJIKA /HYDROLOGY/DISSERTATION WORK                          APPENDIX(X)(ii)
C      PROGRAMME FOR FINDING OUT STATISTICAL PARAMETERS
C      MEAN,STANDARD DEVIATION,VARIANCE AND AND C.V.
C      DATA USED Y(I),YR(1),YR(2),YR(3),YR(4),YR(5),YR(6)
       DIMENSION YR(600)
       READ 10,N
10     FORMAT(I5)
100    READ 11,(YR(I),I=1,N)
11     FORMAT(6F12.6)
       SUM1=0.0
       DO 44 I=1,N
       SUM1=SUM1+YR(I)
44     CONTINUE
       AN=N
       AMEAN=SUM1/AN
C      THAT  IS AMEAN=MEAN= SIGMA(X)/N
       PUNCH 13,AMEAN
13     FORMAT(10X,17HMEAN OF THE DATA=,E16.8)
       SUM2=0.0
       DO 14 I=1,N
       SUM2=SUM2+(YR(I)-AMEAN)**2
C      THAT IS   SUM2=(SIGMA(X-BAR(X)))**2
14     CONTINUE
       SD=SQRTF(SUM2/(AN-1.0))
C      THAT IS   SD=STANDARD DEVIATION=(SIGMA(X-BAR(X)))**2/N-1
       PUNCH 25,SD
25     FORMAT(10X,31HSTANDARD DEVIATION OF THE DATA=,F12.4)
C      VARIANCE = SQUARE OF STANDARA DEVIATION
       VAR=SD**2
       PUNCH 15,VAR
15     FORMAT(10X,17HVARIANCE OF DATA=,E16.8)
C      C.V.=COEFFICIENT OF VAIIATION
       CV=100.*SD/AMEAN
       PUNCH 16,CV
16     FORMAT(10X,14HC.V. OF DATA =,E16.8)
       GO TO 100
       END
```

(ix)

```
C  C  LUHUMBIKA-HYDROLOGY-DISSERTATION WORK
C     SEPERATION OF STOCHASTIC AND RANDOM COMPONENTS
      DIMENSION YR(400),ST(400),RD(400)
      READ 10,N
   10 FORMAT(I5)
      READ 12,A1
   12 FORMAT(3F10.4)
      READ 11,(YR(I),I=1,N)
   11 FORMAT(6F12.4)
      DO 22 I=2,N
      K=I-1
      ST(K)=A1*YR(K)
      RD(K)=YR(I)-ST(K)
   22 CONTINUE
      PUNCH 13
   13 FORMAT(25X,30HVALUES OF STOCHASTIC COMPONENT)
      L1=N-1
      PUNCH 14,(ST(M),M=1,L1)
   14 FORMAT(6F12.4)
      PUNCH 15
   15 FORMAT(27X,26HVALUES OF RANDOM COMPONENT)
      PUNCH 14,(RD(M),M=1,L1)
      STOP
      END
```

```
  C   LUHUMBIKA, HYDROLOGY DESSERTATION WORK                         APPENDIX(X)
      DIMENSION Z(1200),T(1200)
      READ 10,X
0     FORMAT(F12.4)
      READ 48,N
8     FORMAT(I5)
      DO 100 I=1,N
      CALL RANDOM(X)
      RANDOM(X) SUBROUTINE AVAILABLE WITH COMPUTER CENTRE IN PACK (4)
      Z(I)=X
00    CONTINUE
      PUNCH 11
1     FORMAT(22X,36HUNIFORMLY DISTRIBUTED RANDOM NUMBERS,///)
      PUNCH 20,(Z(I),I=1,N)
0     FORMAT(4E16.8)
      READ 50,(Z(I),I=1,N)
0     FORMAT(4(F12.8,4X))
      PIE=3.1415
      DO 30 I=1,N,2
      ARGMT= 2.0*PIE*Z(I+1)
      XI=Z(I)
      AA=-2.0*LOGF(XI)
      BB=((AA)**0.5)
      T(I)=BB*COSF(ARGMT)
      T(I+1)=BB*SINF(ARGMT)
0     CONTINUE
      PUNCH 22
2     FORMAT(23X,31HGENERATED NORMAL RANDOM NUMBERS,///)
      PUNCH 21,(T(I),I=1,N)
1     FORMAT(4E16.8)
      END
```

RESULTS OF VREGMGT$REND COMPONENT YT AND YR VALUES

VALUES OF TREND COMPONENT YT SEPERATED

1554.7275	1548.9684	1542.4092	1536.2501	1530.0909	1523.9318
1517.7727	1511.6135	1505.4564	1499.2952	1493.1361	1406.9769
1480.8170	1474.6587	1468.4995	14[illegible]2.3404	1456.1812	1490.0221
1443.0690	1437.7030	1431.5447	14[illegible]5.3855	1419.2264	1413.0672
1406.9081	1400.7490	1394.5890	13[illegible]8.4307	1382.2715	1376.1124
1369.9533	1363.7941	1357.6350	1351.4758	1345.3167	1339.1575
1332.9934	1326.8393	1320.6801	1314.5210	1308.3618	1302.2027
1296.0436	1289.8844	1283.7253	1277.5661	1271.4070	1265.2478
1259.0887	1252.9296	1246.7704	1240.6113	1234.4521	1228.2930
1222.1339	1215.9747	1209.8156	12[illegible]3.6564	1197.4973	1191.3381
1105.1790	1179.0199	1172.0607	11[illegible]6.7016	1160.5424	1154.3833
1148.2242	1142.0650	1135.9059	1129.7467	1123.5876	1117.4284
1111.2693	1105.1102	1098.9510	1092.7919	1086.6327	1080.4736
1074.3145	1068.1553	1061.9962	1055.8370	1049.6779	1043.5187
1037.3596	1031.2005	1025.0413	1018.8822	1012.7230	1006.5639
1000.4048	994.2456	988.0865	981.9273	975.7682	969.6090
963.4499	957.2908	951.1316	944.9725	938.8133	932.6542
926.4951	920.3359	914.1768	908.0176	901.8585	895.6993
889.5402	883.3811	877.2219	871.0628	864.9036	858.7445
852.5854	846.4262	840.2671	834.1079	827.9488	821.7896
815.6305	809.4714	803.3122	797.1531	790.9939	784.8348
778.6757	772.5165	766.3574	760.1982	754.0391	747.8799
741.7208	735.5617	729.4025	723.2434	717.0842	710.9251

VALUES YR NAMELY VALUES OF Y(I) AFTER REMOVING TREND COMPONENT YT FROM

-1104.7275	31.4316	-942.4092	-996.2501	-650.0909	-103.9318
-857.7727	-1071.6135	-1140.4544	-1199.2952	-1208.1361	-1166.9769
2989.1822	2295.3413	2981.5005	2317.6596	743.8188	-575.0221
-813.8630	-987.7039	-1046.5447	-925.3855	-634.2264	-83.0672
1193.0919	2284.2510	3575.4102	12[illegible]1.5693	-502.2715	-716.1124
-869.9533	-943.7941	-977.6350	-1011.4758	-1015.3167	1140.8425
1147.0016	3673.1607	2459.3199	725.4790	-498.3618	-702.2027
-786.0436	-829.8044	-893.7253	-977.5661	-721.4070	-565.2478
-579.0337	-932.9296	1533.2296	939.3887	-634.4521	-788.2930
-822.1339	-835.9747	-859.8156	-893.6564	-787.4973	-531.3381
-205.1790	150.9801	1627.1393	1[illegible]3.2984	-550.5424	-674.3833
-748.2242	-882.0650	-825.9059	-779.7467	1356.4124	-567.4284
-551.2693	-425.1102	561.0490	747.2081	-136.6327	-540.4736
-654.3145	-688.1553	-741.9962	-615.8370	-569.6779	3306.4813
1802.6404	3708.7993	4569.9587	2781.1178	317.2770	-506.5639
-390.4048	-504.2456	-570.0865	-681.9273	-655.7682	30.3910
116.5501	662.7092	1188.8684	-234.9725	-498.8133	-552.6542
-596.4951	-620.3359	-644.1768	-5[illegible]8.0176	108.1415	1844.3007
2210.4598	2056.6189	362.7781	-2[illegible]1.0628	-404.9036	-458.7445
-492.5854	-536.4262	-560.2671	-5[illegible]4.1079	-447.9488	-371.7896
-175.6305	-229.4714	-15.3122	-27.1531	-304.9939	-384.8348
-251.6757	-422.5165	-456.3574	-370.1982	-184.0391	-57.8799
-191.7208	-335.5617	-98.4025	-111.2434	-231.0842	-225.9251

VALUE OF A= 1560.00660

VALUE OF B= -6.15914

IV.b

RESULTS

CORRELOGRAM

CALCULATION OF K AND R(K) FOR K1=1 AND K2=60

1	2	3	4	5	6	7	8	9	10	11	12	13	14
15	16	17	18	19	20	21	22	23	24	25	26	27	28
29	30	31	32	33	34	35	36	37	38	39	40	41	42
43	44	45	46	47	48	49	50	51	52	53	54	55	56
57	58	59	60										

.6838	.3741	.0468	-.2007	-.3543	-.3924
-.3415	-.2125	-.0474	.1393	.3483	.3869
.3085	.0755	-.1035	-.2625	-.3403	-.3771
-.3159	-.1510	.0512	.2188	.3603	.4207
.3139	.0463	-.1599	-.2839	-.3594	-.3650
-.3157	-.1963	-.0617	-.0355	.0812	.1583
.1061	.0313	-.1134	-.2416	-.3208	-.3358
-.2752	-.1320	.0421	.2181	.3210	.4239
.3324	.1592	-.0643	-.2282	-.3258	-.3247
-.2650	-.1068	.1215	.2912	.3640	.4643

RESULTS

CALCULATION OF K AND R(K) FOR K1=1 AND K2=120

6	12	18	24	30	36	42	48	54	60	66	72	78	84
90	96	102	108	114	120								

-.3924	.3869	-.3771	.4207	-.3650	.1583
-.3358	.4239	-.3247	.4643	-.3457	.6636
-.3423	.1575	-.3406	.3255	-.2682	.0236
-.2402	.1842				

VALUES OF M= 12

AXN	BXN	CXN	THETA
336.53	952.16	1009.69	.33

YR VALUES AFTER REMOVAL OF MONTHS PERIOD

-1872.2623	-961.4395	-1694.1763	-1652.5797	-634.7191	232.6124
-90.2330	-75.7411	-188.7898	-542.9713	-1023.5152	-1503.5281
2221.6378	1302.4675	2029.2383	1061.3413	559.2071	-238.4638
-46.3137	5.1713	-94.7850	-269.0728	-449.6210	-419.6324
425.5378	1291.3745	2623.1530	595.2623	-686.8735	-379.5401
-102.3943	49.0837	-25.6802	-395.1744	-830.7230	804.2633
379.1379	2680.2815	1507.6676	69.1832	-682.9453	-365.6164
-18.4750	162.9962	58.6245	-221.2759	-536.8298	-901.6410
-1346.6621	-1925.0115	631.4623	-316.6957	-819.0220	-451.6927
-54.5556	156.9066	62.7293	-237.3775	-602.9346	-867.9453
-972.7620	-841.5045	674.9569	-592.9748	-735.0977	-337.7690
19.3636	190.6210	126.1340	-123.4791	1540.9634	-904.0496
-1318.8619	-1417.9975	-391.1884	90.9461	-321.1734	-203.8453
115.2830	304.7334	210.4388	40.4193	-385.1444	3049.8461
1035.0381	2715.9095	3017.1262	2124.8672	132.7506	-169.9217
377.2023	488.6458	374.0435	-29.6822	-471.2493	-306.2582
-651.0618	-330.1836	236.7409	-891.2118	-683.3248	-215.9980
171.1217	372.5582	307.6403	118.2161	292.6497	1507.6376
1442.8382	1063.7234	-589.3445	-957.2909	-589.4005	-122.0743
275.0410	456.4706	591.6530	152.1145	-263.4591	-708.4667
-943.2617	-1222.3696	-967.4298	-683.3699	-409.4762	-48.1506
515.9603	570.3830	495.7578	286.0129	.4359	-394.5710
-919.3117	-1328.4626	-1050.6152	-767.4489	-415.5519	110.7730

VALUES OF M= 6

AXN	BXN	CXN	THETA
-237.28	165.67	289.51	-.96

YR VALUES AFTER REMOVAL OF MONTHS PERIOD

-1897.2790	-1223.7363	-2131.8553	-1627.5609	-572.4214	469.8903
-115.2539	-341.0397	-425.5664	-417.9482	-761.2157	-1266.2528
2196.6127	1040.1671	1792.0642	1686.3686	821.5044	-1.1908
-71.3431	-257.1306	-331.6567	-244.0412	-187.3168	-182.3619
430.5041	1029.0705	2385.6636	580.2980	-424.5656	-142.2720
-127.4322	-213.2220	-262.7470	-330.1343	-568.4182	1041.5269
354.5957	1417.9739	1269.9032	94.2274	-420.6377	-128.3532
-43.5214	-99.3132	-176.5374	-296.2274	-274.5194	-664.5802
-1371.7127	-1168.1227	393.6227	-291.8430	-556.7038	-214.4344
-79.6105	-105.4044	-174.1277	-212.3205	-340.6207	-630.6894
-997.8211	-1104.2194	437.7422	-527.9136	-472.7819	-100.5156
-5.6996	-71.4956	-111.0181	-98.4136	1803.2779	-666.7986
-1343.9294	-1680.3159	-628.1382	.16.0158	-58.8541	33.4031
88.2112	42.4131	-27.1085	65.4932	-122.8233	3287.0922
1009.9621	1453.5875	3380.0613	2149.9453	395.0718	67.3219
352.1220	226.3220	136.8010	-.5998	-208.9215	-69.0169
-676.1464	-592.5092	-.1991	-866.1292	-420.9913	21.2408
146.0329	110.2308	70.4107	143.3070	554.9741	1744.6739
1417.7453	601.3942	-626.4797	-912.1997	-327.0794	115.1596
249.3438	194.1396	154.6203	177.2139	-1.1271	-471.2352
-968.3682	-1484.7024	-1404.4601	-638.2663	-227.1425	169.0784
490.8546	308.0484	258.5299	311.1208	262.7715	-157.3444
-944.4716	-1590.7990	-1287.7406	-742.5368	-153.2116	347.9972

VALUES OF N= 4

AXN	BXN	CXN	THETA
46.00	-195.70	102.63	-.29

YR VALUES AFTER REMOVAL OF MONTHS PERIOD

-1741.4984	-1176.9364	-2207.1365	-1674.3597	-416.6400	516.6079
-271.0357	-307.0362	-269.7343	-471.1520	-910.9901	-1313.0470
2332.3936	1000.9002	1636.2012	1639.5767	977.2899	49.5999
-227.1260	-303.9205	-175.8720	-197.2527	-343.1013	-229.1493
556.2009	1075.0567	2230.1967	533.5129	-260.7031	-95.4801
-203.2180	-260.0040	-106.1300	-283.3527	-724.2047	994.7404
510.1026	2464.7532	1114.1160	47.4492	-264.0502	-01.5761
-199.3093	-146.0891	-23.0491	-249.4526	-430.3000	-711.3939
-1215.9230	-2141.3502	236.0035	-330.6144	-400.9202	-167.6642
-239.4004	-152.1735	-19.2374	-165.5526	-496.4113	-677.4562
-842.0301	-1057.4538	201.9509	-574.6781	-316.9912	-93.7522
-161.4916	-118.2570	44.7742	-51.6526	1647.4042	-713.5585
-1180.1364	-1093.5572	-784.1316	69.2502	96.9376	80.1996
-67.5028	-4.3421	120.6359	112.2474	-278.6130	3240.3392
1165.7572	2500.3395	3226.7099	2103.1946	590.0696	114.0716
196.3239	179.5735	292.5975	46.1474	-364.7214	-115.7631
-520.3492	-549.7642	-156.2966	-912.0691	-265.2004	67.9835
-9.7652	63.4092	226.5092	190.0474	399.1732	1690.1346
1573.9449	840.1323	-902.3792	-950.9320	-171.2735	161.8955
94.1435	147.4040	310.4209	223.9474	-156.9241	-517.9677
-012.5619	-1437.5712	-1360.4617	-704.9965	-71.3405	235.0074
335.0923	261.3205	414.3326	357.8475	106.9685	-204.0700
-708.6602	-1544.0746	-1443.1442	-709.0601	2.5893	394.7193

VALUES OF N= 3

AXN	BXN	CXN	THETA
134.94	7.05	135.12	1.51

YR VALUES AFTER REMOVAL OF MONTHS PERIOD

-1600.1404	-1103.3519	-2422.3795	-1612.9999	-343.0572	381.7446
-209.6741	-314.2350	-404.7277	-609.7894	-843.4186	-1447.9905
2613.7605	1160.5380	1501.0376	1700.9436	1050.8660	-89.3437
-165.7302	-230.3460	-310.5164	-135.8023	-269.5284	-364.0932
617.6611	1149.4270	2095.2547	594.0069	-195.2106	-230.4322
-221.8423	-166.4370	-241.9050	-221.9792	-650.6385	859.0040
571.5619	2530.5177	979.1716	108.6303	-191.2873	-216.3207
-157.9264	-72.5279	-157.5938	-108.0600	-356.7485	-846.2986
-1154.5375	-2067.7924	103.0806	-777.2262	-327.3610	-302.6092
-174.0105	-76.6189	-154.0325	-304.1609	-622.0585	-812.4013
-780.6367	-903.9026	147.0056	-913.2029	-243.4497	-180.6976
-100.0947	-44.7099	-90.3712	9.7461	1721.0314	-840.5041
-1126.7359	-1560.0127	-919.8773	130.6605	170.4825	-54.7061
-6.1788	69.1990	-6.2599	173.6932	-205.0785	3105.3933
1227.1647	1573.0771	3089.8599	2164.6039	624.4098	-20.8745
297.7370	253.1080	157.6512	107.5603	-291.1885	-250.7095
-458.9345	-472.2330	-291.2431	-851.4526	-191.6738	-66.9630
51.6529	137.0170	91.1625	251.4674	472.7014	1563.1070
1634.9662	921.6569	-1117.3261	-897.5093	-97.7476	26.9484
155.9680	226.9261	175.4730	239.3745	-83.4005	-652.9149
-731.1331	-1564.4534	-1495.4090	-643.5659	2.1756	100.0600
396.4347	934.0351	279.2051	419.2016	180.4813	-339.0176
-727.2324	-1470.5634	-1578.6919	-727.6225	76.0919	259.7715

N(d

PERIODOGRAM ANALYSIS

CALCULATION OF AXN , BXN, CXS AND FREQUENCY (F)

AX_N	BXN	CXS	F
-378.1640	-300.8739	233533.1700	0.0416
- 96.4439	- 70.1824	14226.9990	0.0625
336.5373	952.1683	1019881.9000	0.0833
-202.8917	-343.8619	159406.0900	0.1000
24.3798	-145.0569	21635.8840	0.11111
- 38.1687	-317.0874	102001.3200	0.1250
-261.0107	177.4185	99603.9440	0.1666
46.0787	-162.6556	28580.1200	0.2500
128.802	11.3237	16730.6000	0.3333
- 11.4640	0.0013	131.4242	0.5000

RESULTS OF POWER SPECTRUM ANALYSIS

Values of			
F	VD	F	VD
0.5000	3.04625	0.12500	0.413801
0.4500	1.821043	0.1000	0.277984
0.4000	0.277695	0.08300	3.382764
0.3500	0.278171	0.07500	-1.038877
0.33300	3.389455	0.07000	-0.285306
0.30000	-1.822218	0.06500	0.799424
0.25000	3.405599	0.06000	1.931568
0.20000	-1.821449	0.05000	-1.821816
0.16700	3.383297	0.04200	0.414021
0.15000	0.277881	0.03500	-0.75509

RESULTS

REMOVAL OF 12 MONTHS

MEAN OF THE DATA= -0.23310400E+02
STANDARD DEVIATION OF THE DATA= 962.7033
VARIANCE OF DATA= 0.92679766E+06
C.V. OF DATA = -0.41299290E+04

REMOVAL OF 6 MONTHS

MEAN OF THE DATA= -0.23310097E+02
STANDARD DEVIATION OF THE DATA= 940.5224
VARIANCE OF DATA= 0.88458232E+06
C.V. OF DATA = -0.40348282E+04

REMOVAL OF 4 MONTHS

MEAN OF THE DATA= -0.21842316E+02
STANDARD DEVIATION OF THE DATA= 933.4476
VARIANCE OF DATA= 0.87132436E+06
C.V. OF DATA = -0.42735741E+04

REMOVAL OF 3 MONTHS

MEAN OF THE DATA= -0.21842528E+02
STANDARD DEVIATION OF THE DATA= 928.5085
VARIANCE OF DATA= 0.86212801E+06
C.V. OF DATA = -0.42509204E+04

STOCHASTIC COMPONENT

MEAN OF THE DATA= -0.14446111E+02
STANDARD DEVIATION OF THE DATA= 565.1790
VARIANCE OF DATA= 0.31942730E+06
C.V. OF DATA = -0.39123263E+04

RONDOM COMPONENT

MEAN OF THE DATA= 0.47705115E+01
STANDARD DEVIATION OF THE DATA= 727.0126
VARIANCE OF DATA= 0.52971119E+06
C.V. OF DATA = 0.15230940E+05

RESULTS

C C LUHUMBIKA-HYDROLOGY-DISSERTATION WORK

PROGRAM ACCEPTED2 36930 50090 99699 59990

VALUES OF STOCHASTIC COMPONENT

-1329.0374	-671.6940	-1475.1006	-902.1956	-208.0079	232.4443
-127.6706	-191.3499	-246.4387	-249.5200	-513.5576	-551.6814
1469.7337	706.6546	914.1669	1035.7049	639.0723	-54.4016
-170.9302	-161.2577	-189.2961	-82.7987	-164.1190	-221.0969
376.0930	699.0066	1277.8005	362.2266	-118.0697	-140.3102
-135.3790	-113.5219	-147.2960	-135.1607	-396.1738	523.5367
340.3240	1549.5816	596.2176	66.2608	-116.4740	-131.0399
-83.9834	-46.1622	-96.2026	-114.5146	-217.2242	-515.3112
-732.9979	-1259.0747	62.7706	-168.0030	-199.3319	-104.2987
-139.9590	-47.8710	-93.0200	-63.4236	-257.4789	-494.6712
-479.3297	-599.0913	89.5117	-312.9800	-140.2310	-114.0900
-30.9477	-27.2219	-54.9052	9.9344	1047.9960	-916.6941
-636.7699	-949.0917	-959.6262	79.5592	103.8060	-33.3993
-3.7623	42.1303	-3.0117	105.7374	-124.0723	1090.0739
747.2206	1567.2387	1881.4035	1310.0273	380.2007	-12.7109
136.9361	154.1179	91.9938	65.4939	-177.3047	-152.6570
-279.4492	-287.3417	-177.3379	-510.4495	-116.7004	-40.7738
31.4519	89.4297	55.7324	153.1105	287.8279	951.0291
989.5909	561.1969	-600.9299	-546.4934	-59.5105	16.4009
94.7293	134.9219	106.0460	173.7649	-50.7074	-397.5599
-437.3049	-830.8117	-910.5545	-391.0671	1.3247	61.6137
211.4199	209.0011	170.1176	255.3006	109.8931	-206.4270
-442.0110	-895.4281	-961.1437	-443.0491	46.3366	

VALUES OF RANDOM COMPONENT

-30.3145	-1750.7436	-137.8913	639.0934	590.6321	-442.1104
-185.9845	-213.3778	-163.3907	-393.0970	-934.4230	3295.4419
-399.2037	794.6061	706.7792	15.1619	-729.2160	-111.3568
-129.4190	-170.9517	59.3730	-156.7097	-199.9774	839.3974
773.3360	1995.9612	-680.9136	-557.4372	-111.3669	-81.5321
-31.3572	-120.3045	-74.6792	-519.6778	1255.9777	40.0279
2190.2937	-566.4100	-487.9079	-257.5541	-100.0499	-6.0870
11.4555	-113.8326	-91.0696	-242.2339	-629.0744	-639.2269
-1364.7946	1362.1673	-339.9960	-190.5610	-103.2772	10.2402
27.3361	-106.2149	-10.3401	-359.4349	-554.9220	-205.9656
-590.5729	746.1039	-602.7946	69.0979	-40.4666	14.8033
16.2378	-62.9473	64.6513	1715.0970	-1096.4401	-610.0018
-573.9433	30.8144	699.2967	90.9233	-198.9929	27.1005
72.9613	-40.3932	177.4640	-310.8199	3230.2655	-663.7092
1326.6566	1522.6042	289.2004	-699.6215	-401.0792	270.4475
96.1720	3.9337	11.5665	-356.6820	-73.4048	-306.2779
-192.7078	-3.7004	-674.1147	326.7787	49.7494	92.4267
109.5696	0.1330	199.7150	319.5029	1279.3600	683.1412
-73.0749	-1670.5229	-217.1694	449.7490	86.4669	139.1599
126.2003	47.9519	171.5285	-257.1730	-602.1279	-353.5722
-937.0089	-664.9934	261.9386	394.0429	99.9353	335.0711
93.4156	79.9040	269.1640	-74.0193	-446.9127	-520.8046
-1327.7516	-633.0639	293.5212	519.1432	213.4349	

0 STOP END AT S. 0019 + 02 L. 2

CORRELOGRAM ANALYSIS

RESULTS

1	2	3	4	5	6	7	8	9	10
15	16	17	18	19	20	21	22	23	24
25	26	27	28	29	30	31	32	33	34
35	36	37	38	39	40	41	42	43	44
45	46	47	48	49	50	51	52	53	54
55	56	57	58	59	60				

Removal of 12 months

.5840	.3116	.0859	-.0107	-.0301	-.0020
-.0003	-.0190	-.0665	-.0821	.0206	-.0056
-.0450	-.1881	-.1546	-.0879	.0130	.0254
.0303	.0602	.0688	.0249	.0614	.0546
-.0186	-.2080	-.2120	-.1025	-.0246	.0160
-.0219	-.0762	-.1659	-.4053	-.4100	-.3224
-.2847	-.1626	-.0798	-.0122	.0305	.0385
.0177	.0114	-.0043	-.0014	-.0183	.1002
.0788	.0476	-.0094	-.0074	.0070	.0309
-.0079	-.0074	.0571	.0664	.0613	.1825

Removal of 6 months

.5924	.3589	.1474	.0158	-.0638	-.0656
-.0336	.0098	-.0061	-.0508	-.0034	-.0642
-.0845	-.1787	-.1102	-.0661	-.0180	-.0356
-.0006	.0926	.1348	.0605	.0386	-.0010
-.0526	-.1962	-.1720	-.0859	-.0550	-.0315
-.0378	-.0416	-.1197	-.5090	-.4655	-.3940
-.3197	-.1412	-.0393	-.0011	.0010	-.0032
.0038	.0395	.0366	.0125	-.0443	.0667
.0748	.0873	.0362	.0089	-.0121	.0049
-.0106	.0305	.1000	.0727	.0293	.1547

Removal of 4 months

.6000	.3762	.1465	-.0032	-.0669	-.0494
-.0324	-.0076	-.0082	-.0342	-.0021	-.0843
-.0885	-.1597	-.1094	-.0906	-.0217	-.0138
.0012	.0712	.1357	.0854	.0417	-.0248
-.0559	-.1763	-.1727	-.1117	-.0579	-.0060
-.0337	-.0681	-.1278	-.3939	-.4734	-.4303
-.3306	-.1166	-.0346	-.0285	-.0057	.0221
.0101	-.0142	.0310	.0393	-.0391	.0415
.0685	.1149	.0462	-.0096	-.0158	.0278
.0012	.0082	.0902	.0977	.0419	.1340

Removal of 3 months

.6089	.3844	.1367	.0016	-.0595	-.0606
-.0293	-.0008	-.0197	-.0315	.0042	-.0978
-.0819	-.1545	-.1233	-.0820	-.0138	-.0294
.0096	.0790	.1216	.0948	.0499	-.0416
-.0475	-.1704	-.1895	-.1040	-.0528	-.0245
-.0257	-.0622	-.1466	-.3915	-.4744	-.4510
-.3260	-.1089	-.0505	-.0208	.0020	.0073
.0182	.0221	.0164	.0469	-.0326	.0262
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

PERIODOGRAM ANALYSIS

RESULTS AFTER REMOVAL OF 12,6,4,3 MONTHS

AXN	BXN	CXS	F
-403.9401	-262.4641	232055.0400	.0416
-101.7903	-25.0191	10987.2320	.0625
-.0018	.0020	0.0000	.0833
-144.3958	-436.6507	211514.0400	.1000
56.4277	-120.8999	17800.8880	.1111
-3.1716	-303.7885	92297.5660	.1250
-237.2802	165.8798	83818.0130	.1666
46.0763	-162.6576	28580.5340	.2500
138.7806	8.2342	19327.8590	.3333
-2.8880	.0057	8.3408	.5000

AXN	BXN	CXS	F
-406.9384	-267.3737	237087.5200	.0416
-107.8205	-30.6220	12562.9850	.0625
-9.1293	-4.6957	105.3941	.0833
-144.8007	-443.7501	217881.4200	.1000
44.3811	-122.2533	16915.5690	.1111
-15.7881	-302.9127	92005.4130	.1250
-.0002	.0008	0.0000	.1666
46.8011	-155.7805	26457.9140	.2500
138.7806	8.2350	19327.8860	.3333
-2.8887	.0067	8.3450	.5000

AXN	BXN	CXS	F
-405.3399	-268.7900	236548.5000	.0416
AXN	BXN	CXS	F
-107.2638	-31.6844	12509.4300	.0625
AXN	BXN	CXS	F
-9.1289	-4.6956	105.3868	.0833
AXN	BXN	CXS	F
-142.2961	-445.0768	218341.5700	.1000
AXN	BXN	CXS	F
45.0413	-120.5415	16558.9860	.1111
AXN	BXN	CXS	F
-14.1916	-300.6384	90584.8930	.1250
AXN	BXN	CXS	F
5.1929	1.1755	28.3489	.1666
AXN	BXN	CXS	F
.0006	-.0002	0.0000	.2500
AXN	BXN	CXS	F
134.9431	7.0599	18259.4840	.3333
AXN	BXN	CXS	F
-4.4689	.0059	19.9717	.5000
AXN	BXN	CXS	F
-404.3650	-257.6693	235319.7400	.0416
AXN	BXN	CXS	F
-105.8333	-30.5659	12134.9620	.0625
AXN	BXN	CXS	F
-7.2377	-3.9796	68.2223	.0833
AXN	BXN	CXS	F
-142.0616	-443.8753	217206.8200	.1000
AXN	BXN	CXS	F
46.6860	-120.6134	16745.9080	.1111
AXN	BXN	CXS	F
-12.6774	-301.0071	90766.0100	.1250
AXN	BXN	CXS	F
5.1926	1.1751	28.3448	.1666
AXN	BXN	CXS	F
1.7792	3.9116	18.4667	.2500
AXN	BXN	CXS	F
.0003	-.0006	0.0000	.3333
AXN	BXN	CXS	F
-4.4678	.0048	19.9617	.5000