STRUCTURAL ANALYSIS OF TIME SERIES OF MONTHLY RAINFALL AND RUNOFF SEQUENCES

A Dissertation
submitted in partial fulfilment of
the requirements for the award of the degree
of
MASTER OF ENGINEERING
in
HYDROLOGY

By
H.S. VASUDEVA MURTHY

UNESCO SPONSORED
INTERNATIONAL HYDROLOGY COURSE
UNIVERSITY OF ROORKEE
ROORKEE (INDIA)
April, 1976

ACKNOWLEDGEMENT

The author wishes to express his deepest sense of gratitude to Dr. S.M. Seth, Reader in Hydrology, University of Roorkee, for the constant supervision, and excellent guidance given to him, at every stage of the present study.

The author is grateful to Dr. Satish Chandra, Professor and Co-ordinator, International Hydrology Course, School of Hydrology, University of Roorkee, for the encouragement rendered to him, throughout his stay at Roorkee.

The author is also grateful to Sri C.P. Bhat, Chief Engineer, P.W.D., Water Resources Development Organisation, Bangalore, and the Government of Karnataka for deputing him to the course.

The assistance of all persons concerned with computational facilities : IBM 1620 of S.E.R.C., Roorkee, and IBM 360 of I.I.T., New Delhi, is thankfully acknowledged.

H.S. Vasudeva Murthy
(H.S. Vasudeva Murthy)

Roorkee
April 9, 1976

C E R T I F I C A T E

This is to certify that the dissertation entitled 'Structural Analysis of Time Series of Monthly Rainfall and Runoff Sequences' being submitted by Sri N.S. Vasudeva Murthy in partial fulfilment of the requirements for award of the degree of Master of Engineering in Hydrology of the University of Roorkee, is a record of the candidate's own work carried out by him under my supervision and guidance. The material embodied in this dissertation has not been submitted for the award of any other degree or diploma.

This is further to certify that Sri N.S. Vasudeva Murthy has worked for a period of 6 months since 1st October 1975 to 31st March 1976 in the preparation of this dissertation under my guidance at this University.

(S.M. Seth)
Reader in Hydrology
International Hydrology Course
University of Roorkee
Roorkee.

ROORKEE

April 1, 1976.

C O N T E N T S

LIST OF TABLES

LIST OF FIGURES

S Y N O P S I S

Hydrologic data of sufficiently long duration is essential for better decisions in planning, design, and operation of water resources projects for optimum development. However generally hydrologic data, such as stream flow, is rarely available for a sufficiently long period. To help overcome this deficiency, the modelling techniques developed in the field of stochastic hydrology are employed to generate a number of synthetic sequences for a longer period using statistical information of observed data.

The objective of this study is to analyse and mathematically describe the time series structure of hydrological sequences of monthly rainfall and monthly runoff of 20 years data from 1959-78 of Lakshmanathirtha subbasin of Cauvery basin.

The time series of monthly rainfall and monthly runoff constitute non-stationary time series, since each monthly value of a calendar month has its own expected value, variance etc. Two different approaches are generally used in stochastic modelling of monthly sequences. In the present study, the approach of time series models wherein the non-stationary process is decomposed into deterministic component, and stationary stochastic process, and ii) the approach which considers non-stationarity in an implicit manner (Thomas-Fiering Model) have been used, to develop stochastic models. For the time series, the parametric method of logarithmic transformation and/no transformation, as well as, non - parametric method of standardization have been used. The correlogram, periodogram, and variance spectrum analysis have been used for identification of significant harmonics in deterministic periodic component, which

is represented by Fourier coefficients. The determination coefficient method has been used for the analysis of dependence model of the stochastic component, and the best fit distribution has been found for the independent component using chi-square and Kolmogorov - Smirnov tests of goodness of fit.

The stochastic models based on time series approach have been developed for monthly rainfall and runoff sequences.

This study indicates that when 12-month cycle, and its significant subharmonics are removed from monthly rainfall series, the remaining stochastic component can be considered as an independent stochastic process. However for monthly runoff sequence the stationary stochastic component consists of both dependent and independent stochastic components.

Two sequences of monthly runoff for 50 years and 100 years were generated for each of the models using the stochastic models developed on the basis of stationarity approach as well as non-stationarity approach (no transformation, square root transformation and Logarithmic transformation) of Thomas Fiering Model.

The results from Logarithmic transformation model of the stationarity approach were comparatively better than the other two models. For the non-stationarity approach both square root and Logarithmic transformation gave reasonably good results.

C H A P T E R - 1

INTRODUCTION

1.1. GENERAL:

In the planning, design, and operation of a water resources system an important aspect is the prediction of future hydrologic sequences. In practice, the historical data usually cover a short period of time compared with the economic life of the system. Even in the case of available longest records, the most extreme drought or flood event can be far different from the next most extreme event. There is often serious question as to whether the extreme event is representative of the period of record. Therefore historical data can provide one set of samples, which gives only one set of responses to aid, in the planning, design and, operation of the system.

In order that some estimate of the likelihood of more severe sequences be made, the stochastic analysis of available hydrologic time series is required, so that its components can be mathematically described and thereby the sequential synthetic hydrologic records can be generated.

The design of water resources projects are mostly based on the assumed recurrence of past hydrologic events. By generating a number of hydrologic sequences, say ten to twenty, the length of which is corresponding to the period of project amortization, it is possible to create much broader base for hydrologic design, and obtain a large set of responses and the systems variability to different inputs.

The use of structural analysis and mathematical description of the components of hydrologic time series, as a simulation technique facilitates, examination of the operation of proposed project, by assuming the repetition of recorded streamflows or the occurrence of equally likely simulated streamflows having the statistical characteristics of the recorded flows.

1.2. HYDROLOGIC SYSTEMS:

Hydrologic phenomena are truely stochastic, that is, the hydrologic phenomena changes with time in accordance with the laws of probability, as well as, with the sequential relationship between its occurrences. For example the occurrence of a flood is considered to follow, the law of probability and, also the relationship with the antecedent flood condition.

Most of the conventional methods for hydrologic designs are deterministic, that is, the behaviour of the hydrologic process is assumed independent of time variations. For example, a unit hydrograph derived for a given river basin, for flood control project design, is based on historical flood records. Once derived, the unit hydrograph is used for analysis of future design floods. This is automatically assumed unchanged with time (from past to the future) and, therefore is deterministic.

Some conventional methods employ the concept of probability to the extent that no sequential relationship is involved in the probability. For example, the flood record is analysed and fitted with certain probability distribution to determine the recurrence intervals of the flood or the flood frequencies. Such methods are probabilistic, but not in the true sense stochastic(9).

The stochastic method, employing the concept of probability as well as its sequential relationship, has not been well introduced in the practical design and planning of hydrologic projects because such methods have not been fully developed. Conventional methods, deterministic and probabilistic, which do not conform closely to the natural stochastic phenomenon, in a particular case may not lead to proper designing of major hydrologic projects, where the assumption of deterministic process may lead to erroneous results. Depending upon the nature of the problem, the deterministic, probabilistic and stochastic techniques are complementary tools of the hydrologist.

1.3. <u>SELECTION OF TIME SERIES SAMPLE:</u>

Analysis of continuously recorded hydrologic time series are currently performed, by transforming the continuous series into discrete time series with time interval Δt, so that the length of discrete series = $T/\Delta t$, where T is the period of observation. Daily, monthly and annual time series are widely used in hydrology. The time series measure Δt, should produce the most statistical information with due consideration to the data processing and computation time. The choice of the discrete series of monthly values is, therefore, a compromise. Since the monthly values are extensively used in engineering applications, an analysis of time series of monthly precipitation and monthly river flow is fully justified. (10)

Thus the significant variations in system response, do not go undetected, while needless details in the analysis are avoided by choosing monthly time series. Hence, the monthly time series of rainfall and run off have been chosen, for the present study, to structurally analyse and, mathematically describe the components of the time series.

1.4. PROBLEM AND DATA:

Rainfall and Runoff for a period of 30 years are available for the Lakshmanathirtha river subbasin, of Cauvery basin. The problem of present study is structural analysis and mathematical description of the components of this data record, employing approaches based on stationarity, as well as, nonstationarity of time series and, the performance of stochastic models of run off series in generating synthetic sequences.

1.5. OBJECTIVE OF THE STUDY:

Stationarity of the time series indicates that the generating mechanism is of the same nature at each point in time, but generally, the hydrological sequences such as monthly flows are non-stationary i.e. time-variant. The formulation of time series models by stationarity approach is recommended by Yevjevich, a pioneer in the field of Stochastic Hydrology.(59).

The expected value of mean, variance etc. of each calendar month come from a different population, and thus sequence of monthly values constitute a sample from 12 different populations, and represents a non-stationary time series. The stationarity approach is based on the assumption that, after removal of deterministic components of trend and periodicity in parameters, the remaining series constitute a stationary stochastic series. The stationary stochastic series may be further subdivided into dependent, and independent components. In the non-parametric method of separating periodic and stochastic components, the 12 monthly means and 12 monthly standard deviations are used for standardising the observed sequence to derive the stochastic component and the number of parameters involved in the process is 24. But parametric method of modelling the deterministic component by Fourier series approach involves lesser number of parameters, still giving the similar results as that obtained by the non- parametric method. The periodogram analysis is used for determining the Fourier coefficients. The correlogram analysis and variance spectrum analysis are used for the identification and removal of significant harmonics. The stationary stochastic component derived from the original series after removal of deterministic component, is analysed for dependence and suitable autoregressive model is fitted. The stochastic stationary independent component computed from the dependence model of the stochastic component, is fitted with the appropriate probability distribution function.

Thomas and Fiering (99) method implicitly allows for the non stationarity of the monthly flow data. The Thomas and Fiering method consists of the use of twelve linear regression equations comprising of the statistics of the recorded flows.

The stationarity approach (non parametric and parametric) and non-stationarity approach are both adopted in the present study for structural analysis and description of the components of time series and synthetic sequences are generated in order to make a comparative study of these approaches.

CHAPTER - 2

ANALYSIS OF HYDROLOGIC TIME SERIES BY STATIONARY APPROACH - I (TREND COMPONENT)

2.1. GENERAL

A time series is a sequential record in time of a given phenomenon. Sequences of mean daily flows, mean monthly flows, annual instantaneous peaks etc., are examples of hydrologic time series, each of which displays a different structure, the manner in which the observations vary in time.

Two broad approaches are in vogue for the structural analysis and mathematical description of its components. One is, that the monthly series is composed of a stationary process combined with deterministic components, and the other that the monthly series is completely a nonstationary process.

2.2. STATIONARY APPROACH

Stationarity of a time series means the non-variance, except by chance fluctuations of the statistical parameters, such as mean, standard deviation, covariance, and other higher order moments of a time series. The stationarity is the key assumption implied in the formulation of time series models. Here a transformation of original series is required to produce the desired stationarity. A typical hydrologic time series may be composed of

i) a trend or a long term movement with oscillations around the mean.

ii) a seasonal or cyclic component

iii) the effect of serial correlation

iv) a random i.e. an unsymmetrical irregular component.

These components combine with each other in a manner in which the causative forces responsible for the production of these components interact with each other. For the simplicity of mathematical description, it is often assumed that the causative forces operate independently, so that a time series can be represented by the sum of the component forces.

2.3. COMPONENTS OF MONTHLY HYDROLOGIC TIME SERIES

The components of time series may thus be an additive model, such as

$$X_t = T_t + P_t + \epsilon_t \text{ (or } S_t) \quad \ldots\ldots\ldots\ldots \quad (2.1)$$

where X_t = observed monthly river flow sequence

T_t = trend component

P_t = periodic component

S_t or ϵ_t = stochastic component

where t is representing the month. This representation and also the representation like $X_{p,\tau}$, $\epsilon_{p,\tau}$ (where p indicates year and τ month) are used synonimously in the present study. The trend and periodic components form deterministic component where as, the stochastic component is a non-deterministic part comprising of autoregressive component and, a random component (Z_t)

T_t is deterministic in the sense that its future values may be predicted more accurately, and P_t is deterministic on the assumption that the oscillations have fixed phases and amplitudes. ϵ_t is stochastic since one can make statistical statements about their future behaviour.

The removal of trend component of the time series, makes the time series homogeneous and the removal of both trend, and periodic component will make the process stationary.

2.4. PHYSICAL EXPLANATION OF THE COMPONENTS OF HYDROLOGIC TIME SERIES:

The components of hydrologic series are still far from being well understood, particularly the interaction between periodic and stochastic components. Although periodicities are explained by astronomic cycles, and consequently by the periodicity in the energy supply from the sun over various areas of the earth's surface, and further interactions and responses of various earth's environments, the complexity of periodic components needs a much better physical analysis than is presently available. Similarly, though the stochastic component may be explained by various random processes in air, over oceans and at the continental surfaces, i.e. over various geophysical environments, many more efforts are needed to improve its physical understanding, explanation, and description (57).

Trends and cyclicity are often the results of sampling fluctuations for short time series. The long range trends, and periodicity, due to sampling variation alone, should not be considered as a permanent property of any series of annual values. The hydrologic time series which is nonstationary, can be decomposed into deterministic, and stationary stochastic process. Many present day approaches in hydrology for the analysis of time series and generation of new samples are based on the above approach and also non-stationary approach to the treatment of stochastic aspects of these series.

2.9. ERRORS OF INCONSISTENCY AND NON-HOMOGENEITY

Inconsistency in data is defined as systematic errors in measurement, and computation, which make a difference between, the figures produced by observations and computations, and true values. Non-homogeneity in data is defined as irregularities in time series resulting from substantial changes in environment introduced either by man-made measures or by accidents in nature, which produce differences between virgin values (values produced in observations if the causative factors remain unchanged with time) and the true values (49).

2.9.1. Trends

A trend is defined as systematic, and continuous change over an entire sample, in any parameter of a series, excluding periodic changes, and produced by factors other than the expected sampling variation of stochastic process. Yevjevich (53) states that trends, as deterministic components, often occur in hydrologic series. Usually trend is assumed to be found in mean only.

If a stream flow series contains trend, and if future samples are not likely to experience the same or similar trend, direct use of recorded data for further analysis or generation of equally likely future sequences may be highly biased. Elimination of trend is, thus necessary to avoid this bias. If the same trend, as recorded, is expected in future sequences it can be incorporated back in the generated sequences. The synthesis of stream flow data, thus requires detection, and removal of trends from recorded sequences by reducing the recorded sequences to virgin data or homogeneous samples.

When a hydrologic series has no significant trend in the mean, and standard deviation, the entire series may be inferred as homogeneous.

2.9.2. Tests of homogeneity:

Tests of homogeneity can be carried out by split-sample approach in ascertaining whether differences between the means of the two unequal sub samples are or are not, significantly different from zero on the 95 pct. probability level of significance. Only if the probability is less than 5 pct. that a difference is greater than the critical value of these differences, the two subsample means are considered not to be from the same population, or the series considered to be non homogeneous (99).

The t - statistic is used for testing whether the difference of the two means, $\bar{X}_1$ and $\bar{X}_2$ is significant with

$$t = \frac{\bar{x}_1 - \bar{x}_2}{S\sqrt{\frac{n_1 + n_2}{n_1 n_2}}} \quad \ldots\ldots\ldots\ldots (2.2)$$

$$S = \sqrt{\frac{\sum_{i=1}^{n_1}(x_i - \bar{x}_1)^2 + \sum_{j=1}^{n_2}(x_j - \bar{x}_2)^2}{(n_1 + n_2 - 2)}} \quad \ldots (2.3)$$

S, in the denominator of Eqn. 2.2 is the pooled standard deviation of the two sub samples, n_1 and n_2 are sub sample sizes, X_i are values of the series in the n_1 sub sample and X_j in the n_2 sub sample. This t has the Student t - distribution. The critical value, t_c, for the significance probability level of 95 pct. is then taken from the Student t - distribution tables.

2.5.3. Tests for the significance of trends

The trend in hydrologic time series must be removed if it is not expected either to be repeated or if it will not occur at all in future. The removal of trend, if found significant, and makes the time series homogeneous.

Normally only linear trends are used, because any non-linear trend, though easy to fit, may have small justification, and because the difference between the non-linear and linear trends may be partly or fully the result of sampling variations. (59)

Trend can be removed from the hydrological time series either by the method of least squares or by the moving averages.

The method of least squares is used in this study and is briefly explained as follows:-

Assuming a linear fit, the best straight line through the points (X_i, Y_i), is by choosing the two parameters m and c of a straight line $y = mx + c$ . . . . (2.4)

in such a way that the sum of the squares of the errors in y is minimum. This is achieved by making

$$m = \frac{n\Sigma XY - \Sigma X\Sigma Y}{n\Sigma X^2 - (\Sigma X)^2} \quad \ldots \ldots \ldots \ldots \ldots \quad (2.5)$$

$$c = \frac{\Sigma X^2\Sigma Y - \Sigma X\Sigma XY}{n\Sigma X^2 - (\Sigma X)^2} \quad \ldots \ldots \ldots \ldots \ldots \quad (2.6)$$

where n is the number of paired observations, and X_i are prescribed values over monthly ranges, and y_i are the consequential observational values. It is assumed here that the observations y_i are subjected to error. This line (Equation 2.4) is called the line of regression of y on x, and one of its properties, is that it passes through the centroid $(\bar{x}, \bar{y})$ of the observed points.

2.5.4. Confidence limits of mean value and slope of regression line:

The usual statistical question now arises concerning the confidence limits which should be applied to the calculated line, which is an estimate of a relationship. To examine this problem the errors or deviations must be calculated. At every observation point (x_i , y_i) which does not actually lie on the calculated line there is an e_i . The variance of y estimated by the regression line is then (5).

$$s_y^2 = \frac{\Sigma e_i^2}{u} \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (2.7)$$

where is u i.e. the degree of freedom.

Since calculation of a and b impose two restraints, the value of u is given by u = n-2. The variance of the mean value $\bar{y}$ is given by

$$s_{\bar{y}}^2 = \frac{s_y^2}{n} \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (2.8)$$

So that the confidence limits of $\bar{y}$ are

$$\bar{y} \pm t\, s_{\bar{y}} \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (2.9)$$

where, similar to sample mean, the value of t is found from tables, using the appropriate number of degrees of freedom.

The variance of the slope b is given by

$$s_b^2 = \frac{s_y^2}{\Sigma(x_i - \bar{x})^2} \quad \ldots\ldots\ldots\ldots \quad (2.10)$$

and the confidence band for slope is given by

$$m \pm t\, s_m \quad \ldots\ldots\ldots\ldots\ldots \quad (2.10a)$$

It is sometimes necessary to compare the regression line with another theoretical one, to see if there is any significant difference between the theoretical slope, m_0, and the observed slope m. This test is performed by calculating a t - statistic;

$$t = \frac{m - m_0}{s_m} \ldots\ldots\ldots\ldots (2.11),$$ and comparing with the tabulated values.

Therefore a straight line trend $y = mx + c$, fitted as a regression line through y time series, is then tested for m being significantly different from zero. This approach of trend detection and description is more reliable than the moving average approach. The moving average technique creates distortions of periodic components in various parameters and also spurious oscillations produced in smoothing the stochastic component which is removed simultaneously with the removal of trend. Yevjevich (58) states that because of these distortions, moving average schemes as an investigating technique, is not suitable in the analysis of hydrological time series.

CHAPTER - 3

TIME SERIES ANALYSIS BY STOCHASTIC APPROACH II (PERIODIC AND STOCHASTIC COMPONENTS)

3.1. GENERAL:

The analysis of periodic component consists of two parts

i) identification of the harmonics or periods present

ii) removal of the same from residual series, obtained after the removal of trend component, by Fourier series approach.

The identification of significant periods is generally carried out with reference to confidence interval, which is defined as an interval around the computed parameter within which a given percentage of parameters of a large number of samples is expected to be found. This given percentage is the "level of confidence". The confidence interval at 95 pct. level means that, out of 100 samples of equal size, it is expected that 95 values of a parameter would be inside that interval. It is equivalent to 5 pct. level in which case expected percentage of parameters falling outside the confidence interval is considered. The confidence limits are numerical values describing the boundaries of the confidence interval (8).

A value of a variable x in the year p and at the position τ inside the year is $x_{p,\tau}$, with $p = 1,2 \ldots\ldots n$, and $\tau = 1,2 \ldots \omega$. This $x_{p,\tau}$ value is for the month τ, of year p following the beginning of records, with n the number of years of record, and ω the total number of discrete values in a year ($\omega = 12$, then τ represents months)

3.2. THE PERIODIC COMPONENTS IN MEAN STANDARD DEVIATION AND [illegible]:

The estimates of the means at any discrete position τ of the periodicity ω from n values of $X_{p,\tau}$ – series, with $\tau = 1, 2, \ldots, \omega$ and $p = 1, 2, \ldots, n$, are

$$m_{\tau} = \frac{1}{n} \sum_{p=1}^{n} x_{p,\tau} \quad \cdots\cdots \quad (3.1)$$

The estimates of the standard deviations are

$$s_{\tau} = \left[\frac{1}{n} \sum_{p=1}^{n} (x_{p,\tau} - m_{\tau})^2 \right]^{1/2} \cdots (3.2)$$

if n is sufficiently large, 'n' in the denominator substituted by (n-1) for smaller values of n (i.e. when n < 30)

The variance is given by the square of the population standard deviation of equation (3.2).

The estimates of the coefficient of skewness are given by

$$\gamma_{\tau} = \frac{\frac{1}{n} \sum_{p=1}^{n} (x_{p,\tau} - m_{\tau})^3}{\left[\frac{1}{n} \sum_{p=1}^{n} (x_{p,\tau} - m_{\tau})^2 \right]^{3/2}} \quad \cdots\cdots (3.2a)$$

3.3. NON-PARAMETRIC METHOD OF SEPARATING PERIODIC AND STOCHASTIC COMPONENTS:

The simple transformation

$$\epsilon_{p,\tau} = \frac{x_{p\tau} - m_{\tau}}{s_{\tau}} \quad \cdots\cdots\cdots\cdots (3.3)$$

in which m_{τ} and s_{τ} are the sample means and sample standard deviations at the position τ, computed by equations 3.1 and 3.2 respectively, is the non-parametric method of standardization of the $X_{p,\tau}$ variable. This is also a way to remove the periodic components in m_{τ} and s_{τ}. It requires the use of 2ω statistics, ω of m_{τ} and ω of s_{τ}. For monthly values $2\omega = 24$, and for daily values $2\omega = 730$. The non-parametric method removes from a series the periodicity in parameters but also removes all sampling variations associated with the coefficients of the periodic functions of parameters.

3.4. PARAMETRIC METHOD OF SEPARATING OF PERIODIC AND STOCHASTIC COMPONENTS:

To economize on the number of statistics needed for the mathematical description of a series, the periodic means m_τ and s_τ may be approximated for large ω by a relatively small number of harmonics of ω. For example, if the periodic components of daily means and daily standard deviations are well approximated each by six harmonics, and all other fluctuations in m_τ and s_τ are assumed to be sampling variations, then the Fourier series approximation of a periodic parameter requires only the mean plus 12 values of A_j and B_j Fourier coefficients for each parameter, with a total of 26 statistics. This is a significant savings in the number of statistics used, 26 instead of 730 for the case of daily flows.

The classical approach in estimating the significant harmonics in computed series is of the type.

$$x_{p,\tau} = \mu_\tau + \sigma_x \epsilon_{p,\tau} \qquad (3.4)$$

in which μ_τ is the periodicity in the mean and σ_x is the standard deviation assumed to be a constant. The periodic component is then given in the form

$$\mu_\tau = \mu_x + \sum_{j=1}^{m} (A_j \cos \lambda_j \tau + B_j \sin \lambda_j \tau) \qquad (3.5)$$

for m harmonics, and μ_x is the general mean of $x_{p,\tau}$.

The coefficients A_j and B_j, $j = 1,2, \ldots, m$, in equation (3.5) are estimated from $n\omega$ values of $x_{p,\tau}$ by

$$A_j = \frac{2}{n\omega} \sum_{p=1}^{n} \sum_{\tau=1}^{\omega} (x_{p,\tau} - \mu_x) \cos \frac{2\pi j \tau}{\omega} \quad \ldots (3.6)$$

and

$$B_j = \frac{2}{n\omega} \sum_{p=1}^{n} \sum_{\tau=1}^{\omega} (x_{p,\tau} - \mu_x) \sin \frac{2\pi j \tau}{\omega} \quad \ldots (3.7)$$

in which $\lambda_j = \frac{2\pi j}{w}$ with sin $\lambda_j = \frac{2\pi}{w}$, n is the number of significant harmonics in the range of variation of j with j = 1,2 ...... w/2, and τ is the time series sequence inside each period, τ = 1,2 ....w with N = nw being the size of sample series.

The square of amplitude C_j^2 of any harmonic is given by $C_j^2 = A_j^2 + B_j^2$ . . . .(3.7a) and the mean square of deviations from the mean, μ_x, as the variance of that harmonic designated by h_j, is

$$\text{Var } h_j = \frac{C_j^2}{2} = \frac{A_j^2 + B_j^2}{2} \quad \ldots\ldots (3.7b)$$

If the summation of explained variances is made for all the significant harmonics of the time series, the difference between variance of the time series of $x_{p,\tau}$, and the total explained variance by these harmonics, is the variance attributed to the Stochastic Component of the time series. The larger this explained variance, the more seasonal is the character of monthly precipitation, and so on. However, if the explained variance is only a very small fraction of the total variance, it may be assumed that the time series has small seasonal variation.

Because σ_x of Eq. (3.6) is rarely a constant in hydrology, if μ_τ is periodic, two cases arise

(i) $\mu_\tau(x)$ and $\sigma_\tau(x)$ as the population periodic components, estimated by m_τ and s_τ, are proportional, and

(ii) $\mu_\tau(x)$ and $\sigma_\tau(x)$ are not proportional.

In the first case $\sigma_\tau \approx \eta_0 \mu_\tau$ with η_0 the proportionality constant, so that

$$x_{p,\tau} = \mu_\tau + \sigma_\tau \epsilon_{p,\tau} = \mu_\tau (1 + \eta_0 \epsilon_{p,\tau}) = \mu_\tau \epsilon^*_{p,\tau} \quad \cdots (3.8)$$

This is the case of a multiplication of a periodic parameter and a stochastic component, with $\epsilon^*_{p,\tau} = 1 + \eta_0 \epsilon_{p,\tau}$ which is a linear transformation of $\epsilon_{p,\tau}$. Equations (3.5, 3.6, 3.7) are not applicable in this case. However by using

$$\log x_{p,\tau} = \log \mu_\tau + \log \epsilon^*_{p,\tau} \quad \cdots (3.9)$$

for $x_{p,\tau} > 0$, $\mu_\tau > 0$ and $\epsilon^*_{p,\tau} > 0$ the case of Equation (3.8) is reduced, to the case of applying Equations (3.5) to (3.7) to logarithms of equation (3.9). The case of applying Equations (3.8) and (3.9) explains why studying logarithms of a hydrologic $x_{p,\tau}$ variable may give, in some cases, more meaningful results than studying the $x_{p,\tau}$ values (37).

The logarithms of $X_{p,\tau}$ further serves a two fold purpose. First, the largest values could be compressed or reduced, and second, on a log scale, the variation of the low flows would be magnified with respect to the high flow variations. These features make possible the description of the periodic components of logarithms of the monthly flows with fewer harmonics than is possible by use of the monthly flows themselves. (50)

The transformation $\log X_{p,\tau}$ and $\log \epsilon_{p,\tau}$ make their distributions less skewed, or they are close to symmetrical distributions. (57)

If σ_τ is not proportional to μ_τ, the simple composition model of the periodic and stochastic components in $X_{p,\tau}$ of

$$x_{p,\tau} = \mu_\tau + \sigma_\tau \epsilon_{p,\tau} \quad \text{------} \quad (3.10)$$

in which case equations (3.5) to (3.7), are not directly applicable, because μ_τ and σ_τ may have different significant harmonics, and/or different phases in the case of the same significant harmonics.

To avoid these difficulties in the application of the classical approach of Equation (3.4), or its application in using logarithms in the form of Equation (3.9), when the model of Equation (3.10) is required, various parameters that may be periodic along the sequence points $\tau = 1, 2, \ldots, \omega$ should first be computed and the significant harmonics fitted to them.

The periodic component in any parameter V may be approximated by a harmonics of its basic period ω in the form

$$V_\tau = V_x + \sum_{j=1}^{m} (A_j \cos \lambda_j \tau + B_j \sin \lambda_j \tau) \quad \cdots\cdots \quad (3.11)$$

in which $\lambda_j = \frac{2\pi j}{\omega}$ is the angular (circular) frequency, ω is the basic period in V, m is the number of harmonics inferred as significant in the Fourier series mathematical description of the periodic parameter V_p and V_n is the mean of v_τ

The standardization by Equation (3.9), but using the mathematical models of μ_τ and σ_τ, with a limited number of harmonics of Equation (3.11) as the fitted periodic components to m_τ and s_τ is defined as the parametric method of standardization.

$$y_{p,\tau} = \frac{x_{p,\tau} - \mu_\tau}{\sigma_\tau} \quad \text{- - - - - - - -} \quad (3.12)$$

Because of the difficulties in estimating the coefficients A_j and B_j of equation 3.11 directly from the $x_{p,\tau}$ series, they can be estimated from the ω values of v_τ by

$$A_j = \frac{2}{\omega} \sum_{\tau=1}^{\omega} v_\tau \cos \frac{2\pi j \tau}{\omega} \quad \ldots\ldots\ldots\ldots \quad (3.13)$$

$$B_j = \frac{2}{\omega} \sum_{\tau=1}^{\omega} v_\tau \sin \frac{2\pi j \tau}{\omega} \quad \ldots\ldots\ldots\ldots \quad (3.14)$$

The maximum number of harmonics in this discrete series of ω values of v_τ for monthly series is $m = 6$, and for daily series is $m = 182$. However the daily series rarely show significant harmonics beyond the first 6 to 12 harmonics.

The physical considerations of the hydrologic periodicities indicate that there is definitely one cycle per year, that of 12 months. Very often, another cycle, that of 6 months, is also clearly detectable from the observed data. In order to fit the trigonometric functions of the Fourier series to the shape of these two basic periodic movements (12 months and 6 months), subharmonics are necessary, and usually those of 4, 3, 2.4 and 2 months. The number of subharmonics of the main 12 - month cycle depends on the shape of the periodic movement. If the 12 - month periodic movement of μ_τ and σ_τ can be approximated well by a sine or cosine function, the 12 - month cycle without any of its subharmonics is sufficient. If the 12 - month periodic movement is far from a sine function, say with sharp peaks etc. of μ_τ and σ_τ, not only is the 6 - month harmonic necessary, but all other harmonics may be needed (59).

For the parametric method, Equation (3.12) is only approximately a standardized variable, because $E(Y_{p,\tau})$ and Var $Y_{p,\tau}$ are somewhat different from the expected value of mean zero, and the variance of unity respectively. To obtain a standardized variable in case the parametric method is used, a further transformation produces.

$$\epsilon_{p,\tau} = \frac{y_{p,\tau} - \mu_Y}{\sigma_Y} \quad \text{- - - - .(3·15)}$$

in which μ_Y is the mean of $Y_{p,\tau}$ (estimated by $\bar{y}_{p,\tau}$) and σ_Y is its standard deviation (estimated by s_y). The autocorrelation coefficients, and the skewness for each month are not affected by the transformation of Equation (3.15)

3.5. TESTING AND IDENTIFICATION OF HARMONICS OF PERIODIC PARAMETERS

The following approaches for determining the significant harmonics in the periodicity of parameters are considered in the present study.

i) Classical Fisher's approach of a process composed of the sum of a harmonic and a normal independent process.

ii) approximate approaches by using the first n harmonics until p percent of the variation of v values of v_τ about the mean V_n of the parameter V_τ is explained by these first n harmonics.

iii) Correlogram analysis.

iv) periodogram analysis

v) variance spectrum analysis

9.9.1. Fisher's approach to testing the significance of harmonics:

The parameter that can be used in testing the significance of various harmonics of Equations (9.4) and (9.9), is the variance of individual harmonics, $C_j^2/2$, provided the Fourier coefficients A_j and B_j are estimated by Eqs (9.6) and (9.7). If a test shows that a given $C_j^2/2$ value is not greater than a critical $C_c^2/2$ value of an independent stochastic process, this j-th harmonic is considered significant (92).

The Fisher's test has facilitated the detection of significant harmonics when Eq (9.4) is applicable (93) which is rarely a case in hydrology if μ_τ is periodic. Fisher's procedure is made biased due to several factors, particularly the time dependence in the stochastic component $\epsilon_{p,\tau}$ of Eq. (9.4) and the periodicity in μ_τ.

Equation (9.4) is not current in hydrologic practice, though it is often assumed and treated as such (97). However Equation (9.9) is a much more current case in hydrology, because the assumption of μ_τ and σ_τ being proportional may be close to physical reality. Fisher's approach is then applicable provided that $\log \epsilon^*_{p,\tau}$ is an independent variable. This is often satisfied for monthly precipitation series, but rarely fulfilled for the monthly run off series for which $\epsilon^*_{p\tau}$ is a time dependent random variable.

9.9.2. An approximate empirical approach for testing the significance of harmonics

The maximum number of potential significant harmonics in a series of monthly values is six. It is assumed in this empirical procedure that only the first six harmonics of a periodic parameter for time series of any interval $\Delta t \leq 30$ days should be tested for significance.

The ratio $\Delta P_j = \frac{\text{Var } h_j}{S^2(V)}$ . . . . (9.16) represents the part of the Variance V_τ which is explained by the j-th harmonic. The sum $\sum \Delta P_j$, $j = 1, 2, \ldots, 6$, gives P, the part of variation of V which is explained by the first six harmonics.

Two critical P-values, P_{min} and $P_{max} = (1 - P_{min})$ have to be selected. If $P \leq P_{min}$ no significant harmonic exists in the sequence of V_τ values, or $V_\tau = V_x$ is a non periodic parameter. If $P_{min} < P \leq P_{max}$, all six harmonics are inferred to be significant. However, if $P > P_{max}$, only some of the six harmonics are considered significant. The values of Var h_j are then sorted by magnitude from the highest to the lowest. Only those harmonics with the highest Var h_j are selected, which when summed up first exceed P_{max}. (As an example, if the three harmonics with highest Var h_j have $\sum \Delta P_j < P_{max}$, but the four harmonics with highest Var h_j have $\sum \Delta P_j > P_{max}$, these four harmonics are inferred to be significant).

The empirical expressions of P_{min} and P_{max} are

$$P_{min} = a\sqrt{\frac{N}{cn}} \quad \ldots\ldots\ldots\ldots\ldots\ldots (9.17)$$

and consequently $P_{max} = (1 - P_{min})$ . . . . (9.18)

c = highest moment used in the definition of V

Yevjevich (33) has suggested empirical constant to use in a = 0.033. The practical ranges of using Equations (3.17) and (3.18) are 12 ≤ v ≤ 335, Or, Δt = 1 day to Δt = 30 days, and 10 ≤ n ≤ 100.

3.5.3. Auto correlation Analysis :

The tendency for high flows to follow high flows and for low flows to follow low flows, is referred to as hydrologic persistence, as is attributed to storage processes in the atmosphere or in the drainage basin either surface or subsurface. The detection and quantitative measurement of persistence is important in the design of long term storage reservoirs. This persistence can be described by the structure of serial dependence of a stream flow sequence.

Autocorrelation analysis is used to determine the linear dependence among the successive values of a series that are a given lag apart. In case of two series, the lag cross correlation, with the positive or negative lag, gives the linear dependence of the successive values of the two series that are a given lag apart. A measure of this dependence is given by the serial - correlation coefficient. If the values of X_t are linearly dependent upon the values of X_{t+K}, then the correlation between X_t and X_{t+K} may be taken as a measure of dependence. This correlation is referred to as the K^{th} order serial correlation and is given by the open series approach as

$$r_k = \frac{(N-k)\sum_{t=1}^{N-k} x_t x_{t+k} - \left(\sum_{t=1}^{N-k} x_t\right)\left(\sum_{t=1}^{N-k} x_{t+k}\right)}{\left[(N-k)\sum_{t=1}^{N-k} x_t^2 - \left(\sum_{t=1}^{N-k} x_t\right)^2\right]^{1/2}\left[(N-k)\sum_{t=1}^{N-k} x_{t+k}^2 - \left(\sum_{t=1}^{N-k} x_{t+k}\right)^2\right]^{1/2}} \quad \cdots\cdots (3.19)$$

where N = Sample size, k = lag intervals 1.2.3.4 . . . . .

The value of r(K) for K=0 gives no information about the time series, as r(0) is always one, and as the observed time series may be considered as two identical sequences, each being the observed identical sequences, each being the observed time series itself.

The value of r(K) for K ≠ 0 reflect the structure of time series, and they are dimensionless, and oscillate between + 1 and − 1.

To facilitate the analysis of the structure, the values of r(K) as a function of K are depicted graphically, with r(K) as the ordinate and K as the abscissa, which is called correlogram. In order to reveal the features of the correlogram better, the plotted points are joined each to the next by a straight line (93).

In fact, whether the cycle of the correlogram is dampened or not, may be used as the criterion for identifying the periodicity present in the time series. Initially, the high frequency harmonic components, necessary to approximate well the shape of periodic movement, may not be readily discernible. As successive larger periods are removed from the time series, the smaller periods become clearly visible on correlograms.

3.5.3.1. Tests of significance:

A test of significance for serial correlation coefficient r(K), has been proposed by Anderson, is based on circular series approach, and the confidence limits which is modified by Yevjevich (93) for open series approach of equation (3.19) is

$$r_K(\alpha) = \left(\frac{-1 \pm n\alpha \sqrt{N-K-1}}{N-K} \right) \quad \ldots\ldots (3-20)$$

where N is the number of observed values in the time series, K is the lag, and n_α is the normal standard deviate from the standard normal distribution for a two tail test at the significance level α .

Common values of α and the corresponding values of n_α are (10).

α = 80 pct., n_α = 1.28
α = 90 pct. n_α = 1.64
α = 95 pct. n_α = 1.96

Thus the tolerance limits of r_K diverge as K increases.

For the sum of - harmonics, the correlogram itself is a harmonic with periods equal to those of harmonic components of time series and it will therefore show the same oscillations.

3.3.4. Periodogram Analysis (Line spectrum analysis) :

Schuster's periodogram is a plot of the squared amplitude as ordinate versus the frequency as abscissa. The amplitudes of various harmonics are given by equation (3.7c). periodogram is a discrete function of variances of harmonics against the frequencies of those harmonics. The value is known as the variance of the harmonic with coefficients A_j and B_j.

Hart-Lee (9) demonstrated that the periodogram is not an efficient estimator of the continuous spectrum.

3.3.5. Power Spectrum (Variance Spectrum) analysis:

Another diagnostic tool for the analysis of time series in the frequency domain, which can help develop an appropriate time series for the hydrologic process, is the power spectrum analysis.

The variance is a measure of dispersion of observations about a mean value, in a sense, it gives the average intensity of fluctuation of the phenomenon about a statistically steady mean value. In a single trace, the total variance is decomposed to find the frequency bands (periods) in which the variance contribution is statistically significant. The resulting variance (or power) spectrum is a sample spectrum that can be interpreted to better understand the physics of generating process (10).

For the discrete time series used in this study, the population variance density spectrum $\gamma(f)$, is defined as

$$\gamma(f) = 2\left[1+2\sum_{k=1}^{\infty} \rho(k) \cos 2\pi fk\right] \quad . \; . \; . \; . \; . \; . \quad (3.21)$$

in which f is the frequency $f = k/2m\Delta t$, $\rho(k)$ is the auto correlation function, and k the lag. The range of f in practice is 0 to 0.5. The lower limit of frequency is zero because $f_{min} = 1/T$ and for $T \to \infty$, $f_{min} = 0$. The upper limit of frequency is $f_{max} = 1/(2\Delta t)$; hence if Δt is expressed as a unit time period of one, then $f_{max} = 0.5$ Peaks in spectrum correspond to frequencies which account for a large percentage of the total variance. Inverses of these frequencies give the significant cycles, in Fourier series approach of time series description, that contribute heavily to the total variance.

In practice the sample variance density spectrum is obtained from an equation containing a smoothing function $D(k)$

applied to equation (3.21). The equation of the "smoothed" variance density spectrum is then

$$\hat{g}(f) = 2\left[1+2\sum_{k=1}^{m} D(k)\,(\cos 2\pi fk)\, r(k)\right] \ldots \quad (3.22)$$

in which m is the maximum lag used for r(k). Smoothing function is often referred to as weighting, filter, kernel function or spectral window. The smoothed power spectrum may be thought of as weighted average of the true spectrum taken over some frequencies.

The selection of m determines (m+1) discrete estimates at equidistant points of the spectrum. The larger m is, the larger the variance of the estimate at each point. Usually $m < N/3$ is selected and often $m = \frac{N}{10}, \frac{N}{6}, \frac{N}{5}$ or similar numbers (50). (where N = sample size). The desired resolution of variance spectrum and the required accuracy of the spectral density should also be considered for choosing m. (50)

The procedure followed for spectral analysis for obtaining significant periods necessary to define μ_τ is to be repeated for obtaining the significant periods to define σ_τ, except that in the latter case, the spectrum analysis is performed on the series Q_t, where $Q_t = (X_t - m_\tau)$, and m_τ is the monthly mean value for month τ. The periods found in Q_t are then used to define σ_τ. (50)

Many mathematical forms have been proposed for the spectral window. The design or selection of window shapes is termed window carpentry. The Hanning, Bartlett, Parzen and Tukey are common forms used in window carpentry (50). All the four forms were examined in this study.

The function proposed by Hanning is

$$D(k) = \left(1 + \cos \frac{\pi k}{m}\right)/2 \quad \ldots\ldots\ldots\ldots \quad 3.23)$$

Bartlett proposed the smoothing function in the lag domain as

$$K(k) = 1 - \left(\frac{k}{m}\right) \quad \ldots\ldots\ldots\ldots \quad (9.24)$$

The Parzen window in the spectral domain has the form

$$K(k) = 1-6\left(\frac{k}{m}\right)^2 + 6\left(\frac{k}{m}\right)^3, \text{ for } 0 \leq k \leq m/2 \text{ and}$$

$$= 2\left(1-\frac{k}{m}\right)^3, \text{ for } \frac{m}{2} \leq k \leq m \quad \ldots\ldots \quad (9.25)$$

Tukey (50) also suggested other models such as

$$K(k) = 1-2a + 2a\cos(\pi k/m) \text{ with } a = 0.23 \quad \ldots \quad (9.26)$$

The maximum lag controls the band width of the spectral window. The smoothed spectral estimates with a wide band width is to be computed first and then smaller band-widths to explore details of the spectrum is to be used. This technique is called window closing (6).

9.5.3.1. Significance tests

Jenkin's as quoted in Yevjevich (53) has shown that the distribution of g(f) may be approximately by a chi-square distribution with the equivalent degrees of freedom (EDF) depending on the spectral window used.

For the parzen window of Equation (9.25) the EDF is

$$EDF = \left(\frac{[illegible]}{m}\right) \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (9.27)$$

For the filter of Equation (9.24) proposed by Bartlett

$$EDF = \left(\frac{[illegible]}{m}\right) \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (9.28)$$

For the filter of Eq^n (9.26) and (9.29) proposed by Tukey and Hanning, $EDF = \left(\frac{[illegible]}{m}\right)$ .... (9.29)

The upper and lower tolerance limits for testing significance are given by

$$T_1 = \frac{1-\chi^2_{\alpha}(NDF)}{NDF} \qquad T_2 = \frac{\chi^2_{\alpha}(NDF)}{NDF} \quad \ldots\ldots(3.39)$$

Where T_1 is the upper tolerance limit and T_2 is the lower tolerance limit and $\alpha = 0.05$ for 95 pct. tolerance limits (χ^2 = chi-square). For values of T_1 and T_2, the corresponding values of spectrum are $T_1\, r(f)$ and $T_2\, r(f)$, with $r(f) = E[\hat{g}(f)]$. In case the limits are put to estimates $\hat{g}(f)$, the confidence limits are $T_1\, \hat{g}(f)$ and $T_2\, \hat{g}(f)$ with T_1 meaning the exceedance by 95 pct. of all values, and T_2 the nonexceedance by 5 pct. of all values (53).

For practical application, Yevjevich (6) suggests that the values of T_1 and T_2 should be increased at the two extremes by 41 percent because variances of spectral estimates at $f = 0$ and $f = 0.5$ are twice as large as they are at other frequencies.

3.5.6. Limitations of correlation and power spectrum methods in hydrology:

The use of correlation or power spectrum techniques depends upon the analysis of time series in time domain or frequency domain. Most prefer power spectrum methods because it is felt that structural components of the time series are more easily discerned in the frequency domain. In contrast to correlogram sampling, hypothesis testing is considerably simpler in power spectrum analysis. The test of hypothesis for serial correlation coefficients is originally based on a circular series approach of autocorrelation function, whose physical interpretation is obscure.

The geophysical processes of the hydrologic cycle are known to be non linear and to contain nonstationary components. Many problems can profitably be treated as linear and stationary and are amenable to analysis by spectral techniques (17).

Thus both the correlogram and the spectrum can be used to explore the internal structure of a time series. The correlograms reveal more about the relationship between values of the series which are separated in time. The spectrum provides the extent to which the series exists in some fundamental periodicities, and the peak of the spectrum is identified with the harmonic term in the system (24), and identifications of dependence models.

In fact spectral analysis essentially involves all the calculation in obtaining the correlogram plus the multiplication of each covariance function by a smoothing function, and then the fourier transformation of the products. Thus, the spectral analysis requires much more computational effort than the auto correlation analysis. The main advantage being able to detect noncommensurable harmonics which is not of much utility so far the monthly hydrological sequences are considered, because of the predominance of cycles induced by astronomical phenomena, whose presence in the sequence can be determined by auto correlation analysis also. However in the case of daily, weekly or ten-daily hydrological sequences where the significant harmonics have to be chosen from a large number of harmonics the spectral analysis plays a key role.

3.6. [illegible]

3.6.1. Investigation of Dependence Models :

The variable $\epsilon_{p,\tau}$ obtained in the previous section, by removing the periodicity in the mean and standard deviation, is only approximately a second order stationary dependent or independent time series (due to the presence of dependent stochastic component). The dependence can be, often approximated by the first-, second-, third-, or higher-order autoregressive linear models. The autoregressive models provide satisfactory representation of the data and have physical significance.

The general m - th order autoregressive(A.R.) linear model

$$\epsilon_{p,\tau} = \sum_{j=1}^{m} \alpha_{j,\tau} \, \epsilon_{p,\tau-j} + \sigma Z_{p,\tau} \quad \ldots\ldots (3.31)$$

with $\alpha_{j,\tau}$ the A.R. coefficients, either periodic as $\alpha_{j,\tau}$ or non periodic as constants α_j, which enables $Z_{p,\tau}$ to be a second order stationary and standard (0,1) random independent variable if

is a standard random but dependent variable.

The value of σ for nonperiodic autocorrelation coeff. is

$$\sigma = \left[1 - \sum_{k=1}^{m} \sum_{j=1}^{m} \alpha_{k} \, \alpha_{j} \, \rho_{|k-j|} \right]^{1/2} \quad \ldots\ldots (3.32)$$

For most hydrologic time series the above model using either the first -, second-, third-order will result in a good fit.(6)

3.6.2. Parameter Estimations:

Estimates a_j of the autoregressive coefficients α_j as a function of the estimated autocorrelation coefficients r_k are of the correlogram of $\epsilon_{p,\tau}$ are summarized below for the models of the first three orders.

For the first order model when n = 1

$$a_1 = r_1 \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (3.33)$$

The new series of nw - values of the standardised $Z_{p,\tau}$ series is computed from Equation (3.31) and (3.32) for n = 1 by

$$Z_{p,\tau} = \frac{\epsilon_{p,\tau} - a_1 \epsilon_{p,\tau-1}}{\sqrt{(1-a^2)}} \quad \ldots\ldots \quad (3.34)$$

For the second order model where n = 2

$$a_1 = (r_1 - r_1 r_2) / (1 - r_1^2) \quad \ldots\ldots\ldots \quad (3.35)$$

$$a_2 = (r_2 - r_1^2) / (1 - r_1^2) \quad \ldots\ldots\ldots \quad (3.36)$$

The new series of nw - values of $Z_{p,\tau}$ is computed from Equation(3.31) and(3.32) for n = 2 by

$$Z_{p,\tau} = \frac{\epsilon_{p,\tau} - a_1 \epsilon_{p,\tau-1} - a_2 \epsilon_{p,\tau-2}}{\sqrt{1 - (a_1^2 + a_2^2 + 2a_1 a_2 r_1)}} \quad \ldots\ldots\ldots \quad (3.37)$$

For the third - order model where n = 3

$$a_1 = \frac{(1-r_1^2)(r_1-r_3) - (1-r_2)(r_1 r_2 - r_3)}{(1-r_2)(1 - 2r_1^2 + r_2)} \quad \ldots \quad (3.38)$$

$$a_2 = \frac{(1-r_2)(r_2 + r_2^2 - r_1^2 - r_1 r_3)}{(1-r_2)(1-2r_1^2 + r_2)} \quad \ldots \quad (3.39)$$

$$a_3 = \frac{(r_1 - r_3)(r_1^2 - r_2) - (1-r_2)(r_1 r_2 - r_3)}{(1 - r_2)(1-2r_1^2 + r_2)} \quad \ldots \quad (3.40)$$

The new series of nw - values of $Z_{p,\tau}$ is computed from Equation(3.31) and(3.32) for n = 3 by

$$Z_{p,\tau} = \frac{\epsilon_{p,\tau} - a_1 \epsilon_{p,\tau-1} - a_2 \epsilon_{p,\tau-2} - a_3 \epsilon_{p,\tau-3}}{\sqrt{1 - (a_1^2 + a_2^2 + a_3^2 + 2a_1 a_2 r_1 + 2a_1 a_3 r_2 + 2a_2 a_3 r_1)}} \quad \ldots \quad (3.41)$$

in which r_k are the autocorrelation coefficients for lag k of ϵ_T series, and a_j are the autoregressive coefficients.

3.6.3. Selection of the order of the Autoregressive model:

Several methods can be used, to determine the appropriate orders of the a.r. models, for describing the time series under consideration. This involve the use of variance density spectra and correlograms.

The first approach uses the variance density spectrum of the $\epsilon_{p,\tau}$ series. A visual examination of the spectrum shows whether its shape is typical of first - order or second order auto regressive process.

The second approach is by analysing $\epsilon_{p,\tau}$ series by fitting first, second or third order a.r. models separately and testing the adequacy of the fitted scheme through a comparison of the residual variance i.e., of the $\epsilon_{p,\tau}$ series for the first, second and third order scheme.

The third approach is a simplified method of coefficient of determination R_j^2, $j = 1,2,3$ ...., proposed by Yevjevich (50). The coefficient of determination indicates what portion of the total variation of ϵ_t is explained, by the autoregressive part of the model, the remaining portion of the variance of ϵ_t being explained by the residual term σz_t in which σ is the standard deviation of Eq. (3.33), and z_t is the independent random variable. The coefficient of determination of the first three order a.r. models are given by

$$R_1^2 = r_1^2 \quad \dots\dots\dots\dots\dots\dots\dots \quad (3.42)$$

$$R_2^2 = \frac{r_1^2 + r_2^2 - 2r_1^2 r_2}{1 - r_1^2} \quad \dots\dots\dots\dots \quad (3.43)$$

$$R_3^2 = \frac{r_1^2 + r_2^2 + r_3^2 + 2r_1^3 r_3 + 2r_1^2 r_2^2 + 2r_1 r_2^2 r_3 - 2r_1^2 r_2}{1 - 2r_1^2 - r_2^2 + 2r_1^2 r_2} -$$

$$- \frac{4r_1 r_2 r_3 + r_1^4 + r_2^4 + r_1^2 r_3^2}{1 - 2r_1^2 - r_2^2 + 2r_1^2 r_2} \quad \dots\dots\dots \quad (3.44)$$

The first - order model is selected if

$$\left(R_2^2 - R_1^2\right) \leq 0.01 \quad \text{and} \left(R_3^2 - R_2^2\right) \leq 0.02 \quad \dots\dots \quad (3.45)$$

The second-order model is selected if

$$\left(R_2^2 - R_1^2\right) > 0.01 \text{ and} \left(R_3^2 - R_2^2\right) \leq 0.01 \quad \dots\dots \quad (3.46)$$

The third-order model is selected if

$$\left(R_2^2 - R_1^2\right) > 0.01 \text{ and} \left(R_3^2 - R_2^2\right) > 0.01 \quad \dots\dots \quad (3.47)$$

It is quite likely that the hydrologic physical reality imposes non linear auto - regressive models.

3.7. INDEPENDENT STOCHASTIC COMPONENT:

3.7.1. General:

The independent stochastic components $Z_{p,\tau}$ are obtained, by removing the dependence structures from the standardized residual components, using the autoregressive models determined in the previous section. The $z_{p,\tau}$ of the process, thus computed, is a standard (0,1) random variable.

3.7.2. Tests of Independence:

Tests of whether, the $z_{p,\tau}$ series, are random independent time process, can be performed using, either the correlogram or the variance density spectrum. In the correlogram analysis the test is to show that the r_k values for all k's are not significantly different from zero. In the spectral analysis, the test is to show the variance density spectrum is not significantly different from the average density. If the tests demonstrate the independence of the random variables, the second - order stationarity of the process investigated is accepted.

3.7.3. Assumptions:

The following assumptions are generally made to estimate the probabilities of $z_{p,\tau}$ series

1. The observed flows are independently distributed in time.
2. The flows follow a specified distribution function.
3. The estimates of the parameter values of the distribution are unbiased.
4. The sample size is large enough to warrant estimation of the parameter values.
5. No operational decisions, for example how to treat zero flows or negative values of $z_{p,\tau}$ markedly influence the results.

Although the validity of each of these assumptions is questionable, seldom have probability estimates been subjected to sensitivity analysis to assess the effect of the assumptions on the results.

3.7.4. Probability Distributions of Independent Stochastic Component:

3.7.4.1. Fitting probability functions to empirical frequency distributions of independent stochastic components:

Once $z_{p,\tau}$ is accepted as a stationary independent variable, the new values serve to determine a probability function of the best fit to the empirical frequency distribution. The fitting of a probability function to the frequency distribution curve of $z_{p,\tau}$ is the approach followed in this study to structurally analyze, and mathematically describe a hydrologic time series.

The transformation of $x_{p,\tau}$ to produce the standardized variable $\epsilon_{p,\tau}$ and the treatment of $\epsilon_{p,\tau}$ to produce the independent stochastic variable $z_{p,\tau}$ make the positively - valued variable $x_{p,\tau}$ as a $z_{p,\tau}$ variable with both negative and positive values. The auto-regressive linear models transform the variable $\epsilon_{p,\tau}$ which is bounded on the left side to a new variable $z_{p,\tau}$ which theoretically may be unbounded.

Owing to the fact that $z_{p,\tau}$ series included many negative values, fits of normal, three parameter log normal, and three parameter gamma distributions have been considered in the present study.

The importance of the choice of a probability distribution for $z_{p,\tau}$ series is hard to evaluate, because it depends on the purpose for which the generated data have to be used. If the purpose of the study is to reproduce the extreme events, the probability distribution of $z_{p,\tau}$ will play a key role. On the other hand, if the generated data have to be used for storage determinations, the effect of the distribution of $z_{p,\tau}$ may not be very important, as suggested by Thomas and Fiering (55).

3.7.4.2. Criteria for selection of best fit distributions

The theoretical distribution of best fit to observed distribution should have the following characteristics.

1. The function is continuous and defined for all positive values of the variable z_t
2. The lower tail is bounded by zero value or positive value.
3. The upper tail is unbounded.
4. The density curve is asymptotic to the axis for larger values of z_t
5. The basic shape, is one peak, bell shaped two tailed curve, with a large variety of skewness.
6. The number of parameters, which describe the theoretical function is limited to three.

3.7.4.3. Fitting two parameter normal probability function to $Z_{p,T}$ variable :

3.7.4.3.1. Density function:

The probability density function of the normal distribution used is

$$f(z) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2\sigma^2}(z-\mu)^2\right] ; \quad -\infty \leq Z \leq +\infty \quad \ldots (3.48)$$

in which μ is the expected value of Z_t and σ its standard deviation.

3.7.4.3.2. Parameters estimation:

The maximum likelihood estimators of the parameters of the normal density.

$$\hat{\mu} = \frac{1}{N} \sum_{i=1}^{N} Z_t \quad \ldots\ldots\ldots \quad (3.49)$$

$$\hat{\sigma} = \left[\frac{1}{N} \sum_{i=1}^{N} (Z_t - \hat{\mu})\right]^{1/2} \quad \ldots\ldots\ldots \quad (3.50)$$

P_j the probability of any class interval representing the area under the probability curve is known. K_j the class interval limits can be evaluated from the corresponding cumulative distribution obtained by integrating Eqn. 3.48 by standardising the variable as u_j

$$f(u) = j\,P_j = \int_{-\infty}^{u_j} \frac{1}{\sqrt{2\pi}} e^{-u^2/2}\, du \quad \ldots \quad 3.51)$$

with $j = 1,2,\ldots\ldots,N$. From the values of u_j, and estimates of population mean, and standard deviation, $\hat{\mu}$ and $\hat{\sigma}$, the class interval limits K_j of the variable Z_t, are $K_j = \hat{\mu} + u_j\hat{\sigma}$ (3.52) in which u_j are the class interval limits of the variable u_t of the Equation (3.51).

3.7.4.4. Fitting three parameter lognormal probability function to $z_{p,T}$ variable:

3.7.4.4.1. Density function:

The three - parameter log normal probability density function is

$$f(z) = \frac{1}{\sqrt{2\pi}\,\sigma_n (z_t - z_0)} \exp\left\{ - \frac{[\log (z_t - z_0) - \mu_n]^2}{2\sigma_n^2} \right\}$$

in which μ_n is the mean of $\log(z_t - z_0)$ . . . . . . (3.53)

σ_n is the standard deviation of $\log(z_t - z_0)$, and z_0 is the lower boundary or location parameter.

3.7.4.4.2. Parameter estimation:

The lower boundary z_0 is estimated from the following equation by an iterative procedure in which an allowable error must be prescribed (6).

$$\sum_{t=1}^{N} \left(\frac{1}{z_t - z_0}\right) \left\{ \frac{1}{N} \sum_{t=1}^{N} [\log (z_t - \hat{z}_0)]^2 - \frac{1}{N} \sum_{t=1}^{N} \log (z_t - \hat{z}_0) \right\} + \sum_{t=1}^{N} \frac{\log (z_t - \hat{z}_0)}{(z_t - \hat{z}_0)} = 0 \quad \text{. . . (3.54)}$$

Once z_0 is estimated, the other two parameters may be estimated from the following

$$\hat{\mu}_n = \frac{1}{N} \sum_{t=1}^{N} \log (z_t - z_0) \quad \text{. . . . . (3.55)}$$

$$\hat{\sigma}_N = \left\{ \frac{1}{N} \sum_{t=1}^{N} \log (z_t - z_0) - \hat{\mu}_n \right]^{1/2} \quad \text{. . . . . . . (3.56)}$$

The class interval limits are computed from

$$X_j = z_0 + \exp\left[\hat{\mu}_n + u_j \sigma_n\right] \quad \text{. . . . . . . (3.57)}$$

The χ^2 - test can be performed either on

$z_{p,T}$ series or on $(z_{p,T} - z_0)$ series . . . . (3.58)

In the Equation (3.57) u_j are the class interval limits of the variable u_j from Equation (3.31)

3.7.4.4.3. Parameters estimation(alternative approach)

The maximum likelihood estimators of equation (3.56) can be found, but they require iterative simultaneous solution of three coupled, nonlinear equations.

Method-of-moments estimation of parameters which is suggested by Benjamin and Cornell (4) is straight forward. Since the standard deviation, and skewness of ($Z_t - Z_0$) equal to those of Z_t , which are in turn, estimated by the corresponding sample moments,

$$\gamma_1 = \frac{3\sigma_n}{\mu_n} + \left(\frac{\sigma_n}{\mu_n}\right)^3 \quad \ldots\ldots\ldots (3.58)$$

implying

$$\mu_n^3 - \frac{\sigma_n^3}{\gamma_1} + \frac{3\sigma_n \mu_n^2}{\gamma_1} = 0 \quad \ldots\ldots\ldots (3.59)$$

where μ_n = mean of $\log(Z_t - Z_0)$ series.

σ_n^2 = variance of $\log(Z_t - Z_0)$ series.

γ_1 = skewness coefficient of Z_t series.

Since σ_n and γ_1 (of Z_t) are known the solution of cubic equation 3.59 gives an estimate of μ_n. Then the shift parameter, $Z_0 = (\bar{Z} - \mu_n)$ ....(3.60)

Where $\bar{Z}$ = the mean of N values of Z_t

μ_n = the mean of N values of $\log(Z_t - Z_0)$, and N = sample size.

Sangal and Biswas (37) have adopted the following expression for the estimation of shift parameter

$$Z_0 = \nu_x - \frac{\sigma_x^2}{2(\mu_x - \nu_x)} \quad \ldots\ldots (3.61)$$

where ν_x = median of Z_t series computed as the mean of middle fifth of data.

μ_x = mean and σ_x = standard deviation of the Z_t series.

Sangal and Biswas have stated that 3 - parameter log normal distribution is a general skew distribution, and the coefficient of variation of the reduced variable ($z_t + z_0$) is automatically adjusted to 0.964, irrespective of the coefficient of variation of original data. The coefficient of skewness of z_t variable should be positive for adopting this distribution. This approach allows the three-parameter lognormal transformation equations to be solved explicitly-as for the two parameter situation.

3.7.4.5. Fitting the three parameter Gamma probability function to $z_{p,\tau}$ variables:

3.7.4.5.1. Density function:

The probability density function is

$$f(z) = \frac{1}{\beta \Gamma\alpha} \left(\frac{z_t - z_0}{\beta}\right)^{\alpha - 1} e^{(z_t - z_0)/\beta} \quad \ldots\ldots (3.62)$$

in which α = shape, β = scale and z_0 = lower boundary of the three parameter gamma function.

3.7.4.5.2. Parameter estimation:

The parameter estimation by the maximum likelihood estimation method suggested by Yevjevich (57) is by an interactive procedure from

$$\frac{1 + (1 + \frac{4}{3}A)^{1/2}}{1 + (1 + \frac{4}{3}A)^{1/2} - 4A} - (\bar{z} - z_0)\frac{1}{N}\sum_{t=1}^{N}\frac{1}{(z_t - z_0)} = 0 \quad \ldots\ldots (3.63)$$

for the evaluation of location parameter of the lower boundary z_0 in which

$$A = \log(\bar{z} - z_0) - \frac{1}{N}\sum_{t=1}^{N}\log(z_t - z_0)$$

$\bar{z}$ = the mean of N values of z_t

Once Z_0 is estimated, the parameter α is estimated by

$$\alpha = \frac{1 + (1 + \frac{4}{3} A)^{1/2}}{4A} - \Delta\alpha \quad \ldots\ldots \quad (3.65)$$

with A given by Equation (3.64) and $\Delta\alpha$ approximated by $\Delta\alpha = 0.04475\ (0.26)^{\alpha}$ ...(3.66). The parameter β is estimated by

$$\beta = \frac{1}{\alpha}\frac{1}{N}\sum_{t=1}^{N}(z_t - z_0) = \frac{1}{\alpha}(\bar{z} - z_0) \quad \ldots \quad (3.67)$$

with all the three, i.e. α, β and z_0 estimated, the probability density function of $Z_{p,T}$ is

$$f(z) = \frac{1}{\beta\Gamma\alpha}\left(\frac{z_t - z_0}{\beta}\right)^{\alpha-1} e^{(z_t - z_0)/\beta} \quad \ldots \quad (3.68)$$

where $\Gamma\alpha$ is the gamma function of α. It is skewed to the right for all values of the parameter α and β. The method of moments estimators of parameters of Gamma probability function suggested by Benjamin and Cornell (4) are as follows:

The shift parameter z_0 is estimated by solving the three equations.

$$g_1 = \frac{2}{\sqrt{K_1}} \quad \ldots\ldots\ldots\ldots \quad (3.69)$$

$$S = \sqrt{K_1}/\lambda \quad \ldots\ldots\ldots\ldots \quad (3.70)$$

$$z_0 + \frac{K_1}{\lambda} = \bar{z} \quad \ldots\ldots\ldots \quad (3.71)$$

where g_1 = Skewness coefficient of Z_t variable.

S = Standard deviation of Z_t variable.

$\bar{z}$ = mean of Z_t variable.

Z_0 = Shift parameter

The method of moments estimation of parameters is more straight forward than the maximum likelihood estimators which is accurate but involving iterative procedure for estimation of lower boundary location parameter.

The class interval limits are computed by using Equation (3.68) with the lower integral limit to zero. In order to use the existing Tables of Incomplete Gamma Functions, the integral of equation (3.68) is first expressed in terms of shape parameter only by using the scale parameter of equation (3.67) and by suitable transformations (20) the class interval limits are given by

$$X_j = Z_0 + \frac{(1 - z_t)\, u_j}{\sqrt{c}} \quad \ldots\ldots\ldots \quad (3.72)$$

By knowing the estimated parameters, the value of u_j can be calculated for given c from the Tables of Incomplete Gamma Functions.

3.7.4.6. Criteria for best fit distribution:

3.7.4.6.1. Chi-Square Test:

The total range of sample observations is divided into K mutually exclusive class intervals, each having the observed class frequency O_j and corresponding expected class probability E_j ($j = 1, 2, 3, \ldots\ldots\ldots, K$). Using the expected value E_j as the norm of any class interval, it is reasonable to choose the quantity $(O_j - E_j)^2$ as a measure of departure from the norm. However, the magnitudes of the squared deviations $(O_j - E_j)$ would not be comparable from one class to another, since the scale of each is nearly proportional to the expected value. Therefore a suitable measure is expected by $(O_j - E_j)^2/E_j$ and the measure of total discrepancy between observations and expectations, X^2 becomes

$$\text{Chi-Square } \chi^2 = \sum_{j=1}^{K} \frac{(O_j - E_j)^2}{E_j} \quad \ldots\ldots \quad (3.73)$$

This statistic is distributed asymptotically as Chi-Square (X^2) with (K-1-ν) degrees of freedom where ν is the number of parameters already estimated from the sample.

The choice of length of class intervals on the basis of equal probabilities (20) leads to uniform distribution of probability, and as the total value of probability integral is unity, the probability of each class interval is $p_j = \frac{1}{K}$ with $j = 1,2 \ldots K$. For this value of probability, the required length of any class interval can be obtained from the probability integral transformation. (20)

Yevjevich (37) suggest the number of class intervals to be 60 for the monthly hydrologic time series. Fisher (15) suggest that no group should contain fewer than five expected frequencies.

The sample observations should be arranged in an array in increasing order. Then to determine how many observations will fall in each of the K chosen class intervals, (K-1) interval limits are computed for each of the selected functions by the following equations, for Normal function by Eq. (3.52), log-normal-3 by Eq. (3.57) and for Gamma 3 by Eq. (3.72) respectively.

Knowing the class interval limits, the corresponding observed class frequencies are determined, squared and summed and chi - squares computed by Eq. (3.73). Equation (3.73) can be simplified for computational purposes as

$$X^2 = \frac{K}{N} \sum_{j=1}^{K} O_j^2 - N \qquad (3.74)$$

where N is the sample size.

These chi - square values for each sample, one for each of the above three probability density functions, give automatically the measure of goodness of fit of a particular theoretical function to the $Z_{p,T}$ series.

The chi - square test prescribes the critical value X_c^2 for a given confidence level (from X^2 tables), so that for $X^2 < X_c^2$ the null hypothesis of goodness of fit is accepted, and for $X^2 \geq X_c^2$ it is rejected.

The smaller probability of chi- square means the better fitting to Z_t series. However Yevjevich (97) states that, when $X^2 \leq X_c^2$ for the normal function, there is no practical need for testing the fit of other probability functions to $Z_{p,T}$ frequency distribution, which have more parameters, though their X^2 - values may be smaller.

3.7.4.6.2. Kolmogorov - Smirnov goodness-of-fit test:

A second quantitative goodness-of-fit test is based on a second test statistic. It concentrates on the deviations between hypothetical cumulative distribution function (CDF), $F_X(x_i)$ and the observed cumulative histogram (also called empirical CDF). (4)

$$F^{*}(x_i) = \frac{i}{N} \quad \ldots\ldots\ldots\ldots\ldots \quad (3.75)$$

in which x_i is the i-th largest observed value in the random sample of size N.

$$D = \max_{i=1}^{N} \left[F^{*}(x_i) - F_X(x_i) \right]$$

$$= \max^{N} \left[\left| i/N - F_X(x_i) \right| \right] \quad \ldots\ldots \quad (3.76)$$

In the case of normal distribution, for the probability density function of equation (3.68), the cumulative distribution function if the probability that X is less than or equal to x is given by

$$P_x(X \le x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{(z-\mu)^2}{2\sigma^2}} \, dz$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\left(\frac{x-m}{\sigma}\right)} e^{-t^2/2} \, dt \quad \ldots (3.77)$$

for which the power series expansion as given by Milton and Stegun (25) is

$$P(x) = \frac{1}{2} + \frac{1}{\sqrt{2\pi}} \sum_{n=0}^{\infty} \frac{(-1)^n \; x^{2n+1}}{\lfloor n \;\; 2^n (n+1)} \quad \ldots (3.78)$$

This test is based upon the fact that the observed cumulative distribution of a sample is expected to be fairly close to true distribution. The goodness-of-fit is measured by finding the point at which the sample, and the population are farthest apart and comparing this distance with the entry in a table of critical values (1). Its only parameter is n, the sample size.

For a sample size larger than 40, the K-S statistic D in equation (3.76) is given by $1.36/\sqrt{N}$ ..(3.79) where n is the sample size at 95 pct. confidence level (1).

If the distance D is too large, the chance that the observation actually came from a population with the specified distribution is very small. This concludes that the specified distribution is not correct one.

9.7.6.6.3. Comparison of chi-square and K-S tests:

a) The K-S test is superior to chi-square test in the following ways.

i) The exact-distribution of KS test is known and tabled for small sample sizes, while the exact distribution of Chi-square is known and tabled only for infinite sized samples.

ii) The K-S test is used, to test for deviations in a given direction, while the chi-square can be used only for a two-sided test.

iii) The K-S test can be used in a sequential test where data become available from smallest to largest, computations being continued only up-to the point at which rejection occurs (9)

iv) The K-S test has an advantage over X^2 test in that it does not lump data and compare discrete categories, but rather compares all the data in an unaltered form.

v) When data is plotted on probability paper, "Dmax can be computed more easily than X^2.

b) The chi-square test is superior to the K-S test in the following ways :

i) Chi-square can be partitioned and added

ii) Chi-square can be applied to discrete populations.

iii) K-S goodness-of-fit test is strictly valid only for continuous distributions and only when the model is hypothesized wholly independent of the data. Nevertheless, the test is often used for discrete distribution tests

The model: - Having identified all the components of the additive time series model, the $x_{p,\tau}$ series can be represented by

$$X_t = T_t + P_t + S_t \quad \ldots\ldots\ldots \quad (3.80)$$

where T_t = Trend component

P_t = periodic component

S_t = Stochastic component

Generating new samples of time series can be regarded as a reversible process of decomposition of a time series into their various components. The mathematical description of each of the above component, thus facilitates the generation of new samples. In the generation of new samples, for the description of random component, the random numbers of appropriate distribution are required, the generation of which, is discussed in chapter 5.

CHAPTER - 4

TIME SERIES ANALYSIS OF NON-STATIONARITY APPROACH

4.1. GENERAL:

Many investigators have considered the flow series as entirely non-stationary process for the purpose of structural analysis, and mathematical description of hydrologic time series.

The analysis of monthly volume of discharge is more complicated, than the analysis of annual flow data due to the annual cycle associated with the fluctuation in solar radiation received at the earth's surface. The effect of this fluctuation is considered to render the time series of monthly flows non stationary in both mean and standard deviation. The non-stationarity is also rendered in the correlation between successive months discharges(10).

It is clear that any river draining a basin subject to strongly seasonal rainfall is likely to be non-stationary, in the sense that mean monthly flow for one month cannot be regarded as equal, apart from sampling variation to the mean monthly flow for another. Similarly, the variance of flows observed in a particular month will also vary from one month to another.

Thomas - Fiering model is based on the assumption, that the correlations with a lag greater than one is negligible and that the serial correlation is linear, will generate an artificial record of any length (11).

The studies conducted by Matalas (22) validated the use of regression model of lag one (one month) incorporated in the equation (4.1) Matalas found that more elaborate model involving lags of two or more months did not give significantly better results. The studies conducted by Colston and Wiggert (11) also confirm the above hypothesis. Colston and Wiggert also found that the precipitation for more than a month previous did not add significantly to the overall correlation coefficient for the monthly model and thus state that the mean monthly flow is serially correlated to only the previous month's hydrologic parameters. Thomas and Fiering (33) have also assumed the absence of trends in the monthly flow sequences in the formulation of regression equation relating current to antecedent flows, and expressing the effect of a normal deviate, thus encompassing correlations coming from both periodic and stochastic components in monthly flow rates. Yevjevich (20) has shown that in some rivers a significant trend component is present in addition to periodic and stochastic components affect accuracy when any component is neglected in a synthesis of stream flow.

Clarke (10) suggest that the Thomas- Fiering method be used with caution, if fewer than twelve years data is available, in order that parameters in the regression equations be estimated with reasonable precision.

4.2. THOMAS AND FIERING METHOD:

The method of Thomas and Fiering implicitly allows for the non-stationarity observed in monthly discharge data. The method can be used for weekly, monthly, or seasonal flows also, as well as for annual flows. It does not require that the flow data be normally distributed, and may be used with skewed distributions as well; it incorporates serial correlation between successive flows so as to accord with observed stream flows; and by means of the cross-correlation technique, the method is adopted to the design of multi reservoir project also. (35)

In its simplest form, the Thomas and Fiering method consists of the use of twelve linear regression equations. If say, twelve years of record are available, the twelve January flows and the twelve December flows are abstracted and January flow is regressed upon December flow; Similarly, February flow is regressed upon January flow, and so on for each month of the year. Using the Thomas - Fiering notation the model may be written as

$$Q_{i+1} = \bar{q}_{j+1} + b_j (Q_i - \bar{q}_j) + t_i S_{j+1} \sqrt{1-r_j^2} \quad . . . (4.1)$$

In the equation 4.1, Q_i and Q_{i+1} are the volumes of discharge during the i-th and (i+1) th month respectively, $\bar{q}_j$ and $\bar{q}_{j+1}$ are the mean monthly discharges during j-th and (j+1) th months respectively, within a repetitive annual cycle of 12 months; b_j is the regression coefficient for estimating the volume of discharge in the (j+1) th month from the j th month; t_i is a random normal deviate with zero mean and unit variance; S_{j+1} is the standard deviation of discharges in the (j+1)th month; and r_j is the correlation coefficient between flows in the j-th and (j+1)-th months.

Thus the synthetic flows for the (i+1) th month is the sum of three terms. First is the mean of the flows for (i+1)th month of the historical flows, referred to as (j+1)th month, second is the difference between previous flow (month i and its mean multiplied by the regression coefficient $r(j)(\sigma_{j+1}/\sigma_j)$ and third is z_i which is normally distributed with zero mean and unit standard deviation so that when it is multiplied by the standard error have the desired variances.

The equation (4.1) can also be adapted to give flows that are distributed approximately as gamma with mean, μ variance σ^2 and skewness γ_x, when the observed flows show considerable skewness. Fiering and Jackson (15) suggest introduction of γ_z as the skewness of the random component of each flow event. It is different from the skewness of the flows because the sum of gamma variates, are not necessarily gamma.

For $\gamma_z = \frac{1-\rho^3}{(1-\rho^2)^{\frac{3}{2}}}\gamma_x$ (4.2) where ρ is the lag-one correlation coefficient of the flows and γ_x is the skewness coefficient of flows for the month under consideration. The normal random deviate of Eqn. 4.1 is to be replaced by a random variate that has approximately a gamma distribution. Then a modified random sampling variate z_i' is defined by

$$z_i = \frac{2}{\gamma_z}\left(1 + \frac{\gamma_z z_i}{6} + \frac{\gamma_z^2}{36}\right)^3 - \frac{2}{\gamma_z} \quad \ldots\ldots (4.3)$$

This deviate is distributed approximately as gamma with mean zero, variance unity, and coefficient of skewness γ_z. Thus the skewness of the historical record is preserved.

When Thomas-Fiering model is fitted to monthly stream flows, it may be found that the values in the generated sequences are sometimes negative. This may be due to normal distribution assigning non zero probability, to negative values (15).

Negative values, when encountered, should be used to derive the subsequent values in the sequence; once the generated sequence is complete, however, all negative values in the generated sequence are replaced by zero.

Thomas and Fiering (55) recommend the usage of the logarithmic transformation of the observed flows, instead of using the observed flows themselves, or its square root transformation. Thomas and Fiering states that, a logarithmic transformation best fitted a normal probability distribution, and has the further advantage of eliminating the nega-tive flows that occur occasionally when untransformed flows are used in the model. However the study conducted by Thomas and Fiering (55) indicate that the total amount of water added to the system by increasing the negative values found in the generated sequences using untransformed flows was negligible.

The variance of the logarithmic series may be drastically changed depending on the increment added to the original series and therefore it is necessary in every case to test several different increments before choosing the most appropriate one. Rodriguez et al (29) recommend, an increment of 0.012 times the average individual monthly flow to be added to each historical flow before computing its logarithm to avoid infinite logarithms, when flow is zero. Yevjevich (57) recommends the zero flows to be replaced by a very small positive value, such as 0.010 or 0.0001 or some other similar small value.

When the modelling uses the above Markov generation scheme to produce the logarithms of flow we must remember that the procedure reproduces the mean, variance, serial correlation coefficient, the skewness coefficient of the logs of the flows. The serial correlation and the skewness coefficient of the flows themselves are not preserved (13). In practice, the resultant distortion may be important, the resultant distortion may be important and Matalas (22) has suggested procedure for ensuring that the moments of the flows are maintained.

Clarke (10) suggests the procedure to be followed, for the generation of synthetic sequences of monthly flows, when some months, frequently have zero flows.

4.5. GENERALIZATION OF THE THOMAS-FIERING MODEL:

The Thomas Fiering model may be generalized by the inclusion of other variables, such as rainfall, in the equations defining the model.

The model suggested by Colston and Wiggert (11) is as follows:

$$Q_i = a + b_1\, f(P_i) + b_2\, f(Q_{i-1}) + b_3\, f(P_{i-1}) + b_4\, f(\text{cross products}) \quad ..(4.4)$$

Where

Q_i = mean monthly flow of month i ; P_i is the total basin precipitation for month i ; Q_{i-1} is the mean monthly flow for month previous to month i ; P_{i-1} is total basin precipitation for month previous to month i ; i is the index running 1 to 12, denoting months; a is constant; b is coefficient; and f is some function.

The generation of samples, by the non - stationary approach using the Thomas Fiering model and also by the stationarity approach using the time series models require the generation of random numbers of approprite distribution for the description of random component part of the model, is discussed in the next chapter.

CHAPTER - 5

GENERATION OF RANDOM NUMBERS

5.1. GENERAL:

An element, whose occurrence can be entirely attributed to chance is called random element. The random numbers generated in a computer are called pseudo-random numbers, as they are produced by an algorithm, and are not truely random. The method chosen for generating random numbers, must yield random numbers with the following properties.

i) Uniformly distributed

ii) reproduceable

iii) Statistically independent

iv) non-repeating for any desired length

v) capable of generating random numbers in minimum time.

vi) requiring minimum computer memory.(25)

The method to get random numbers (RN) of required distribution consists of first, generating numbers of rectangular distribution (i.e. a distribution of values, continuous over the interval (0,1), such that the probability that a random variate lies in any interval of length dy is itself dy) and then transforming the rectangularly - distributed variates into others having the required distribution. As a part of this two stage procedure the method available for deriving the uniformly - distributed variates is discussed first (10).

5.2. GENERATION OF RECTANGULARLY DISTRIBUTED VARIATES:

Congruential methods (32) use deterministic formulae for the generation of random numbers. These methods have been developed from one originally proposed by Lehmer (19,10). His original multiplicative congruence method used the recurrence relation.

$$x_i = a\, x_{i-1} \ (\text{modulo } m) \quad \ldots\ldots\ldots\ldots \quad (5.1)$$

meaning that x_i is the remainder when $a\, x_{i-1}$ is divided by m, and this has been generalized to the relation.

$$x_i = (a\, x_{i-1} + c) \ (\text{modulo } m) \quad \ldots\ldots \quad (5.2)$$

meaning that x_i is the remainder when $a\, x_{i-1} + c$ is divided by m. In the equations (5.1) and (5.2) m is a large integer determined by the design of the computer usually a large power of 2 or 10, (2 is applicable for IBM 1620). And a,c,x_i are integers between 0 and m-1. The numbers x_i/m then form a sequence having a rectangular distribution.

Much care is necessary in the choice of the values of a,c, and m used; the sequence (x_1, x_2 ......) must eventually repeat itself, so that it is preferable to describe it as a sequence of pseudo - random number rather than a sequence of random numbers. If the sequence repeats itself after x_p (that is, after p pseudo - random numbers have been generated), then p will depend upon the choice of a, c and m; it is therefore particularly important to choose these integers to make p as large as possible.

5.3. THE GENERATION OF SEQUENCES OF PSEUDO-RANDOM NUMBERS WITH DISTRIBUTIONS OTHER THAN RECTANGULAR:

Sequences with distributions other than rectangular are obtained by transforming rectangularly-distributed variates to the required distribution (10) as follows:

5.3.1. Normally distributed random numbers generators:

5.3.1.1. a) Box and Muller's Method:

Box - Muller's method requires the prior generation of variates x_1 and x_2 rectangularly - distributed over the interval (0,1); these are then transformed to values y_1 and y_2, where

$$y_1 = (-2 \log_e x_1)^{1/2} \cos (2\pi x_2) \quad \ldots\ldots \quad (5.3)$$

$$y_2 = (-2 \log_e x_1)^{1/2} \sin (2\pi x_2) \quad \ldots\ldots \quad (5.4)$$

The values of y_1 and y_2 are normally and independently distributed with zero mean and unit variance.

5.3.1.2. b) Method based on central-limit theorem:

The central limit theorem states that, under certain very broad conditions, the sum of a sequence of independent random variables approximates a normal distribution, whatever the distribution of the random variables in the sequence.

In particular, the central-limit theorem holds when the independent random variables summed are sampled from a rectangular distribution. Values $x_1, x_2, \ldots, x_n$ are therefore generated to follow a distribution rectangular over the interval (0,1) and the quantity y is then calculated, where

$$y = x_1 + x_2 + \cdots + x_n - \tfrac{1}{2}n \quad (5.5)$$

If n = 12, the distribution of the values y then closely approximates a normal distribution with zero mean and unit variance.

5.3.2. Normally - distributed pseudo - random numbers; distribution to have a given μ and standard deviation σ :

Having obtained normally distributed pseudo-random numbers with zero mean and unit variance, it is then simple to obtain normal pseudo - random numbers with any mean μ and any standard deviation σ . The procedure is as follows:

(i) Generate a pseudo - random normal deviate y (from N (0,1)).

(ii) Calculate $Y = \mu + y$; then Y is also pseudo-random normal deviate with mean μ and standard deviation σ .

5.3.3. Generator for independent Log-Normal random numbers distribution to have a given and standard deviation σ :

The procedure is as follows:

(i) Generate a pseudo - random normal deviate y (from N (0,1)).

(ii) Transform to $Y = \exp(\mu + \sigma y)$; then Y has a log-normal distribution with parameters μ and σ .

For three parameter log-normal random numbers generation with mean μ_n and standard deviation σ_n, the transform

$$Y = Z_0 + \exp(\mu_n + \sigma_n y) \quad . \; . \; . \; . \; . \; (5.6)$$

where Z_0 = lower boundary has a 3-parameter log Normal distribution with parameters μ_n, σ_n and Z_0 .

5.5.6. Generator for independent gamma distributed random numbers:

Let t_i be normally distributed random numbers with (13) zero mean and unit variance.

Suppose x_i is a variable where $i = 1,2,\ldots\ldots\ldots,N$ Then mean of x_i values is

$$\bar{x}_i = \frac{1}{N}\sum_{i=1}^{N} x_i \quad \ldots\ldots\ldots\ldots\ldots (5.7)$$

Next calculate S_i - standard deviation of variable by

$$S_i = \sqrt{\frac{\sum_{i=1}^{N}(x_i - \bar{x})^2}{N-1}} \quad \ldots\ldots\ldots\ldots (5.8)$$

The lag one serial correlation coefficient r_1 by the equation (5.13) where K = 1

Then calculate $\gamma_x = \frac{1 - r_1^3}{(1-r_1^2)^{3/2}} \; \gamma_x \quad \ldots\ldots\ldots (5.9)$

Where γ_x = skewness coefficient of the variable.

Then G_i, is distributed as gamma with mean zero, variance unity and coefficient of skewness γ_x is defined by

$$G_i = \frac{2}{\gamma_x}\left(1 + \frac{\gamma_x t_i}{6} - \frac{\gamma_x^2}{36} \right) - \frac{2}{\gamma_x} \quad \ldots\ldots (5.10)$$

In most of the computer - centres standard subroutines for the generation of uniformly distributed random numbers with distribution (0,1) are available and the same can be used to generate desired distribution of random numbers.

CHAPTER - 6

DESCRIPTION OF WATERSHED SELECTED FOR ANALYSIS

6.1. GENERAL:

Cauvery is a East flowing river, and takes its birth in Coorg district of Karnataka. After travelling the state it flows through Tamilnadu before it joins Bay of Bengal.

The average annual rainfall in the Cauvery basin of the state, is about 56" , out of which about 67% falls during monsoon period of June to September, about 14 pct. during March to May, and remaining about 19 pct., during the post - monsoon period of October to December. In the Ghat region, the peak rainfall is normally observed in July.

The principal types of soil in the basin are red loams, red sandy soils, laterites and lateritic soils, and black and dark brown soils.

Lakshmana Thirtha is one of the tributaries of Cauvery. Lakshmana Thirtha from the South West direction join the river Cauvery, at the confluence of which, is situated the Krishnaraja-sagar reservoir, famous for its Brindavan Gardens, which is a tourist attraction. Lakshmanathirtha river rises in Western Ghats, in Coorg district, at an altitude of about 6500 feet. It has a total length of 82 miles, and drainage area of 690 Sq. miles, with average annual rainfall, varying from 100" to 29" .

In this study the Lakshmanathirtha subbasin of Cauvery basin, for which long term rainfall , and runoff data were available is considered as the watershed, and treated as the stochastic hydrological system for the analysis of runoff and rainfall in the catchment. The location of gauging station at Undwadi, and nine raingauge stations in the catchment, for which data was available, is shown in figure 1.

6.2. UNDWADI (The Discharge Station):

Undwadi is the only important gauging station located on the Lakshmanathirtha river. The catchment area at this site is 580 Sq. miles, out of which about 511 Sq. miles lie in the maidan area and the balance 69 Sq. miles lie in the Malnad area. The Malnad portion is the hilly area of the Sahyadri Hills in the Western Ghats, having an average width of about 22 miles and mean elevation of about 3300 ft, and the Maidan is a plane area lying to the east of Malnad with mild undulations, low granite mounds out cropping, here and there. No major irrigation project is existing in the upstream of this gauge site. This is a current meter gauge site, and is located at about 26 miles upstream of Krishnarajasagar reservoir.

The following nine reporting raingauge stations in the catchment of, the gauge-discharge station have been considered in the present study for which a long term concurrent data was available for a period of 30 years. i) Hunsur ii) Kushal iii) Nagarhole iv) Karmad (v) Srimangala vi) Balele vii) Ponnampet viii) Titimati, and ix) Piriyapatna

6.3. DATA USED IN THE ANALYSIS:

The monthly rainfall data for 38 years from 1935 to 1972 for the above nine raingauge stations, being within the catchment of Lakshmanathirtha at Unduwadi, and the monthly run-off data measured at Unduwadi, for the con current period, were considered for the study of structural analysis of time series of rainfall and run off, in the present study.

CHAPTER - 7

DATA COLLECTION, ANALYSIS AND DISCUSSION OF RESULTS

7.1. ANALYSIS OF AVAILABLE POINT RAINFALL AND RUN OFF DATA:

The monthly runoff data at Udurodi on Lakshmanthirtha subbasin and the rainfall data of nine raingauge stations in the catchment, for a period of 33 years i.e. from 1939 to 1972, were used in the present study for structural analysis of monthly rainfall and runoff sequences and the mathematical description of the components implied in the sequences. The computer programs for the present study were written in Fortran language and were run on IBM 1620 computer at Structural Engineering Research Centre, Roorkee, and IBM 360 Computer at India/ Meteorological Department, New Delhi.

The following are the nine raingauge stations in the catchment (Figure 1), for which a concurrent data of 33 years were available.

1) Hunsur
2) Piriyapatna
3) Titimati
4) Murkal
5) Ponnampet
6) Kesmada
7) Srimangala
8) Nagarhole
9) Balelev

Thiessen polygons (Figure 1) were constructed for the above nine raingauge stations in the catchment and monthly weighted average rainfall was calculated using the computer program No. 2. In the Thiessen polygon method rainfall at a station is weighted by the area commanded by that station and weighted average rainfall found out (table No. 2 and 3).

TABLE NO. 1

MEAN MONTHLY FLOWS OF LAKSHMANATHIRTHA RIVER AT UNDUWADI
IN CUSEC DAYS

YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
1935	8	10	11	12	16	7	942	630	808	133	4	2
1936	7	6	4	6	4	437	1557	867	132	138	49	6
1937	4	17	10	8	9	2	1534	613	23	169	19	4
1938	4	9	4	7	9	347	1006	299	105	475	2	0
1939	1	0	0	0	0	0	1625	1178	240	316	72	2
1940	6	10	6	6	101	747	2764	1496	195	125	923	74
1941	52	23	13	7	18	641	2207	1527	290	626	9	51
1942	6	15	5	15	99	270	2878	1462	425	160	4	4
1943	36	9	7	3	180	375	3529	520	329	601	379	20
1944	13	14	20	13	16	22	2209	884	63	167	97	11
1945	5	14	23	13	5	8	1956	718	133	156	5	1
1946	9	2	1	3	5	535	2017	2962	840	253	679	240
1947	59	28	92	23	14	2	1128	873	748	803	10	25
1948	14	10	5	13	52	773	1569	3203	434	62	40	17
1949	6	7	5	9	15	290	1004	1003	349	98	8	5
1950	5	8	8	6	5	54	2210	1134	1036	286	30	12
1951	13	16	17	17	30	186	1287	629	47	240	16	16
1952	17	18	18	16	15	137	620	909	666	362	14	14
1953	17	16	12	12	12	206	3416	3121	249	482	33	14
1954	5	14	13	7	25	418	2142	1636	192	479	13	24
1955	17	8	7	7	257	609	560	160	328	451	93	3
1956	8	8	3	2	15	414	2921	1590	162	539	355	15
1957	14	13	20	9	411	642	[illegible]	1193	118	136	292	20
1958	12	7	7	39	366	171	3513	1330	905	360	147	33
1959	13	14	14	14	44	689	8746	1760	1700	373	188	70
1960	10	17	27	46	115	197	2014	1433	92	50	315	15
1961	9	3	16	24	426	1316	15019	4112	1256	1128	566	293
1962	188	113	69	140	726	37	3653	2305	1010	2171	258	310
1963	76	73	55	24	52	154	1726	1158	272	257	20	25
1964	11	17	11	16	40	51	1712	6246	592	831	291	30
1965	28	36	39	41	44	13	2469	239	26	9	5	13
1966	5	7	12	5	15	1	634	394	101	118	83	20
1967	19	16	4	3	15	598	2041	1560	77	27	4	4
1968	2	0	0	3	9	94	1619	1398	104	55	3	7
1969	2	0	0	0	26	18	1680	1251	320	31	46	14
1970	5	4	4	8	26	280	712	891	69	1069	205	21
1971	18	11	9	9	42	1362	763	424	200	123	21	5
1972	3	4	4	5	331	72	1723	292	212	453	54	55

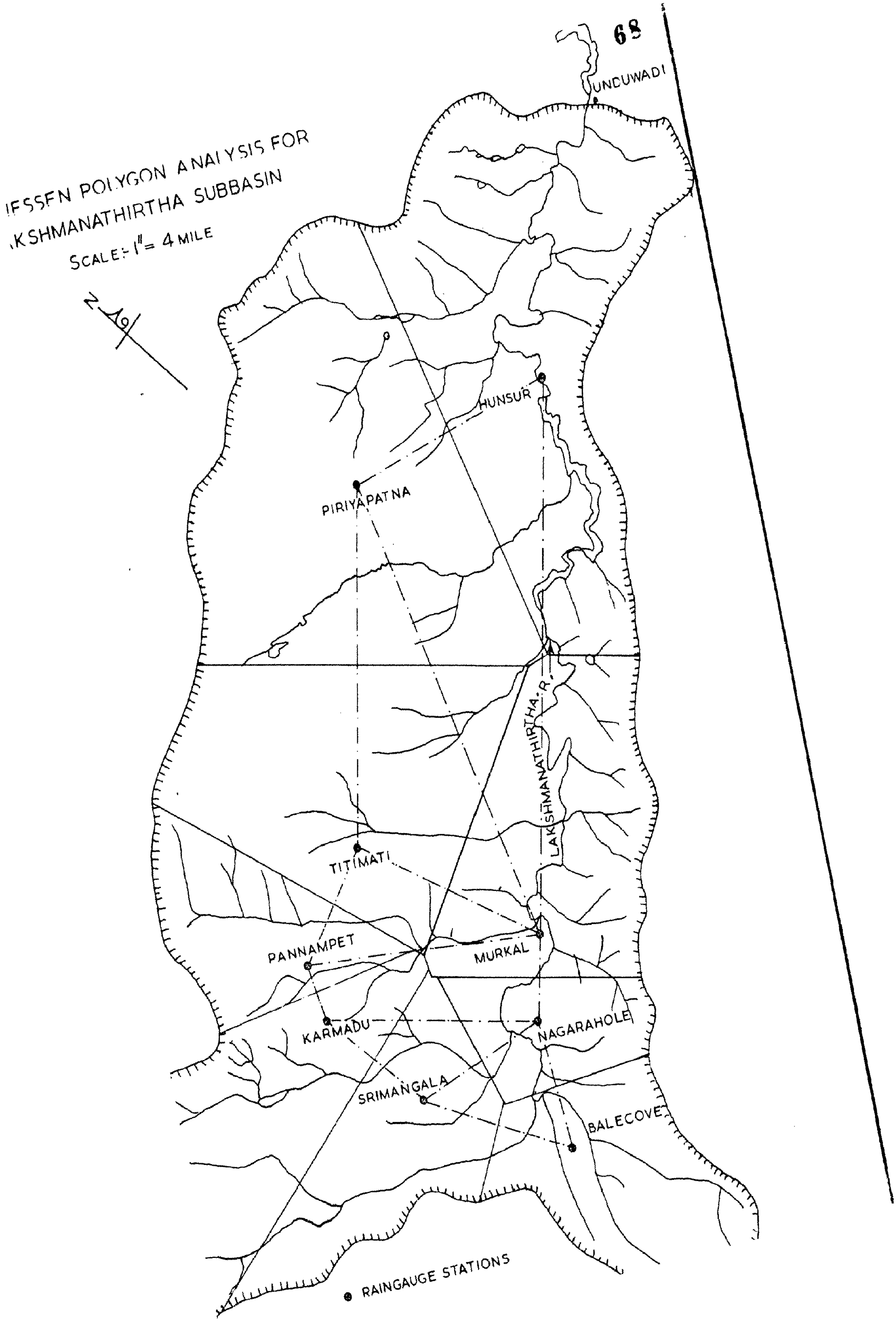

IESSEN POLYGON ANALYSIS FOR
AKSHMANATHIRTHA SUBBASIN
SCALE: 1" = 4 MILE
N
UNDUWADI
HUNSUR
PIRIYAPATNA
LAKSHMANATHIRTHA R.
TITIMATI
PANNAMPET
MURKAL
KARMADU
NAGARAHOLE
SRIMANGALA
BALECOVE
RAINGAUGE STATIONS

FIG.

TABLE NO. 2

TABLE SHOWING THIESSEN POLYGON AREAS

Sl.No.	Name of Station	Thiessen Polygon Area in Sq. Miles
1.	Hunsur	115
2.	Piriyapatna	121
3.	Titimati	83
4.	Murkhal	63
5.	Ponnampet	52
6.	Karmadu	45
7.	Srimangala	41
8.	Nagarhole	25
9.	Palekov	55
	Total area:	580 Sq miles

TABLE NO. 3

RAINFALL OF LAKSHMANATHITHA SUB-BASIN IN INCHES

YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
1935	.44	0.00	.37	3.08	2.96	4.58	14.52	8.40	6.67	6.61	.39	.66
1936	0.00	.19	3.01	1.15	5.96	14.76	9.74	7.38	6.48	5.25	2.66	.03
1937	0.00	.88	2.17	4.71	4.17	4.91	16.78	7.12	2.89	6.67	2.52	.19
1938	0.00	.12	1.81	4.31	.90	10.20	7.70	6.72	3.72	2.94	1.21	.27
1939	.52	0.00	.45	3.83	3.35	5.73	13.40	10.76	3.11	10.79	3.59	0.00
1940	0.00	0.00	0.00	1.89	5.15	14.60	12.39	8.87	2.49	6.20	10.22	.28
1941	0.00	0.00	0.00	2.73	4.43	10.54	12.24	9.61	5.32	6.08	1.66	1.95
1942	0.00	0.00	.30	4.23	7.58	8.09	17.58	6.02	2.73	4.99	1.06	1.59
1943	1.65	0.00	.14	3.01	9.64	5.33	19.70	2.65	3.97	7.67	5.15	0.00
1944	0.00	.42	1.26	.71	4.90	5.00	19.07	3.89	3.31	6.01	3.28	.66
1945	0.00	.08	0.00	2.04	3.35	5.61	14.24	6.16	2.13	4.84	1.51	0.00
1946	0.00	.20	1.83	4.55	2.70	10.96	11.05	15.44	6.62	4.20	6.19	2.17
1947	.25	.20	1.18	3.11	3.50	2.85	12.57	10.15	5.60	7.52	0.00	1.06
1948	.15	0.00	0.00	4.73	5.18	9.15	10.32	16.10	2.24	3.62	2.15	.21
1949	.05	0.00	.12	2.86	6.94	4.77	14.35	6.09	4.83	5.52	.32	0.00
1950	0.00	1.24	0.00	.75	3.86	4.82	20.20	7.13	7.50	5.18	3.29	0.00
1951	0.00	0.00	1.09	5.31	5.18	6.01	11.49	4.70	6.46	4.93	2.19	0.00
1952	0.00	.39	0.00	2.47	4.59	6.15	9.61	6.94	2.77	8.45	0.00	2.25
1953	0.00	.03	.40	2.03	2.63	9.36	26.92	15.14	3.44	9.39	.12	0.00
1954	.02	.24	.94	2.12	6.24	9.28	11.37	10.04	2.96	7.92	.01	.95
1955	0.00	0.00	.68	2.89	10.12	9.16	4.51	5.01	8.13	8.85	.48	.41
1956	.06	.17	.02	5.52	2.17	12.95	13.69	7.36	3.11	9.81	4.49	0.00
1957	0.00	0.00	.86	1.38	12.02	8.87	16.29	7.91	1.48	6.70	4.29	0.00
1958	.24	0.00	1.15	6.43	9.77	5.36	24.20	8.45	4.44	4.84	2.81	0.00
1959	0.00	0.00	0.00	4.68	5.50	19.64	28.81	8.28	9.32	2.79	3.08	.50
1960	.24	0.00	1.38	6.73	6.70	6.09	14.47	6.61	2.55	4.16	6.57	0.00
1961	0.00	.25	.26	5.74	10.32	13.92	46.01	11.07	3.09	6.70	.89	.05
1962	.21	1.05	.52	5.31	10.68	2.14	19.17	10.78	4.94	11.59	1.42	2.79
1963	.29	.58	1.26	1.99	3.26	5.13	15.01	7.60	2.94	7.13	.16	.88
1964	0.00	.02	.15	2.68	3.32	6.04	14.98	21.57	4.20	8.02	2.60	.13
1965	.04	0.00	.51	2.08	3.17	4.03	15.64	2.44	4.54	2.14	1.15	1.44
1966	.04	.06	.21	.65	3.19	2.82	12.38	3.31	6.41	4.49	5.32	.27
1967	2.73	0.00	0.00	2.68	5.70	6.88	19.44	6.32	1.34	2.88	.97	.48
1968	0.00	.38	.70	5.37	5.22	7.16	16.35	5.39	4.56	4.12	1.80	.33
1969	.03	0.00	.73	3.77	6.94	3.96	18.73	9.68	4.33	5.33	2.56	1.30
1970	.27	0.00	.89	5.36	6.95	4.77	18.78	11.57	3.86	13.03	3.97	0.00
1971	.25	.25	.06	4.16	5.51	12.74	10.22	6.03	7.76	5.90	.92	0.00
1972	0.00	.07	0.00	1.43	12.49	6.91	12.46	4.55	7.10	5.76	.58	2.20

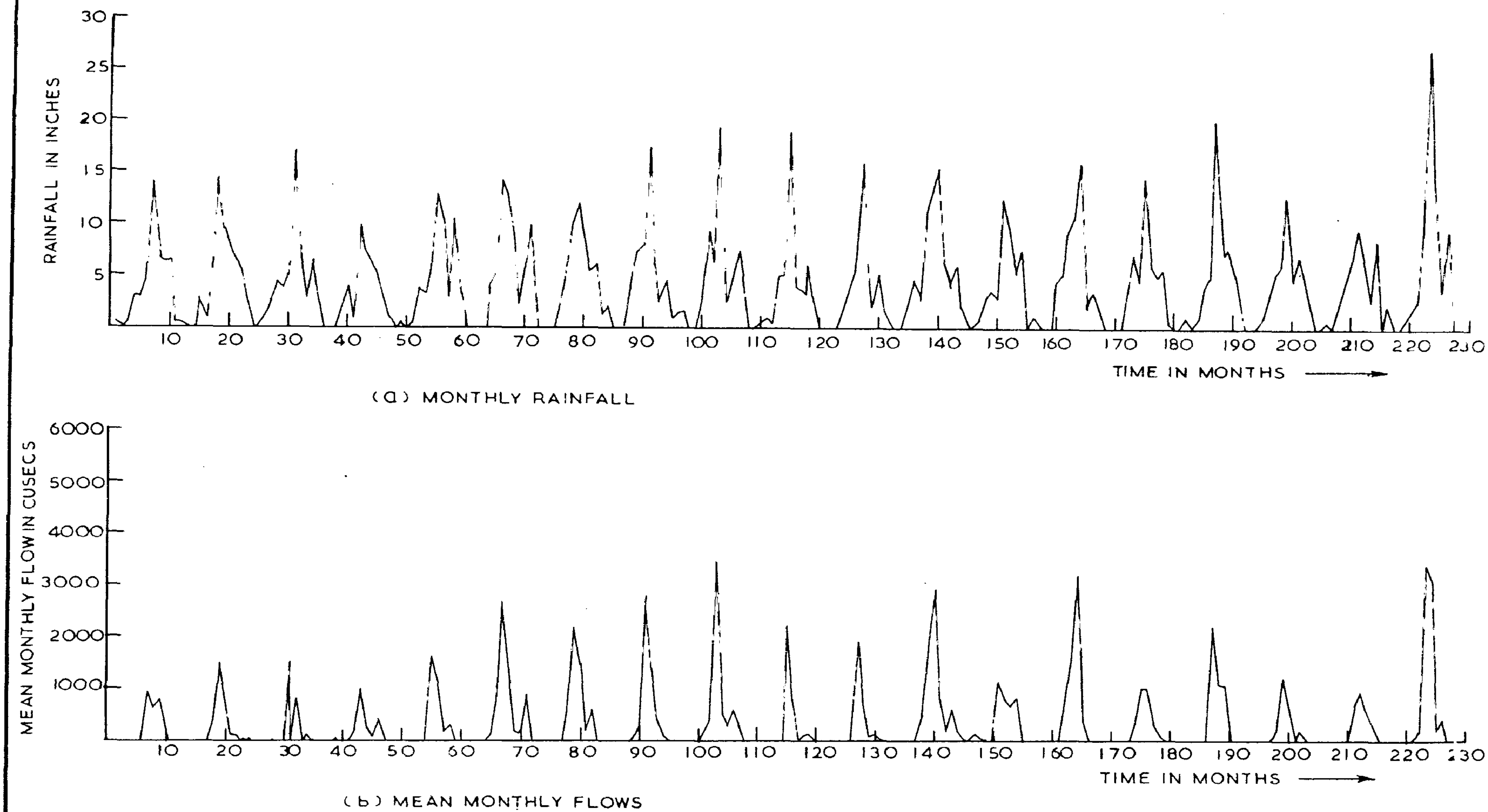


FIG. 2(i) LAKSHMANATHIRTHA SUBBASIN (DATA FROM 1935-54)

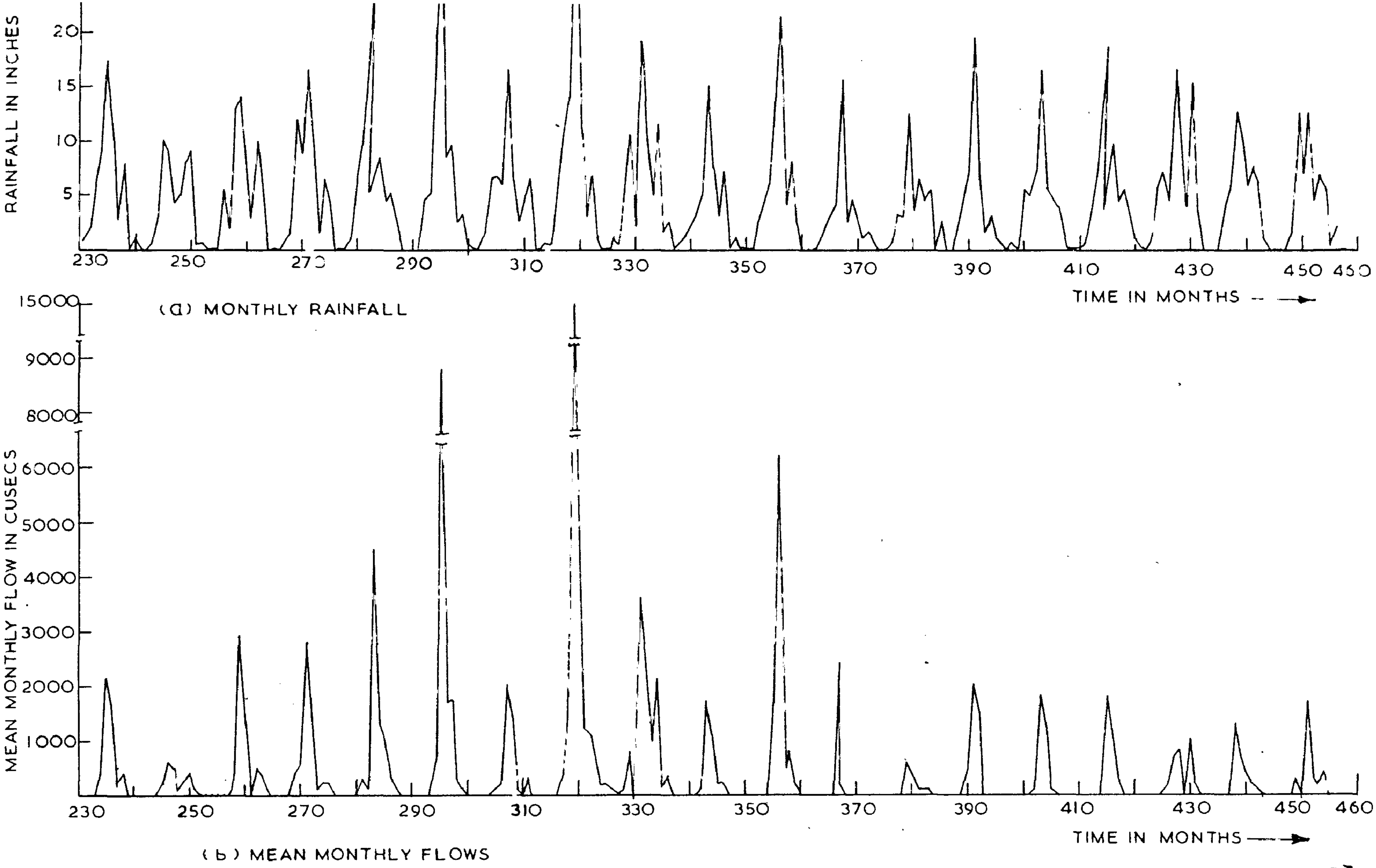


FIG.2(ii) LAKSHMANATHIRTHA SUBBASIN (DATA FROM 1954-72)

In the simulation of stream flows it is immaterial whether total volumes of flow or average rates of flow are synthesized, since one is a known multiple of the other and . If averaged rates are considered for analysis, averaging intervals for successive periods of flow need not be constant. For analysis in this study the mean monthly flows are calendar month averages, in cusecs are used.

The monthly rainfall in inches and mean monthly flows in cusecs, of the catchment, used in this analysis are given in table Nos. 1 and 3 respectively and also plotted in figure 2. A study of the monthly runoff sequence of Fig. 2 reveals that there are relatively high values of flows during 1959, 1961 and 1964.

7.2. [illegible] TREND IN THE DATA

7.2.1. Run Off Sequences:

The method of least squares was used in the present study for fitting a regression line $y = mx + c$ vide equation (2.4), (2.5) and (2.6) and the equation for the trend line was obtained as $y = 0.4096\, x + 348.09$ ....(7.1) from the results of computer program No. 1, where y is representing the flows and x the number of month. The slope of this regression line was compared with a horizontal line, by a t-statistic (equation 2.11) and value of t-statistic was found to be 1.0449 which is less than the tabulated value 1.96, at 95 pct. confidence level (from computer program No. 3). Hence it is inferred that the trend in the data is statistically not significant and is due to chance fluctuations or a sampling bias rather than the non-homogeneity in data. Usually this sampling trend is shown in any sequence. The confidence limits for above regression line were found to be 541.3005 and 541.0294 from equation (2.9), and the regression line for the data period lies within these confidence limits, further confirming that the trend in the data is not significant.

The split sample approach of testing the homogeneity of monthly runoff data, was also used, to calculate t-statistic. The runoff data of 38 years was divided into two unequal sub-samples of 21 and 17 years, and the means of the sub samples were 594.66 and 948.01 respectively. S, the pooled standard deviations of the two sub samples, was 1495.00. The t-statistic of equation (2.2) was 1.02 against the critical value of 1.96 at 95 pct. probability level. Hence the entire series may be inferred as homogeneous, and confirm the absence of significant trend in the data.

As, long term trend was not significant in the observed sequence, the trend component was taken as the original series mean itself i.e. 3.9626 for the log. transformed data and 441.16 for the untransformed runoff sequence.

7.2.2. Rainfall sequence:

The trend component in the monthly rainfall sequence was also examined for significance. The regression line (equations 2.4 to 2.6) was found to be

$$y = 0.0015\,x + 4.3705 \quad \ldots\ldots\ldots\ldots\ldots \quad (7.2)$$

TABLE NO. 4

STATISTICS OF OBSERVED RUNOFF [illegible]

S. No.	Month	Mean in cusecs	Std. dev. in cusecs	Ser. coeff. "r1"	Coeff. of Skewness C_s	Log. Coeff. Ej	Coeff. of var C_v in Pct.	Remarks
1.	January	10.61	31.30	0.94	4.63	0.69	169	
2.	February	15.71	20.17	0.75	3.03	0.70	130	
3.	March	15.00	13.93	0.65	2.76	0.70	123	
4.	April	15.47	29.06	0.60	4.36	4.69	149	
5.	May	92.04	150.53	0.16	2.50	0.36	163	
6.	June	325.65	347.50	0.49	1.49	3.35	106	
7.	July	2459.02	2503.53	0.42	3.33	0.19	101	
8.	August	1433.11	1173.73	0.53	2.40	0.13	83	
9.	September	530.94	306.22	0.40	1.04	0.42	93	
10.	October	370.52	403.71	0.23	2.74	0.12	106	
11.	November	140.84	206.93	0.50	2.27	0.20	146	
12.	December	53.34	73.54	0.03	3.14	0.36	135	

Series mean = 441.16 Cusecs

Series Std. Dev. = 1033.93 Cusecs

Range = 41610.02 Cusecs

First Ser. Cor. Coeff (r_1) = 0.53

and the confidence limits for this trend line were found to be 5.1313 and 4.1047, from equation (2.9) and results of computer program no. 3. The regression line, for the data was found to lie within the above confidence limits.

The t - statistic of equation (2.11) was found to be 0.6991 which is less than the tabulated value of 1.96 at 95 pct. confidence level. Thus it can be inferred that trend was not significant in the monthly rainfall sequence. Consequently in the analysis of monthly rainfall sequences the trend component was taken as the series mean itself i.e. 4.6633 inches.

7.3. PERIODIC AND STOCHASTIC COMPONENTS

The composition model of periodic and stochastic components of the flow series is

$$x_{p,\tau} = \mu_\tau + \sigma_\tau \; \epsilon_{p,\tau}$$

where both μ_τ and σ_τ have periodic components, with significant harmonics, and/or different phases in the case of some significant harmonics. As explained in section (3.4) if μ_τ and σ_τ are proportional to each other, the Fourier analysis for the computation of harmonic coefficients is permissible, by transforming the x_p series into their logarithms.

The runoff sequence in the present study indicates that μ_τ and σ_τ can be assumed to be proportional to each other on the basis of values of monthly means and standard deviations given in table no. 4. Consequently as described in section (3.4) the logarithmic transformation of runoff sequence facilitates (equation 3.4) the application of Fourier series analysis for the estimation of significant harmonics, by fitting periodic functions to μ_τ alone for the log-transformed data.

In this study in addition to the logarithmic transformation of the runoff sequence, i) the classical approach in which no transformation is involved and, ii) non-parametric method of separating periodic and stochastic components, have been used, for the removal of periodic component in order to facilitate a comparitive study of three cases.

The classical approach, normally followed by many investigators, in the estimation of significant harmonics in X_p series, is of the type (from equation 3.4)

$$x_{p,\tau} = \mu_\tau + \sigma_x \epsilon_{p,\tau} \cdots\cdots\cdots\cdots\cdots$$

in which μ_τ is the periodicity in mean, and σ_x is the standard deviation which is assumed to be constant.

The non-parametric method of separating periodic and stochastic components in monthly runoff sequences involve 24 parameters i.e. 12 monthly means, and 12 monthly standard deviations as explained in section (3.3)

The maximum number of harmonics in the monthly runoff and rainfall sequences is taken as six, representing 12 month main harmonic and 6,4,3,2.4 and 2 - month sub harmonics. This does not mean that there are six physical cycles in the monthly flows. Real physical cycles, and the number of harmonics needed by the Fourier analysis to describe the periodic component mathematically are not necessarily identical.

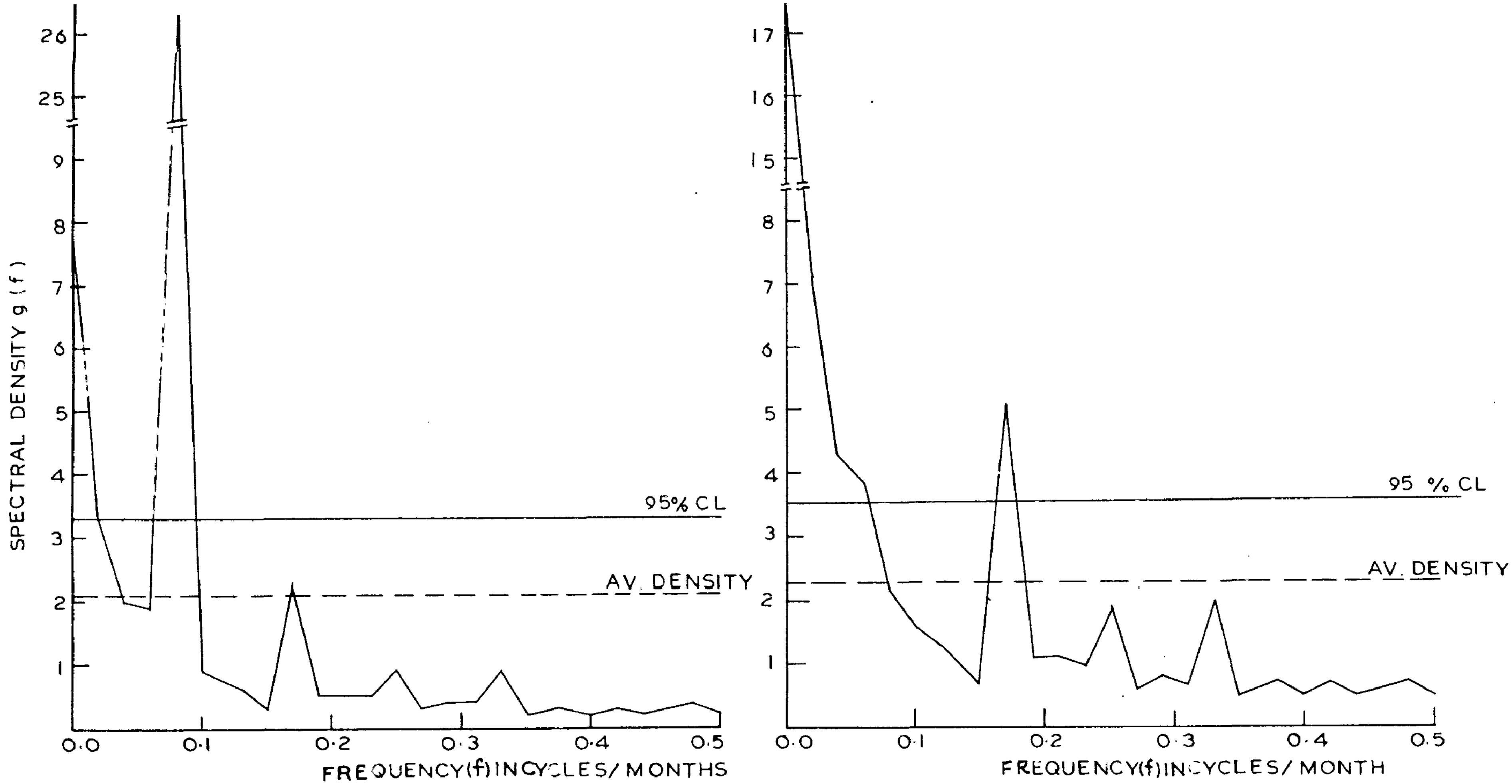
7.3.1. Analysis of Logarithmically Transformed Monthly Runoff Sequence:

The run-off data is transformed into their logarithms and all the further analysis is made with the transformed data only. To avoid infinite logarithm when the flow is zero, the zero flows have been replaced by a small positive value of 0.001.

The next step in the analysis is to detect the periodic movement inside the series, and to approximate it by Fourier series analysis, in specifying the coefficients for the main cycle and its subharmonics. To fit periodic movements about mean, by appropriate mathematical models, the correlogram and variance spectrum analysis (Chapter 3) were used to detect the significant harmonics.

Computer program no. 3 was run, to obtain the Fourier coefficients from the periodogram analysis, and also, for the correlogram analysis, and also, for the correlogram analysis, and the spectrum analysis of the log. transformed runoff data and the series with each of the six harmonics corresponding to 12,6,4,3, 2.4 and 2 - month period, removed in turn.

As the maximum lag m should be between N/10 to N/3 where N is the sample size, a maximum lag equal to 48 and also equal to 96 were chosen for the correlogram and spectral analysis to consider the effect of maximum lag on the spectral density functions. For smoothing the spectral density functions, the smoothing functions suggested by Hanning, Bartlet, Parzen and Tukey (equations 3.23 to 3.26) have been examined in the present study.

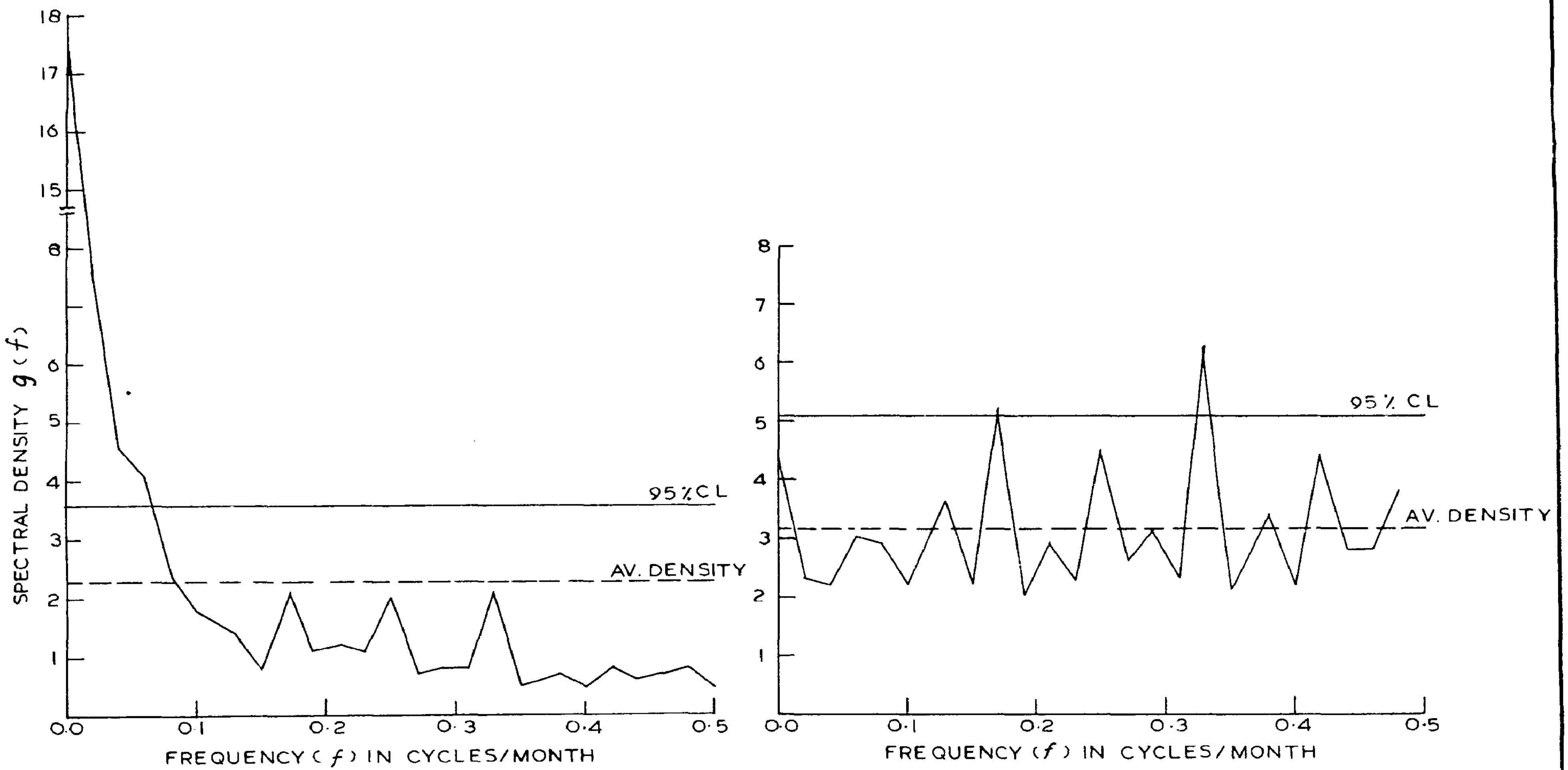


(a) OBSERVED SEQUENCE

(b) 12-MONTH PERIOD REMOVED

MAX. LAGS = 48; HANNING SMOOTHING FN. USED

FIG. 4 (i) VARIANCE SPECTRUM ANALYSIS-RUNOFF SERIES
(LOG. TRANSFORMED)



(c) 12,6-MONTHS PERIOD REMOVED

(d) RANDOM COMPONENT WITH THIRD ORDER AR. REMOVED FROM (C)

MAX. LAGS = 48; HANNING SMOOTHING FN. USED

FIG. 4(ii) VARIANCE SPECTRUM ANALYSIS-RUNOFF SERIES (LOG. TRANSFORMED)

TABLE NO. [illegible]

VARIANCE SPECTRUM ANALYSIS OF [illegible] SEQUENCE-OBSERVED SERIES

Frequency (f)	Smoothed Spectrum Density Function g(f): [illegible] n=96	[illegible] n=48	[illegible] n=96	[illegible] n=48	[illegible] n=96	[illegible] n=48	[illegible] n=96	[illegible] n=48	Remarks period (1/f)
0.00	4.93	3.91	4.46	3.93	4.61	3.20	4.91	4.12	
0.021	0.92	1.94	1.07	1.51	1.16	1.09	0.97	1.94	
0.042	1.24	1.83	1.29	1.70	1.99	1.09	1.30	1.80	
0.063	0.76	1.45	0.91	1.05	1.02	2.42	0.83	1.41	
0.083	29.95	12.71	29.20	12.76	18.24	9.70	29.20	10.08	12 months
0.104	0.72	0.94	0.80	1.00	0.81	2.42	0.77	0.98	
0.125	1.34	1.50	1.55	1.46	1.42	1.40	1.49	1.59	
0.146	0.49	0.95	0.62	0.67	0.64	1.71	0.56	0.99	
0.166	11.12	6.15	10.90	6.20	8.69	4.68	11.90	6.69	6 months
0.188	0.83	0.05	0.07	0.09	0.84	1.50	0.98	0.91	
0.208	1.19	1.15	1.17	1.14	1.10	1.06	1.27	1.22	
0.229	0.25	0.57	0.53	0.34	0.35	1.12	0.32	0.59	
0.250	7.24	4.18	7.11	4.15	3.74	3.29	7.78	4.46	4 months
0.271	0.90	0.73	0.94	0.84	0.85	1.13	0.59	0.87	
0.292	0.96	0.83	0.94	0.89	0.93	0.08	1.05	0.96	
0.313	0.16	0.51	0.26	0.47	0.23	0.91	0.21	0.59	
0.333	5.16	3.08	5.05	3.03	4.17	2.49	5.58	3.33	3 months
0.354	0.81	0.64	0.65	0.69	0.79	0.92	0.90	0.78	
0.375	1.07	0.85	1.06	0.87	0.96	0.70	1.15	0.93	
0.396	0.25	0.62	0.37	0.57	0.41	0.83	0.35	0.89	
0.417	3.12	2.14	3.01	2.11	2.67	1.76	3.30	2.50	2.4 months
0.437	0.33	0.46	0.35	0.38	0.40	0.63	0.39	0.94	
0.458	1.29	0.83	1.21	0.91	1.07	0.70	1.32	0.98	
0.479	0.40	0.50	0.48	0.55	0.45	0.75	0.47	0.69	
0.500	2.25	1.79	2.72	1.70	2.05	1.46	2.99	1.66	2 months
Average g(f)	2.86	2.03	2.05	2.09	2.41	2.14	3.03	2.16	
Confidence limit	5.89	3.28	3.29	3.22	3.82	2.47	5.65	3.42	
Significant harmonics	12, 6, 4	12, 6, 4	12, 6, 4	12, 6, 4	12, 6, 4, 3	12, 6, 4	12, 6, 4	12, 6, 4	

As an example, smoothed spectrum density functions obtained from the above four smoothing techniques for various frequencies ($f = K/2m\Delta t$, where K = lag number, m = maximum number of lags, Δt = one month) pertaining to the untransformed monthly runoff sequence is given in table no. 6 . Similar analysis was conducted for log. transformed data also, at different stages when each of the six harmonics is removed, in turn.

The 95 pct. confidence limit of r_k values in the correlogram have been computed using equation (3.20) and were used with the respective correlograms, in order to identify the significant harmonics.

The equivalent degrees of freedom (EDF) depends upon the spectral window used. For a maximum equal to 48 lags the EDF, when Hamming, Bartlet or Tukey windows are used works out to 19, and when Parzen window is used it is 30 (equations 3.27 to 3.29). The corresponding upper tolerance limit at 95 pct. significance level works out to (1.586 x Average variance density) and, (1.152 x Average variance density) respectively from equation(3.30). For a maximum lag equal to 96, the corresponding values are 9.5 and (1.092 x Average variance density) when Hamming, Bartlet or Tukey windows are used; and 19 and (1.986 x Average variance density) when Parzen window is used.

In the spectrum of logarithmically transformed observed runoff sequence, a sharp peak at 0.083 cycles/month was observed (figure 4) indicating that 12 - month cycle was significant. All the four smoothing techniques confirmed that 12 - month cycle as significant as is evident from table No. 7. After the removal of

[illegible] OF SIGNIFICANT [illegible]

Particulars	Significant periods: Running smoothing n=35	Running smoothing n=48	Partial smoothing n=35	Partial smoothing n=48	[illegible] smoothing n=35	[illegible] smoothing n=48	[illegible] smoothing n=35	[illegible] smoothing n=48	Remarks
I. [illegible]									
Original series	12	12	12	12	12	12	12	12	
12 month period removed.	6	6	6	6	6	6	6	6	
II. [illegible]									
Original series	12,6,4	12,6,4	12,6,4	12,6,4	12,6,4 3	12,6,4	12,6,4	12,6,4	
12 month period removed.	6,4,3	6,4,3	6,4,3	6,4,3	6,4,3	6,4,3	6,4,3	6,4,3	
12,6 month period removed.	12,6,4,3 2.4	12,4,3	4,3	12,4,3,2.8	12,4,3,2.4	12,4,3,2.4	14,3,2.4	12,8,3 2.8	
12,6,4 month period removed.	12,3,2.4	12,3,2.4	12,3,2.4	12,3,2.4	12,3,2.8	12,3,2.4	12,3,2.4	12,3,2.4	
12,6,4,3 month period removed.	12,2.4	12,2.4	12,2.4	12,2.4	12,2.4	12,2.4	12,2.4	12,2.4	
12,6,4,3,2.4 month period removed.	12	12	12	12	12	12	12	12	
12,6,4,3,2.4,2 month period removed.	12	12	12	12	12	12	12	12	
III. [illegible]									
Original series 12 month period removed.	12,3 6,4,3, 2.4	12,3 6,4,3, 2.4	12,3 6,4,3, 2.4	12,3 6,4,3, 2.4	12,3 6,4,3, 2.4	12,3 6,4,3, 2.4	12,3 6,4,3, 2.4	12,3 6,4,3, 2.4	
12,6 month period removed.	4,3,2.4	4,3,2.4	4,3, 2.4	4,3, 2.4	4,3, 2.4	4,3, 2.4	4,3, 2.4	4,3,2.4	
12,6,4 month period removed.	3,2.4	3,2.4	3,2.4	3,2.4	3,2.4	3,2.4	3,2.4	3,2.4	
12,6,4,3 month period removed.	2.4	2.4	2.4	2.4	2.4	2.4	2.4	2.4	

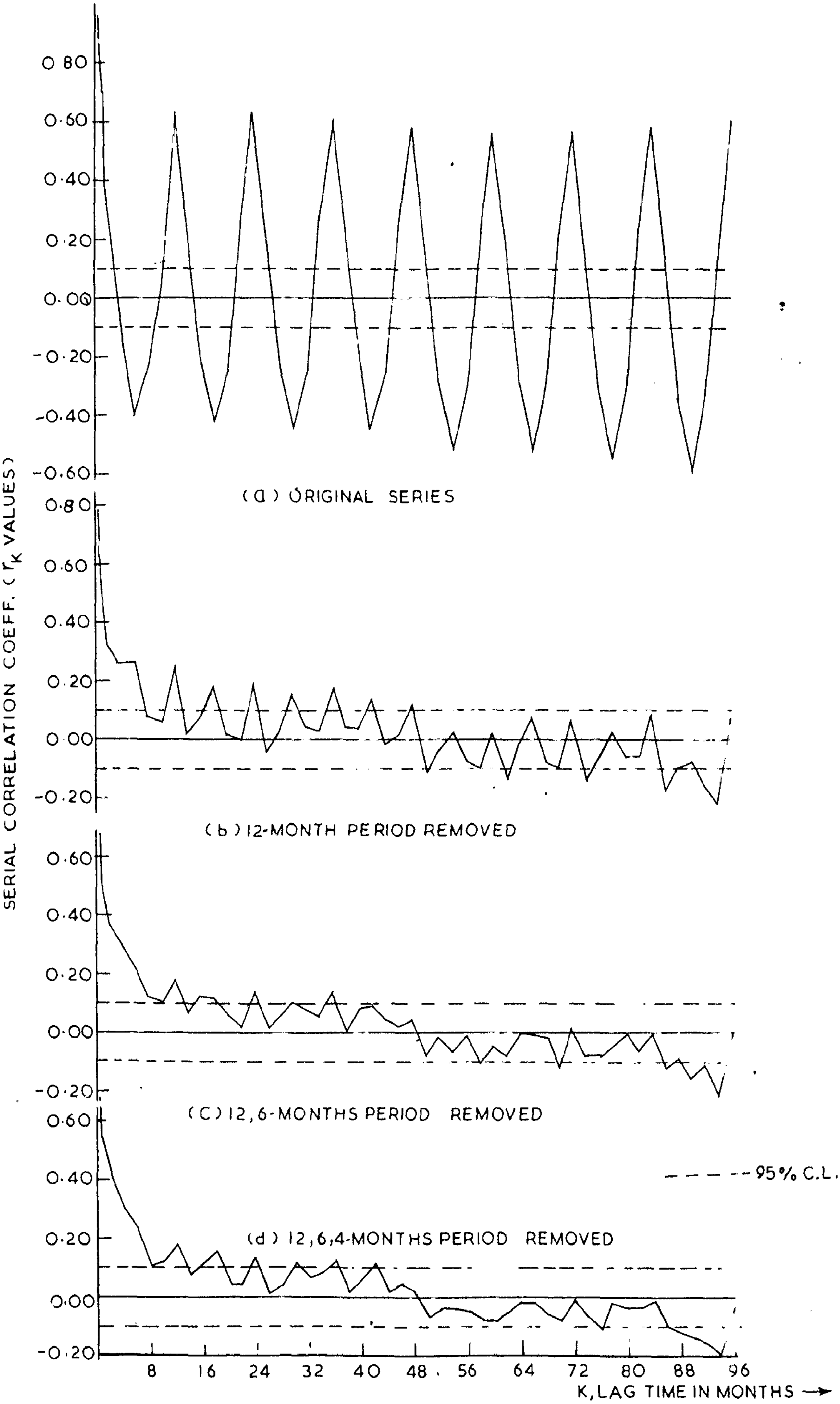


FIG. 3(i) CORRELOGRAM ANALYSIS (RUNOFF SERIES-LOG. TRANSFORMED)

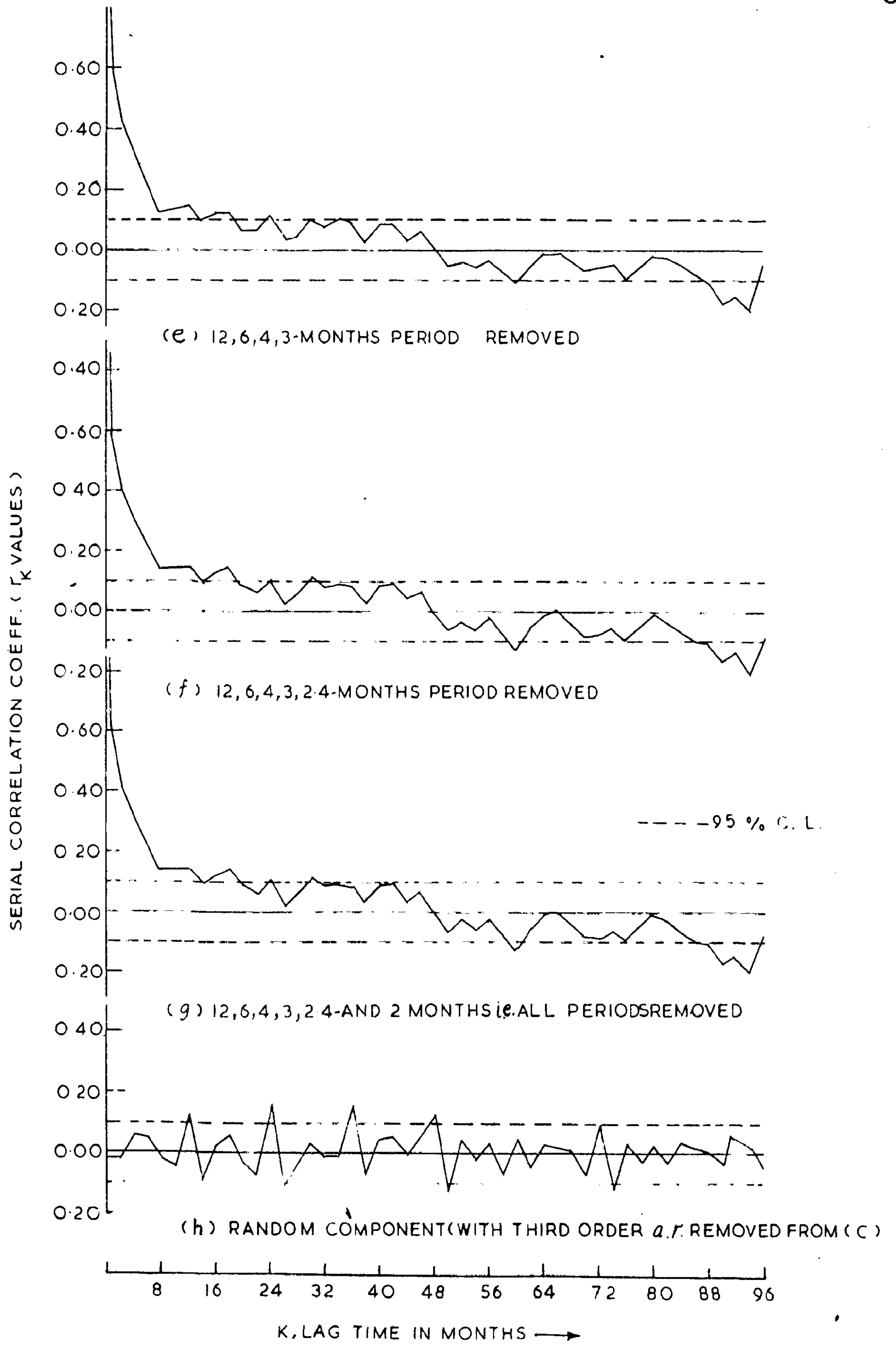


FIG. 3(ii) CORRELOGRAM ANALYSIS — (RUNOFF SERIES LOG. TRANSFORMED) periodic component of 12 - month cycle from the K_0 series, its 6 - month subharmonic became significant as shown in the spectrum plotted in figure 4 . However after the removal of the cyclical component comprising of 12 - month cycle and its 6 - month subharmonic from K_0 series , no significant harmonic was observed in the residual series at the further stages of removal of 4, 3, 2.4 and 2 - month harmonics (Table no.7B)

The correlogram results are however not that definite. The correlogram of original log transformed series (figure 9) indicates presence of 12 and 6 - months period. After removal of 12-month period the correlogram for remaining series indicates presence of 6 - month period as significant. However after removal of both 12 and 6 - month period the correlogram of the remaining series when each of 4, 3, 2.4 and 2 - months period are removed in turn do not reveal a definite information about further periods being significant. Since the spectrum analysis is less subjected to sampling variations in comparison to correlograms, only 12 and 6 - months period were considered as significant, as given by spectrum analysis.

It was observed (table 8) that the 12 - month cycle and its 6 - month subharmonic explains about 96 pct. of the total variance of K_0 series. About 93 pct. of the variance is explained by the 12 - month cycle alone. The variance explained by each harmonic is computed from equation (3.76) (computer program No.3).

The Fourier coefficients A_j and B_j using equations (3.6 and 3.7) were obtained as - 1.1565, - 2.3170 for 12 - month

T A B L E No. 8

VARIANCE EXPLAINED BY EACH OF THE HARMONICS

Sl.No.	Particulars	Fourier Coeffs. A_j	Fourier Coeffs. B_j	Explained variance, $(A_j^2+B_j^2)/2$	Pct. of explained variance.	Remarks
1.	RUNOFF SEQUENCE (LOG. TRANSFORMED)					
	(a) Observed sequence			6.344		
	(b) 12 month period	-1.157	-2.317	3.353	52.84	
	(c) 6 month period	0.032	0.621	0.194	3.06	
	(d) 4 month period	-0.071	0.333	0.058	0.91	
	(e) 3 month period	-0.365	0.191	0.085	1.32	
	(f) 2.4 month period	0.280	0.156	0.033	0.52	
	(g) 2 month period	-0.122	0.000	0.007	0.11	
	Total variance explained by significant harmonics i.e. 12 and 6 month periods.			3.547	55.90	
	Stochastic component			2.797	44.10	
2.	RUNOFF SEQUENCE					
	(a) Observed sequence			1107930		
	(b) 12 month period	-470.032	-522.079	230677	21.0	
	(c) 6 month period	67.596	471.619	113493	9.6	
	(d) 4 month period	129.349	-331.551	60460	5.4	
	(e) 3 month period	-240.603	175.803	43201	3.7	
	(f) 2.4 month period	211.603	- 16.793	22520	1.0	
	(g) 2 month period	-155.719	- 0.039	012124	01.0	
	Total variance explained by significant harmonics i.e. 12,6,4,3 and 2.4 month period.			502499	42.3	
	Stochastic components			605439	57.7	
3.	RAINFALL SEQUENCE					
	(a) Observed sequence			27.69		
	(b) 12 month period	-4.370	-5.207	14.98	53.20	
	(c) 6 month period	1.047	1.001	1.04	4.13	
	(d) 4 month period	-0.325	-1.469	1.13	4.08	
	(e) 3 month period	-1.304	1.374	1.70	6.49	
	(f) 2.4 month period	1.164	-0.103	0.68	1.93	
	(g) 2 month period	-0.521	-0.033	0.13	0.47	
	Total variance explained by all significant harmonics.				69.44	
	Stochastic component			8.01	30.66	

harmonic and, 0.0919 and 0.0212 for 6 - month harmonic respectively. These can be substituted in equation (9.9) to give periodic component in the mean i.e. for log. transformed series.

$$\mu_\tau = \mu_x + \sum_{j=1}^{2} [A_j \cos(2\pi kj/12) + B_j \sin(2\pi kj/12)]$$

where μ_x is the mean of trend free series. Since series mean of the original log. transformed series (9.9623) was removed while analysing trend component, by considering it as a trend with zero slope, hence μ_x in the above equation (7.9) will be zero.

Equations (9.13 and 9.14) for obtaining the Fourier coefficients facilitate computing A_j and B_j from 12 values of monthly means only, instead of all the values of X_t series. Computer program no. 4 was run to compute A_j and B_j values from the monthly means of X_t series alone, and values of A_j and B_j obtained were identical with the corresponding values calculated by using equations (9.6 and 9.7).

Computer program no. 4 was run for testing the significance of harmonics by approximate method of empirical approach described in section 9.5.2 of the chapter 9. The empirical approach for testing the significance of harmonics using equations (9.16 to 9.18) indicate that all the six harmonics are significant, as explained below:

ΔD_j the part of variance in the sequence X_t explained by the j-th harmonic is computed from equation (9.16).

The value of P from equation (9.17) is given by

$$P = \sum_{j=1}^{m} \frac{Var\ h_j}{Var\ x_t} = \frac{3.728}{6.3438} = 0.587$$

(m = maximum number of harmonics) in X_t series)

$$P_{min} = 0.099 \times \sqrt{\frac{12}{1 \times 98}} = 0.0186$$

$$P_{max} = 1.0 - 0.0186 = 0.9814$$

Since $P_{min} < P < P_{max}$, all the eight harmonics are to be inferred significant from this approach. However in the present study the results from correlogram and spectrum analysis, which gave only 12 and 6 - month periods as significant, have been considered for identifying the significant harmonics in the periodic component and only these two periods were removed and Y_t series was obtained from the equation

$$Y_t = (X_t - T_t - P_t) / \sigma_x$$

where σ_x is the standard deviation of the trend removed series ($\sigma_x = 2.9183$) Both Y_t and $y_{p,\tau}$ represent same series in single and double subscript respectively.

For the parametric method of separation of periodic and stochastic components a further transformation of the series is necessary, as the Y_t series has a mean of 0.0001 and variance of 0.4412 (standard deviation = 0.6642), as against their expected values of 0 and 1 respectively.

Using equation (3.19) the ϵ_t series was obtained as

$$\epsilon_t = (Y_t - 0.0001)/0.6642 \quad \ldots\ldots\ldots\ldots \quad (.7.5)$$

It is observed from the results of computer program No. 5 that the autocorrelation coefficients of Y_t and ϵ_t series, were not changed by this transformation and therefore the correlogram and spectrum of ϵ_t series remain similar to that of Y_t series (figure 3 and 4).

The difference between the variance of the time series X_t and the total of explained variance of the significant harmonics is attributed to the stochastic component of the time series, which works out to 44 pct. for this log. transformed series (table no. 8).

The correlogram (figure no. 9) and the spectrum for the stochastic component obtained after removal of periodic component show that, this component is a highly dependent stochastic component approximately of an auto-regressive (a.r.) model.

To identify the order of linear autoregressive model, the determination coefficient approach was adopted. Computer program 5 was run to find the values of determination coefficients R_1^2, R_2^2 and R_3^2 using equations (3.42) to (3.44) respectively and these were found to be 0.2830, 0.2945 and 0.3136 respectively using the autocorrelation coefficients $r_1 = 0.5310$, $r_2 = 0.3796$, and $r_3 = 0.3575$ calculated for ϵ_t (i.e. X_t) series.

In this approach it is assumed that the monthly serial correlation coefficients $r_{k,\tau}$ were not significantly different from a constant independent of τ for a given k, where τ represents the calendar month and k is the lag. This assumption is equivalent to assuming that there was no periodicity in auto-correlation coefficients.

Since $R_2^2 - R_1^2 = 0.0115 > 0.01$ and $R_3^2 - R_2^2 = 0.0193 > 0.01$, using equation (3.47), third order auto regressive (a.r.) model was adopted giving stochastic component as below:

$$\epsilon_t = a_1 \epsilon_{t-1} + a_2 \epsilon_{t-2} + a_3 \epsilon_{t-3} + R_t \quad --- (7.6)$$

T A B L E NO. 9

CHI-SQUARE TEST FOR GOODNESS-OF-FIT

(Independent Stochastic Component-Log. Transformation)

		O_j	O_j^2			O_j	O_j^2
		17	289			11	121
K_1	-2.1217			K_{16}	-0.6217		
		0				6	36
K_2	-1.8517		36	K_{17}	-0.5717		
		6				5	
K_3	-1.6417			K_{18}	-0.5217		49
		9	81			4	
K_4	-1.5017			K_{19}	-0.4817		
		5	25			13	169
K_5	-1.3817			K_{20}	-0.4317		
		4				2	
K_6	-1.2817		49	K_{21}	-0.3817		144
		3				10	
K_7	-1.1917			K_{22}	-0.3417		
		4				13	169
K_8	-1.1217		49	K_{23}	-0.3017		
		3				6	36
K_9	-1.0417			K_{24}	-0.2517		
		6	36			9	81
K_{10}	-0.9717			K_{25}	-0.2117		
		6	36			12	144
K_{11}	-0.9017			K_{26}	-0.1719		
		5				9	81
K_{12}	-0.8417		64	K_{27}	-0.1317		
		3				15	225
K_{13}	-0.7817			K_{28}	-0.0817		
		6	36			14	196
K_{14}	-0.7317			K_{29}	-0.0410		
		7	49			10	100
K_{15}	-0.6717			K_{30}	+0.0017		

		O_j	O_j^2
		6	36
X_{31}	0.0417		
		7	49
X_{32}	0.0817		
		9	81
X_{33}	0.0517		
		13	169
X_{34}	0.1717		
		8	64
X_{35}	0.2117		
		9	81
X_{36}	0.2517		
		10	100
X_{37}	0.3017		
		10	100
X_{38}	0.3417		
		13	169
X_{39}	0.4017		
		4	
X_{40}	0.4317		169
		9	
X_{41}	0.4817		
		6	36
X_{42}	0.5217		
		11	121
X_{43}	0.5717		
		9	81
X_{44}	0.6217		
		12	144
X_{45}	0.6717		
		7	49
X_{46}	0.7317		

		O_j	O_j^2
		7	49
X_{47}	0.7817		
		10	100
X_{48}	0.8417		
		10	100
X_{49}	0.9017		
		9	81
X_{50}	0.9717		
		4	
X_{51}	1.0417		81
		5	
X_{52}	1.1217		
		11	121
X_{53}	1.1917		
		7	49
X_{54}	1.2817		
		6	
X_{55}	1.3817		100
		4	
X_{56}	1.5017		
		6	36
X_{57}	1.6417		
		5	25
X_{58}	1.8317		
		3	
X_{59}	2.1217		64
		5	

$\Sigma O_j^2 = 4556$

X^2 calculated $= (50/456) \times 4556 - 456 = 499.5 - 4560 = 43.5$

$X^2_{0.05}$ for 47 d.f. $= 65.00$

$\therefore X^2$ calculated $< X^2$ critical

Hence Normal distribution fits.

where R_t (zt) is the independent stochastic component (random component)
The values of a_1, a_2, a_3 were calculated using equations (3.39) to (3.40) with r_1 = 0.9910 and r_2 = 0.9796 and found to be 0.4495, 0.0497 and 0.1696 respectively.

The dependent stochastic component was removed from the series by fitting a third order a.r. model using equation 3.41 and random component $\xi_{p,\tau}$ or ξ_t was obtained, mean of which was 0.0017 and standard deviation was 1.00 from the results of computer program No. 9

The independent stochastic component ξ_t was obtained by removing the dependence structure from the standardised stochastic component, using the a.r. model of a particular order. The correlogram (fig. 3) and the spectrum (fig. 4) for this component together with their 95 pct. C.L. were plotted. Both these indicate that the random variables constituting this component can approximately be accepted as independent standardised (0,1) random variables of second order stationarity.

The ξ_t series was tested for best fit distribution by using chi-square and Kolmogorov-Smirnov tests for goodness of fit as explained in chapter 3.

For the chi-square test the ξ_t series was arranged in descending order and arranged into 60 class intervals (97) on the basis of class intervals of equal probabilities (20) and using cumulative standardised normal distribution functions. From the calculations shown in table No. 9 the Chi-square was obtained as 31.00 (using equation (3.74)) where as the chi-square value tabulated for (50-3) = 47 degrees of freedom at 95 pct. confidence level is 65.00 and hence the hypothesis is true that the normal distribution fits the random component i.e. S_t series.

The K-S statistic is calculated using equation (3.76) with the aid of computer program no. 7. As the hypothesised cumulative distribution function $F_H(x_i)$ of equation (3.76) is the normal distribution function, equation (3.70) is used to obtain the value of $F_H(x_i)$, and 'D' was obtained as 0.042. For a confidence level of 95 pct. and a sample size of 456, the critical value of K-S statistic is $1.36/\sqrt{456} = 0.064$ using equation (3.79). Therefore S_t series can be treated as normally distributed.

The Model: Having thus identified all the additive components of the time series model, the model is described as below:

$X_t = T_t$ (trend component) + P_t (Periodic component)
+ S_t (Stochastic component)

Where t is the number of month.

The above components are modelled as

T_t = mean of the original series = 9.9626

$$P_t = A_1 \cos(2\pi t/12) + B_1 \sin(2\pi t/12) + A_2 \cos(4\pi t/12) + B_2 \sin(4\pi t/12)$$

where $A_1 = -1.1509$, $B_1 = -2.3170$,
$A_2 = 0.0319$, $B_2 = 0.6212$

$$\epsilon_t = S_t \times \sqrt{1-(a_1^2+a_2^2+a_3^2+2a_1a_2r_1+2a_1a_3r_2+2a_2a_3r_2)} + a_1\epsilon_{t-1} + a_2\epsilon_{t-2} + a_3\epsilon_{t-3}$$

where $a_1 = 0.4459$, $a_2 = 0.0497$, $a_3 = 0.1694$,

$b_1 = 0.3310$, $b_2 = 0.3796$

R_t = the random component, described as

$R_t = \mu + \sigma(Zt)$ in which $\mu = 0.0017$, $\sigma = 1.0000$

and Zt is the normally distributed random numbers with mean zero and variance unity.

From equation (5.15)

$X_t = AV + STD\ (\epsilon_t)$ and $S_t = X_t \times \sigma_B$

where $AV = 0.0001$, $STD = 0.6642$ and $\sigma_B = 2.5103$

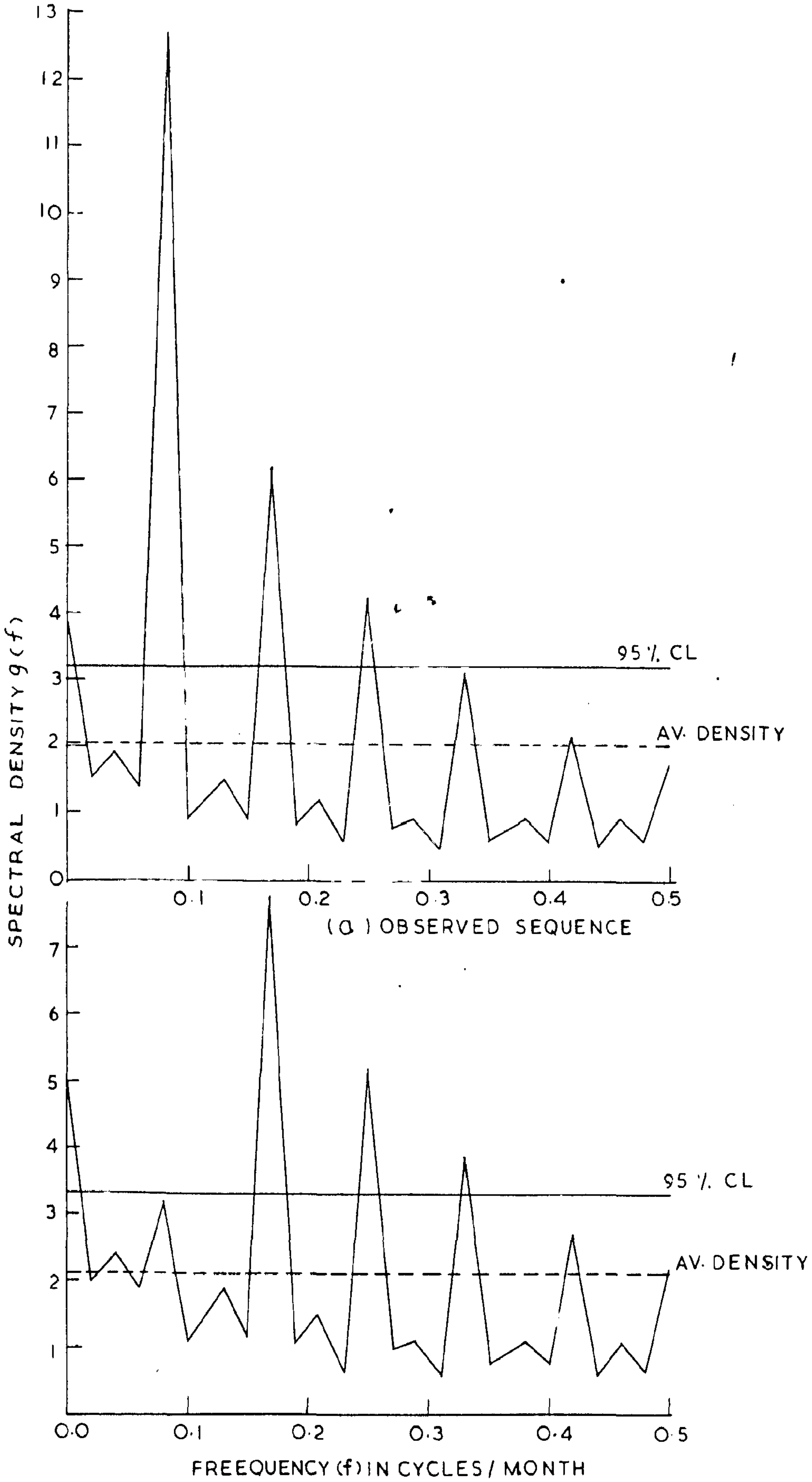
Since the analysis is performed on the logarithmic values of the runoff sequence

$$X_t = \exp\ (S_t + P_t + R_t) \quad \ldots\ldots\ldots\ldots\ldots \quad (7.7)$$

7.5.2. Classical Approach of Analysis of Monthly Runoff Sequences (involving no transformation of the data)

For the detection of periodic movement inside the runoff series and to approximate it by Fourier analysis, the analysis of untransformed data was performed on the same lines as that of logarithmically transformed data.

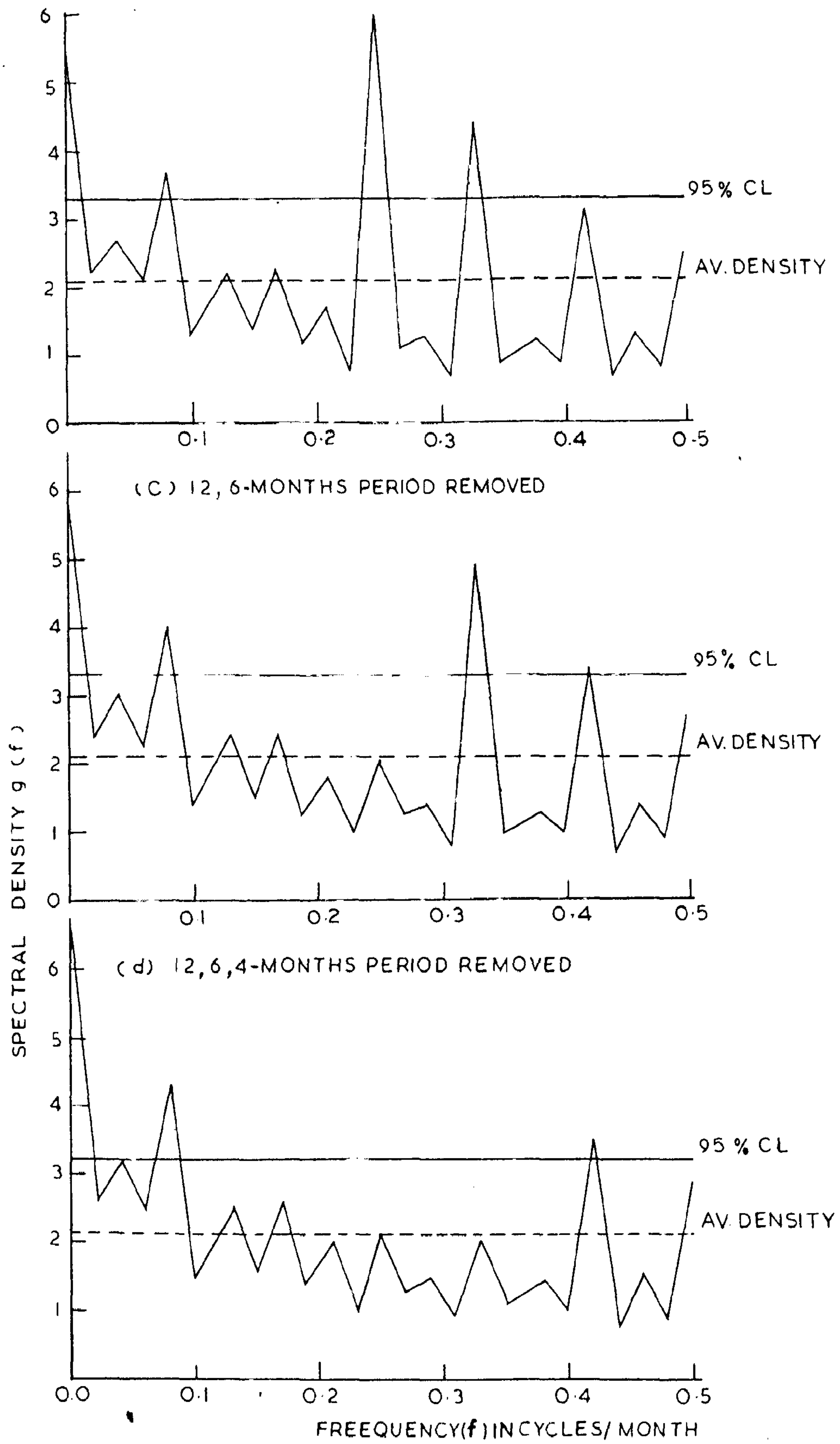
Computer program no. 3 was run, to obtain Fourier coefficients from the periodogram analysis and also for the correlogram and spectrum analysis of runoff sequence and the series with each of the six harmonics, corresponding to 12,6,4,3,2.4 and 2-months period in turn.



(b) 12-MONTH PERIOD REMOVED

MAX. LAGS = 48; HANNING SMOOTHING FN. USED

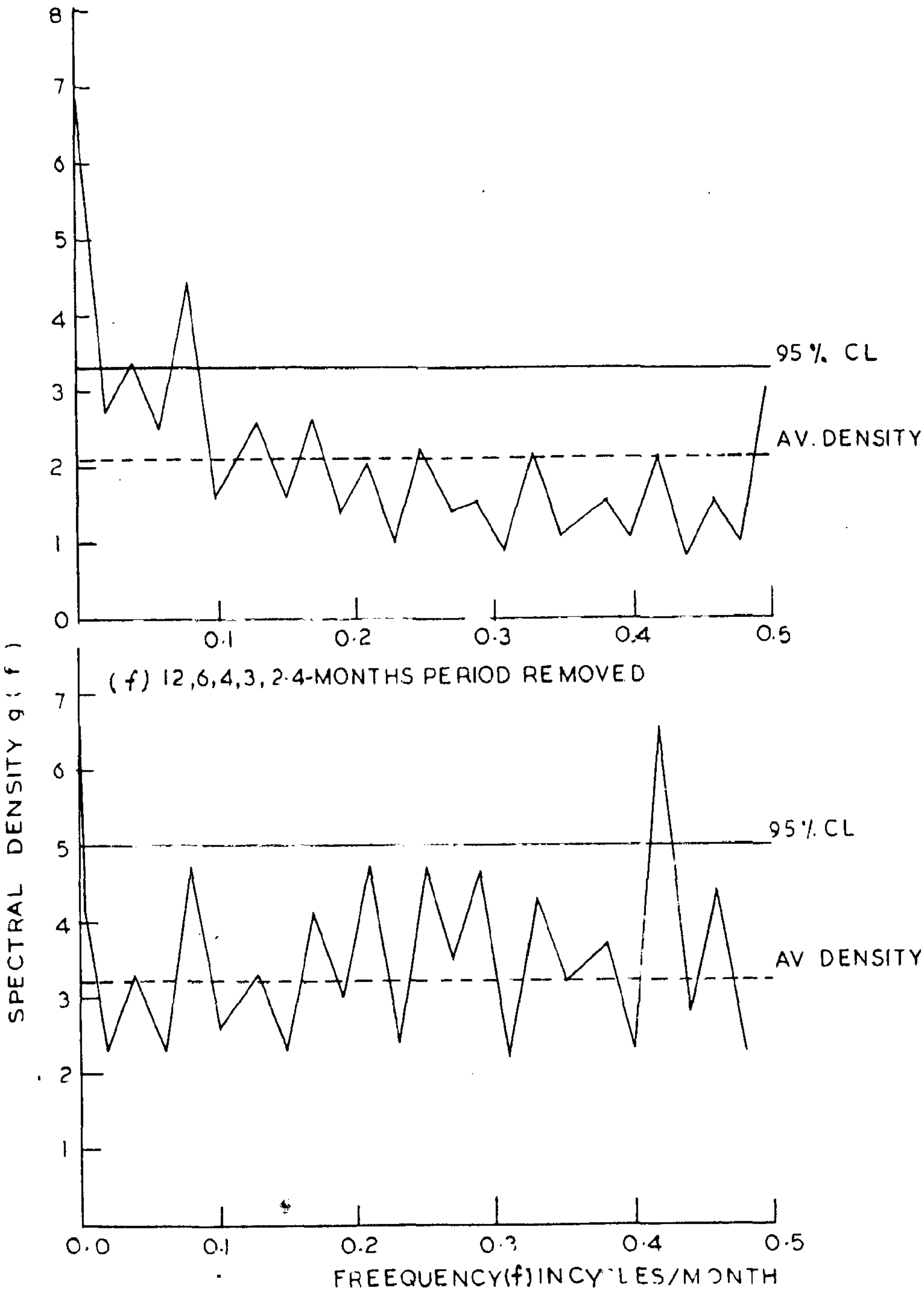
FIG.6(i) VARIANCE SPECTRUM ANALYSIS-RUNOFF SERIES



(e) 12, 6, 4, 3-MONTHS PERIOD REMOVED

MAX. LAGS = 48; HANNING SMOOTHING FN. USED

FIG. 6(ii) VARIANCE SPECTRUM ANALYSIS-RUNOFF SERIES



(g) RANDOM COMPONENT WITH 2ND ORDER A.R. REMOVED FROM (f)

MAX LAGS = 48; HANNING SMOOTHING FN. USED

FIG.6(iii) VARIANCE SPECTRUM ANALYSIS-RUNOFF SERIES

A maximum lag 'm' equal to 40 and also 96 were chosen for the correlogram and spectral analysis to consider the effect of maximum lag on the spectral density functions. As explained for the log. transformed data analysis, for the untransformed run off data also, the Hanning, Bartlet, Parzen and Tukey windows for smoothing the spectral density functions have been used. As an example the smoothed spectrum density functions obtained from the above four smoothing techniques, for the observed runoff sequence for the various frequencies are given in table no. 6.

The significance levels for correlogram and spectrums are drawn as explained for the similar analysis of log.transformed run off data.

From the spectrum analysis of observed sequence (as shown in figure 6 and given in(table No. 6 and 7) 12,6,4-month period were found significant, for the Hanning, Bartlet, and Tukey windows, where as for Parzen window, in addition to 12,6,4-months period, 3 - month period was also found to be significant. However in general, result from spectrum analysis of the series with each of 12,6,4,3,2.4, and 2 - month period removed in turn indicated almost identical findings for all the four . smoothing functions (table 7). The spectrum of 12 - month period removed (fig. 6) indicated the significance of 6,4,3-months period. After the removal of 12,6,4-months period, still 12,3, 2.4-months period was significant. 2.4 - month period remained significant in addition to 12-month period, after the removal of 12,6,4,3-month period (fig.6, table 7).However after the removal &

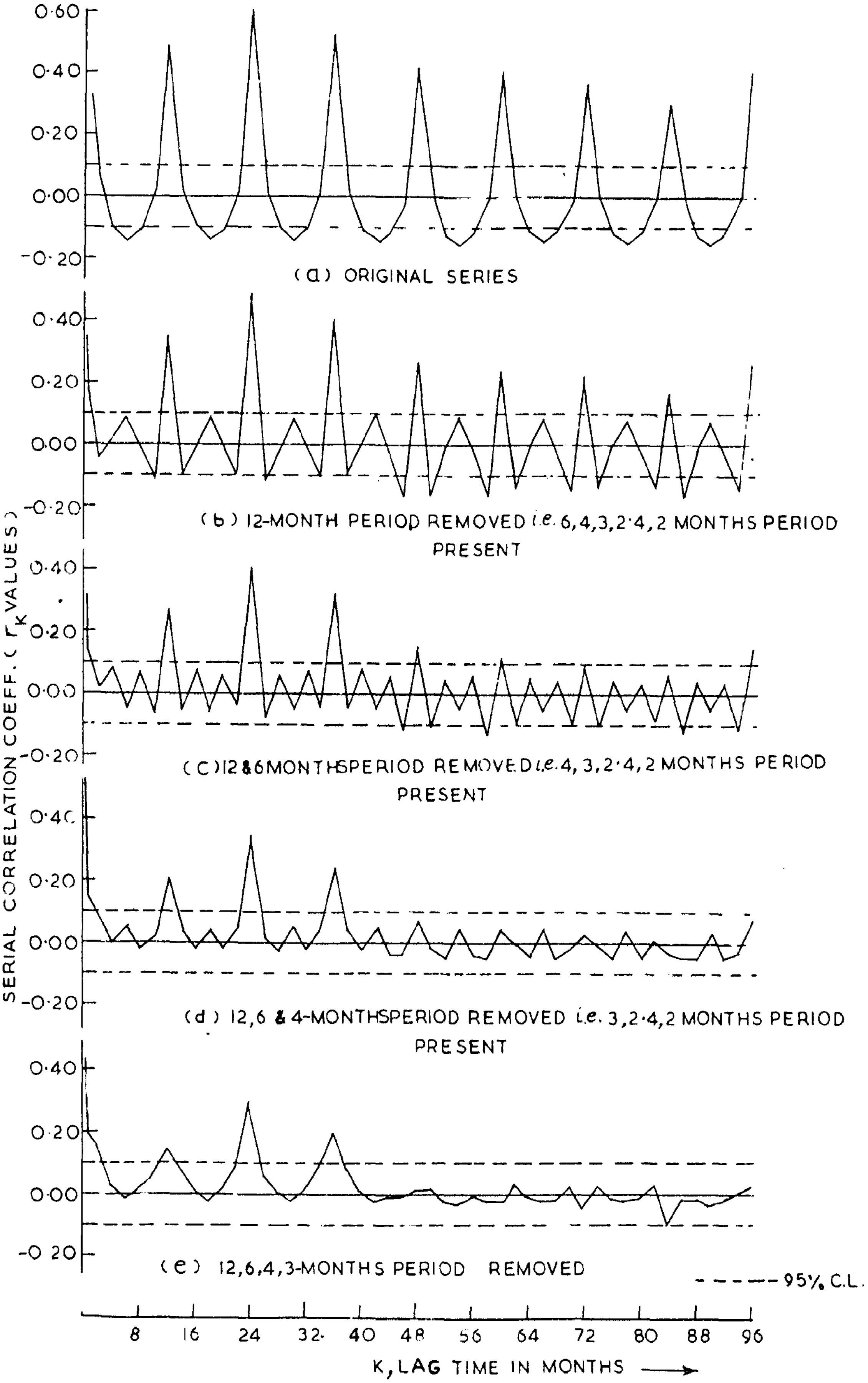


FIG. 5(i): CORRELOGRAM ANALYSIS (RUNOFF SERIES – NO TRANSFORMATION)

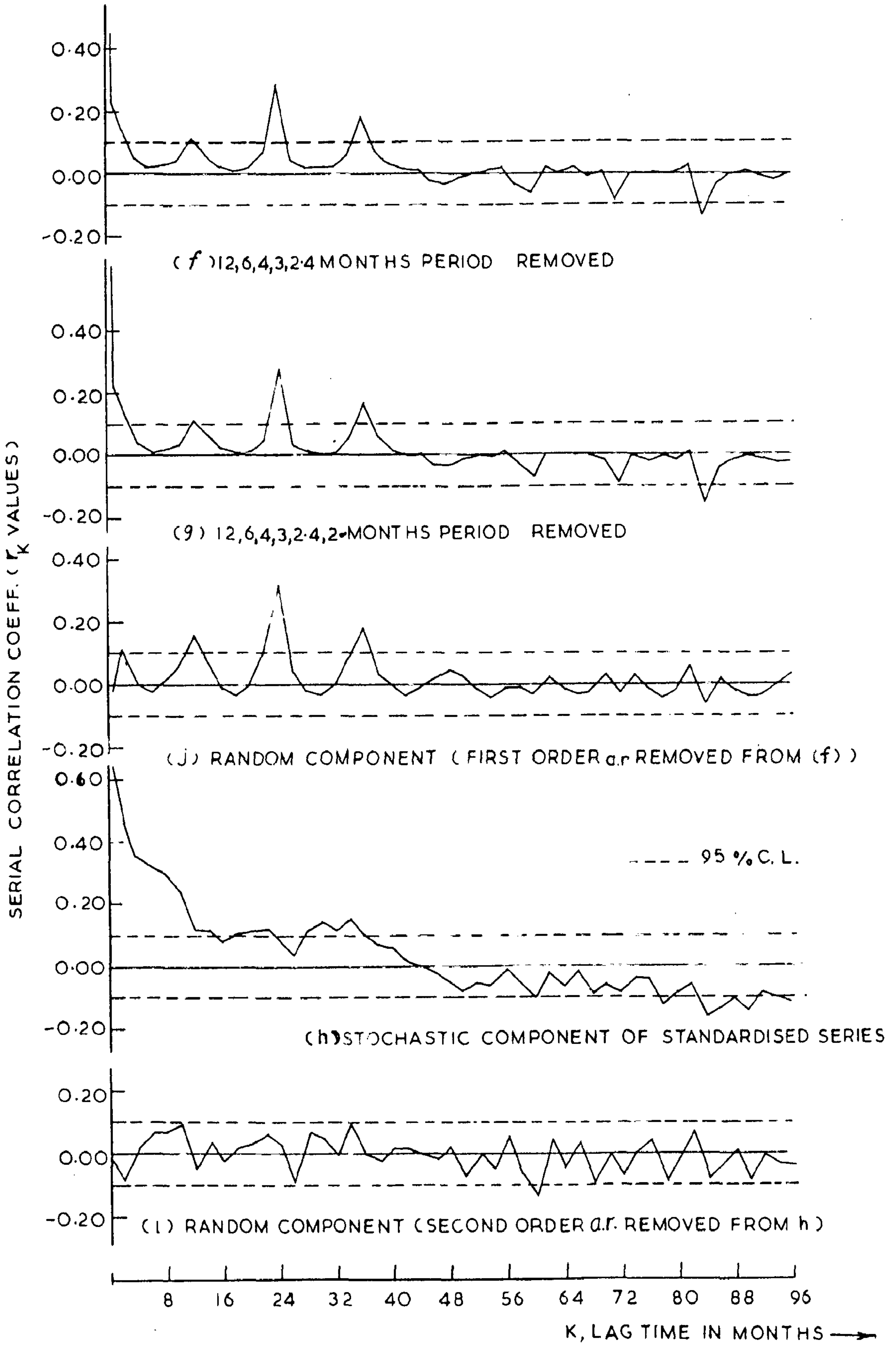


FIG.5(ii) CORRELOGRAM ANALYSIS-RUNOFF SERIES(NOTRANSFORMATION)

12,6,4,3,2.4 - months period , 12 - month cycle was still significant. The presence of 12 - month cycle in the residual series may be attributed to chance fluctuations or the necessity of fitting periodic functions to the standard deviation also. However this requires further investigation.

The correlogram (fig. 5) of observed sequence indicated the significance of 12,6,4 - months period, in conformity with the findings of spectrum (fig. 6) of observed sequence. However after the removal 12 month - period 6,4 - month periods become significant where as 3 - months period is not clearly detectable. The significance of the periods of 3,2.4 and 2 - months cannot be definitely said in the correlograms of figure 5, unlike the spectrum at the corresponding stage of figure 6. The 12 month-period remained significant in the correlogram of stochastic component also.

It was observed from table No. 8, that 12,6,4,3-month harmonics explains about 42 pct. of the variance of H_0 series, out of which about 22 pct. is explained by the 12- month period alone. The variance explained by each harmonic is computed from equation (9.9) and computer program No.5.

Fourier coefficients A_j and B_j were computed using equations (9.6) and (9.7) which were found to be -470.0796 - -522.6721 for 12 - month period; 67.5961, 471.0594 for 6-month period, 129.3403, -331.5503 for 4-month period -240.4093, 175.8023 for 3-month period and 211.5959,-16.7994 for 2.4-month period respectively. These can be substituted in equation 9.5

to give the periodic component in mean i.e. μ_T.

$$\mu_\tau = \mu_x + \sum_{j=1}^{5} \left[A_j \cos(2\pi Kj/12) + B_j \sin(2\pi Kj/12)\right]$$

Where μ_x is the mean of trend removed series.

Identical A_j and B_j values were found, when calculated from the 12 values of monthly means of observed sequence using equations (3.13) and (3.14) and computer program 4.

Computer program 4 was used for testing the significance of harmonics by approximate method of empirical approach from equations (3.16) to (3.18) and all the six harmonics were found to be significant as explained below:

P_j, the part of variance in the sequence X_t explained by j-th harmonic is computed from equation (3.16). The value of P from equation (3.17) is given by

$$P = \sum_{j=1}^{n} \frac{Var\, h_j}{Var\, X_t} = \frac{502/89}{110/993} = 0.626$$

Where n = max. no. of harmonics.

$$P_{min} = 0.093 \times \sqrt{12/1x50} = 0.0186$$

$$P_{max} = 1.0 - 0.0186 = 0.9814$$

Since $P_{min} < P < P_{max}$ all the six harmonics were to be inferred significant from this approach. However the significant harmonics in the periodic component as identified by the correlogram and spectrum analysis were adopted for further analysis. Y_t series was obtained from the equation

$$Y_t = \frac{X_t - T_t - P_t}{\sigma_x}$$

where σ_T = standard deviation of trend free series (σ_T =1009.926)

Since Y_t series obtained after the removal of significant harmonics had the mean = 0.0022 and variance = 0.5969(standard deviation = 0.7720) against their expected values of 0 and 1 respectively.

From equation (3.15)

$$\epsilon_t = (Y_t - 0.0022) / 0.7720 \quad \ldots\ldots\ldots\ldots \quad (.7.0)$$

The autocorrelation coefficients of Y_t and ϵ_t series were similar as observed from the results of computer program no. 9. Therefore the correlogram and spectrum of Y_t and ϵ_t series (fig. 5 and fig. 6) remained same.

The difference between the variance of time series X_t and the total of explained variance of significant harmonics is attributed to the stochastic component, which works to about 90 pct. (table No. 8).

The correlogram (fig. 5) and spectrum (fig. 6) for the stochastic component obtained after the removal periodic component, show that this component is a dependent stochastic component of an a.r. model.

The coefficient of determination approach was used for the identification of the order of the model.

R_1^2 , R_2^2 , R_3^2 of equations (3.42) to (3.44) were found to be 0.0524, 0.0599, 0.0599 respectively , from

r_1 = 0.2203, r_2 = 0.1366, r_3 = 0.0548 calculated for ϵ_t (or Y_t) series.

Since $R_2^2 - R_1^2 = 0.0079 < 0.01$ and

$R_3^2 - R_2^2 = 0 < 0.02$, as per equation (5.69)

first order linear autoregressive a.r. model was adopted, giving the stochastic component as below

$\epsilon_t = a_1 \epsilon_{t-1} + \xi_t$ . . . .(7.9) where ξ_t is the independent stochastic component.

The dependent stochastic component was removed from the series by fitting a first order a.r. model by using equation no. (5.54) (computer program no. 5) and random component ξ_t with mean = 0 and standard deviation = 1.0109 was obtained.

The correlogram (fig. 5) and spectrum (fig. 6) of the residual series can be almost taken as indicating random component.

The ξ_t series was tested for best fit distribution by chi-square and K-S test for goodness of fit.

The chi-square value (calculations similar to that of ξ_t series of log. transformed data) was found to be 89.00 against critical value of 67.50 for (53-3) = 50 degrees of freedom, at 95 pct. confidence level. In this case also ξ_t series was divided 53 class intervals of equal probability.

As the calculated value of chi-square is greater than admissible critical value, the hypothesis is not true i.e. the normal distribution does not fit the random component ξ_t series.

However the K-S statistic obtained from computer program no. 7 was 0.059 against the critical value of K-S statistic of 0.064 at 95 pct. confidence level.

Though the K-S test indicates that the normal distribution fits the Z_t series, as the hypothesis is not confirmed by chi-square test, log-normal distribution was tried. Since the random component Z_t series contain many negative values, 3-parameter log-normal distribution was considered for testing its fitness to the Z_t series.

The calculation of parameters of LN-3 distribution of equation (3.54) involves interactive procedure requiring greater computation time for solving iterative simultaneous solution of three coupled, non linear equations, and therefore the method of moments approach of equation (3.61) was followed in the present study to calculate the shift parameter δ_0. The method of moments approach is not so accurate as the method of maximum likelihood, but it is more straight-for-ward. The shift parameter or lower boundary location parameter δ_0 was obtained from computer program no. 6 and found to be 11.0555, and it was added to all the values of Z_t series so as to transfer in the negative values in Z_t series to positive values to facilitate logarithmic transformation of Z_t series. The mean and standard deviation of log($Z_t + \delta_0$) were found to be 2.5998 and 0.0712 respectively.

The chi-square test was performed on the log($Z_t + \delta_0$) series by dividing the log($Z_t + \delta_0$) series into 60 class intervals of equal probability (calculations are similar to that shown for the Z_t of log. transformed data). The Chi-square value for (53-3)= 50 degrees of freedom calculated as 63.20 against the tabulated value of 67.50 at 95 pct. confidence level. Therefore the hypothesis is true i.e. the 3-parameter log-normal distribution fits into the Z_t series.

The Model:- Having identified all the additional components of the time series model, the model is described below:

$$X_t = T_t \text{ (Trend component)} + P_t \text{ (periodic component)} + Q_t \text{ (Stochastic component)}$$

where t is the number of month.

The above components are described below:

T_t = original series mean i.e. 441.1621

$$P_t = A_1 \cos(2\pi t/12) + B_1 \sin(2\pi t/12) + A_2 \cos(4\pi t/12) + B_2 \sin(4\pi t/12) + A_3 \cos(6\pi t/12) + B_3 \sin(6\pi t/12) + A_4 \cos(8\pi t/12) + B_4 \sin(8\pi t/12) + A_5 \cos(10\pi t/12) + B_5 \sin(10\pi t/12)$$

Where

$A_1 = -478.0754$, $B_1 = -922.0721$

$A_2 = 67.5941$, $B_2 = 471.6394$

$A_3 = 129.5400$, $B_3 = 531.9953$

$A_4 = -240.4033$, $B_4 = 175.2025$

$A_5 = 211.5999$, $B_5 = -16.7994$

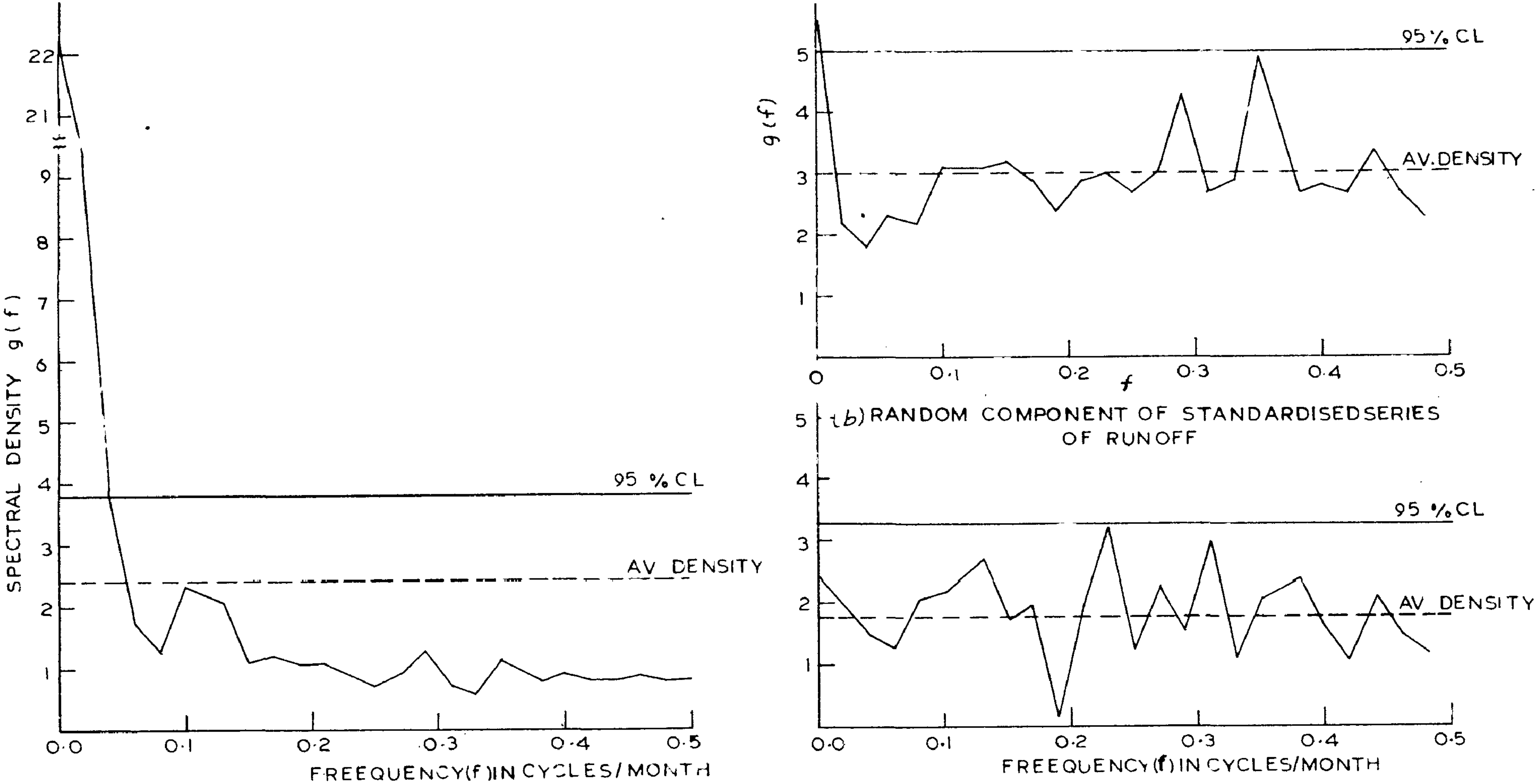
$$\epsilon_t = \left(\xi_t \times \sqrt{1 - r_1^2}\right) + r_1 \epsilon_{t-1}$$

where $r_1 = \rho_1 = 0.2283$

ξ_t the random component is described as

$$\xi_t = \xi_0 + \exp\left[\mu_n + \sigma_n (z_t)\right]$$

Where $\xi_0 = 11.0559$, $\mu_n = 2.5998$, $\sigma_n = 0.0712$ and z_t are the normally distributed random number with zero mean and unit variance.



(a) SANDARDISED SERIES OF RUNOFF

(c) STANDARDISED SERIES OF RAINFALL SEQUENCE (INDEPENDENT STOCHASTIC COMPONENT)

MAX. LAGS = 48; HANNING SMOOTHING FN. USED

FIG. 9 VARIANCE SPECTRUM ANALYSIS OF STANDARDISED SERIES RUNOFF & RAINFALL

Since ϵ_t series was obtained from X_t series using the equation 3.1!

$$X_t = AV + STD\ (\epsilon_t) \text{ and } S_t = X_t \times \sigma_x$$

Where AV = 0.0022, STD = 0.7720, σ_x = 1059.926

X_t is thus obtained by adding S_t, P_t and T_t as

$$X_t = (T_t + P_t + S_t) \quad \ldots\ldots\ldots\ldots\ldots \quad (7.10)$$

7.3.3. Non-parametric method of separating periodic and stochastic components of monthly runoff sequence:

The non-parametric method of standardization of X_t sequence of observed runoff, is by the transformation,

$$\epsilon_t = \frac{X_t - m_\tau}{S_\tau}$$ to obtain second-order stationarity,

where m_τ and the sample means and S_τ are the sample standard deviations at the position τ (monthly) calculated from equations (3.1) and (3.2) respectively.

Computer program no. 5 was run to obtain the stochastic component of X_t from equation (3.3). The monthly means and monthly standard deviations computed are given in table no. 6.

The correlogram and spectrum analysis of the stochastic component (figures 8 and 9) indicated the presence of dependent stochastic component.

In order to identify the order of A.R. model, the deterministic coefficients, were computed by using equations (3.42) and (3.44) from computer program no.3 and found to be R_1^2 =0.2920, R_2^2 = 0.3325, R_3^2 = 0.3619 for ϵ_t series. As per equation (3.46)

Second order a.r. model was adopted, giving the stochastic component as below

$$\epsilon_t = a_1 \epsilon_{t-1} + a_2 \epsilon_{t-2} + z_t \quad \ldots\ldots\ldots\ldots \quad (7.11)$$

The values of a_1 and a_2 were calculated from equations (5.95) and (5.96) with $r_1 = 0.9403$ and $r_2 = 0.4645$ from computer program no.5.

The dependent stochastic component was removed from the series by fitting a second order a.r. model using equation (5.97) and computer program no. 5, and the independent stochastic component z_t with mean = -0.0002 and standard deviation = 1.0001 was obtained. The correlogram (fig. 9) and spectrum (fig. 9) of this component clearly indicate, that it is a random component with second order stationarity.

The z_t series was tested for best fit distribution by Chi-square and K-S test for goodness of fit.

The Chi-square value was found to be 102.0 against the critical value of 56.00 for (45-5)= 40 degrees of freedom at 99 pct. confidence level. Calculations were similar to the previous cases. In this case also z_t series was divided into 60 ~~class~~ class intervals of equal probability. As the calculated value is greater than the critical value admissible, the normal distribution does not fit the random component.

The 3 - parameter log-normal distribution was tried for its fit to the z_t series. The computation of lower boundary parameter, by the method of moments of equation (3.61) was followed in the present case, also. The shift or lower boundary parameter

Z_0 was found to be 2.4694. The mean and standard deviation of log (X_t + Z_0) series was computed using computer program no. 6 and found to be 0.8107 and 0.3636 respectively.

The Chi-square test was performed on log(X_t + Z_0) series and the series was divided into 60 class intervals with equal probability. The Chi-square i.e. value was found to be 55.12 against the critical value of 56.00 for (49-9)= 40 degrees of freedom at 95 pct. confidence limit. Therefore the hypothesis is true that 3 - parameter log-normal distribution fits the X_t series.

The model: Having identified all the additive components of the time series model, the model is described below

$$\epsilon_t = Z_t \times \sqrt{1-(a_1^2 + a_2^2 + 2a_1 \cdot a_2 r_1)} + a_1 \epsilon_{t-1} + a_2 \epsilon_{t-2}$$

where a_1 = 0.4110, a_2 = 0.2591 and r_1 = 0.5409

$$X_t = Z_0 + \exp[\mu_n + \sigma_n (Z_t)]$$

where Z_0 = 2.4094, μ_n = 0.0107

σ_n = 0.3684, Z_t = normally distributed random number with zero mean and unit variance.

$$X_{D,\tau} = \epsilon_t \times STD(\tau) + AV(\tau) \quad \ldots\ldots\ldots (7.12)$$

where AV(τ) and STD (τ) are the monthly means and monthly standard deviation of the τ-th month of observed runoff sequence given in table No. 4.

7.3.4. Comparison of the analysis of time series of monthly runoff by the three approaches (Log. Transformation, classical and non-parametric).

From the analysis of time series structure using non-parametric method and parametric methods (transformation and log transformation, as discussed in the earlier sections 7.31, 7.32, 7.33) a comparison can be made about these approaches as follows:

i) The non-parametric method was used for detecting the character of stationary stochastic component and to use it as the criteria for evaluating the relative performance of Log. transformation and no transformation cases. Though in Log. transformation approach assumption was made regarding monthly standard deviation being proportional to monthly mean, for the data under study this approach gave better results in comparison to no transformation case where monthly standard deviation was assumed to be constant.

ii) In the Log. transformation approach the periodic component was modelled by only two significant harmonics i.e. 12 and 6-month explaining 56 pct. of the variance, of which 12 - month period alone accounted for 53 pct; while in no transformation case 12,6,4,3,2.4 months period were needed.

iii) The correlogram and (figures 3 and 5) and spectrum (figs. 4 and 9) for the Log. transformed case and non-parametric case for stochastic and random components compare reasonably well, while those for no transformation case (fig. 6 and 9) do not compare that well with the non parametric case.

iv) The dependence model for Log. transformation case is of third order a.r. process while that for non-parametric is of second-order and that for no-transformation is of first order. This probably indicates the effect of considering monthly standard deviation as constant or proportional to monthly mean or as periodic. This aspect needs further study.

TABLE NO. 9

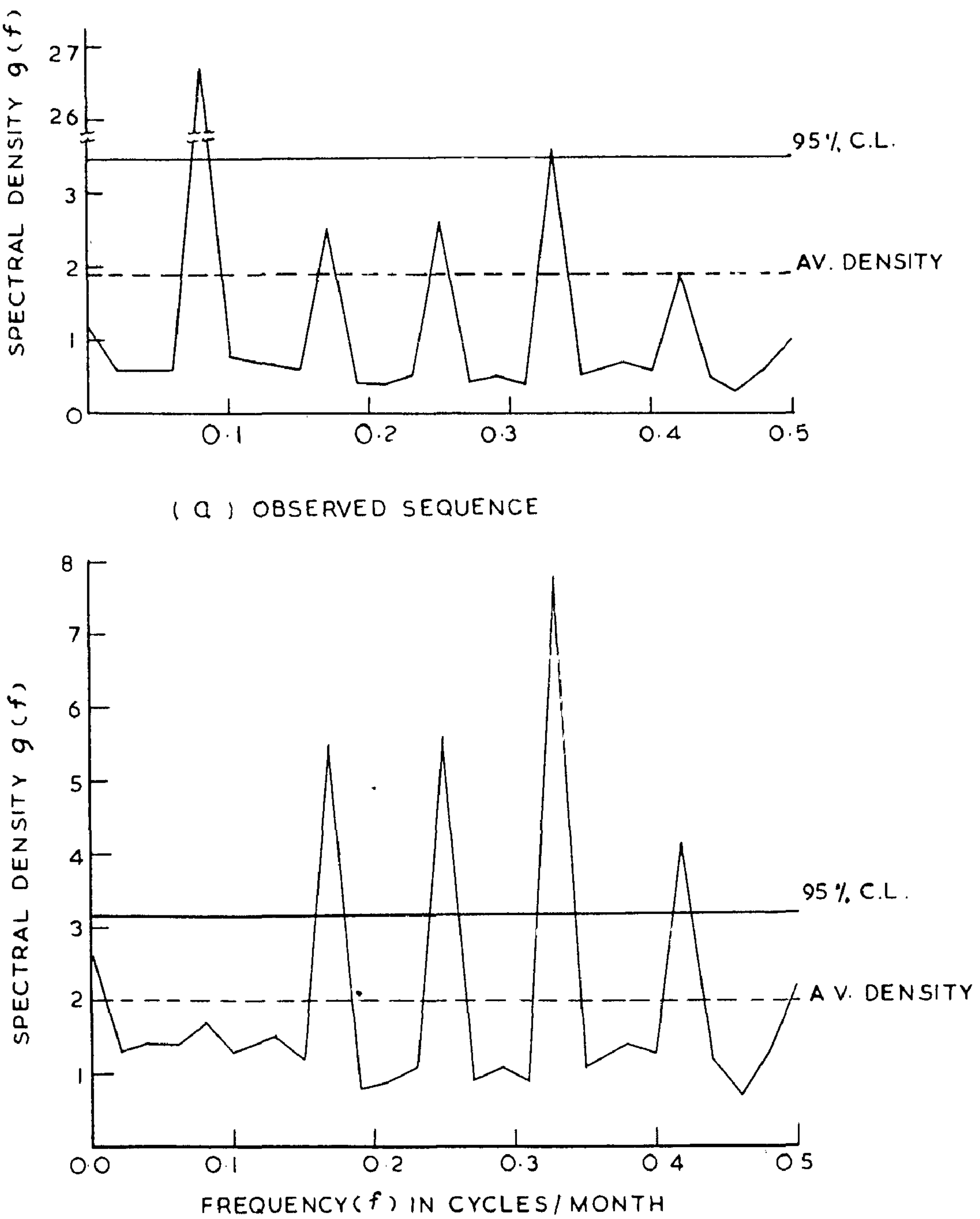
STATISTICS OF OBSERVED RAINFALL SERIES

S. No.	Month	Mean in inches	Std. dev. in inches	Ser. Coeff "r"	Coeff. of Skewness C_s	Reg. Coeff. B_1	Coeff. of var C_v in Pct.	Remarks
1.	January	0.20	0.50	-0.14	0.91	-0.08	248	
2.	February	0.10	0.20	0.16	2.40	0.59	199	
3.	March	0.64	0.70	0.07	1.47	0.19	109	
4.	April	3.33	1.63	0.12	0.17	0.21	47	
5.	May	5.08	2.03	0.001	0.05	0.00	40	
6.	June	7.09	3.03	0.16	1.15	0.26	49	
7.	July	16.11	6.17	0.15	1.79	0.09	38	
8.	August	8.24	3.66	-0.15	1.41	-0.07	44	
9.	September	4.90	1.97	-0.19	0.94	-0.23	40	
10.	October	6.99	2.89	-0.02	1.16	-0.02	41	
11.	November	2.41	2.17	-0.24	1.54	-0.09	90	
12.	December	0.61	0.78	0.08	1.41	0.05	127	

Series mean = 4.668 inches

Series Std. Dev. = 3.297 inches

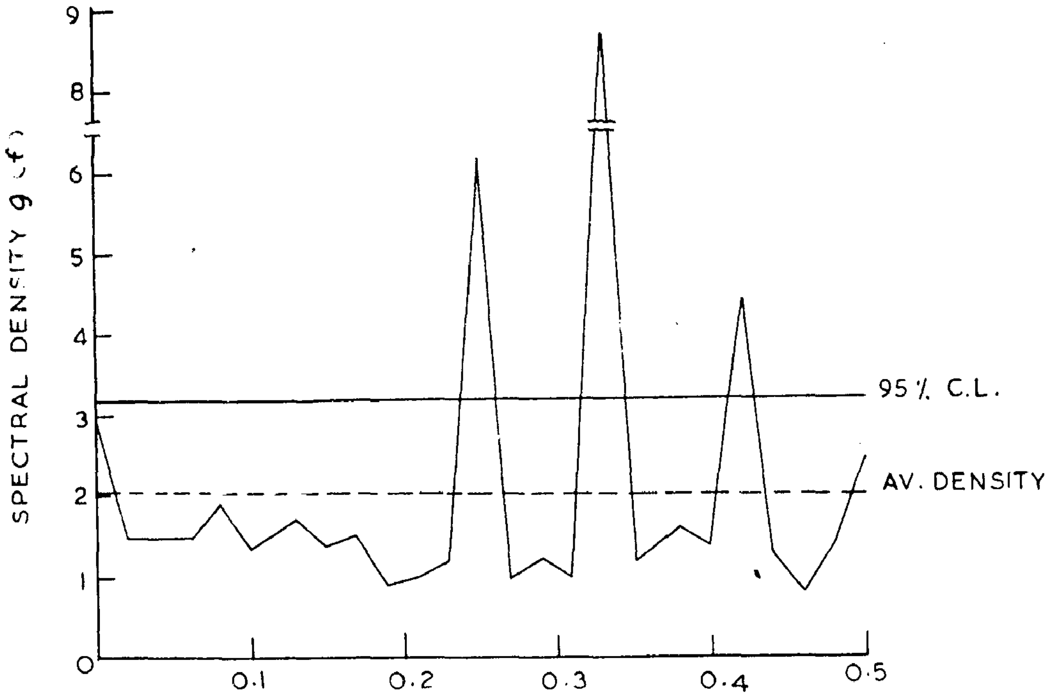
First Ser. Cor. Coeff = 0.49



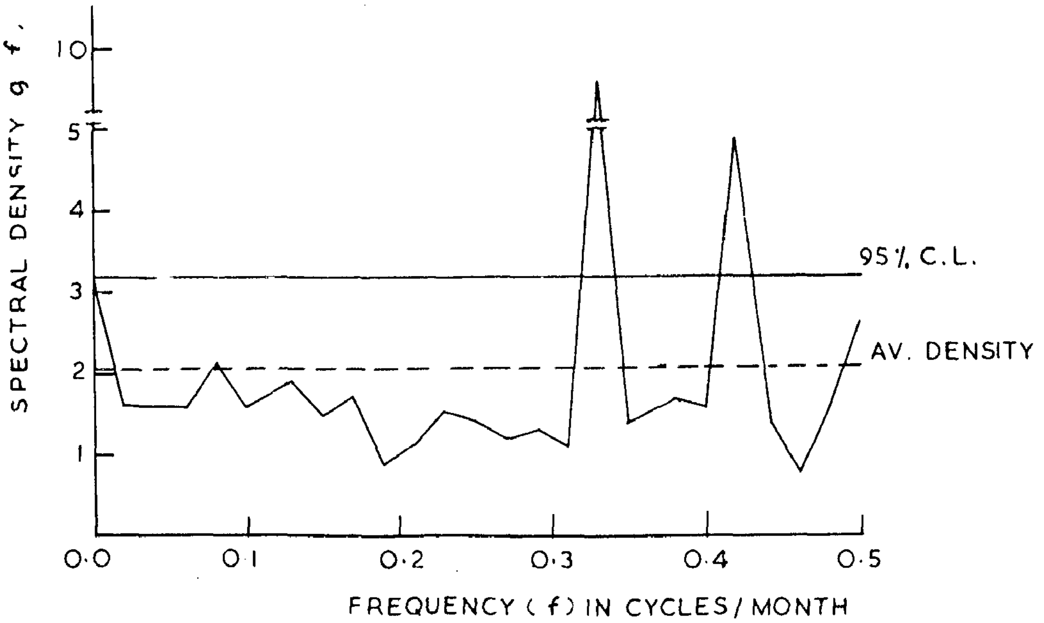
(b) 12-MONTH PERIOD REMOVED

MAX. LAGS=48; HANNING SMOOTHING FN. USED

FIG.8(i) VARIANCE SPECTRUM ANALYSIS FOR RAINFALL SERIES



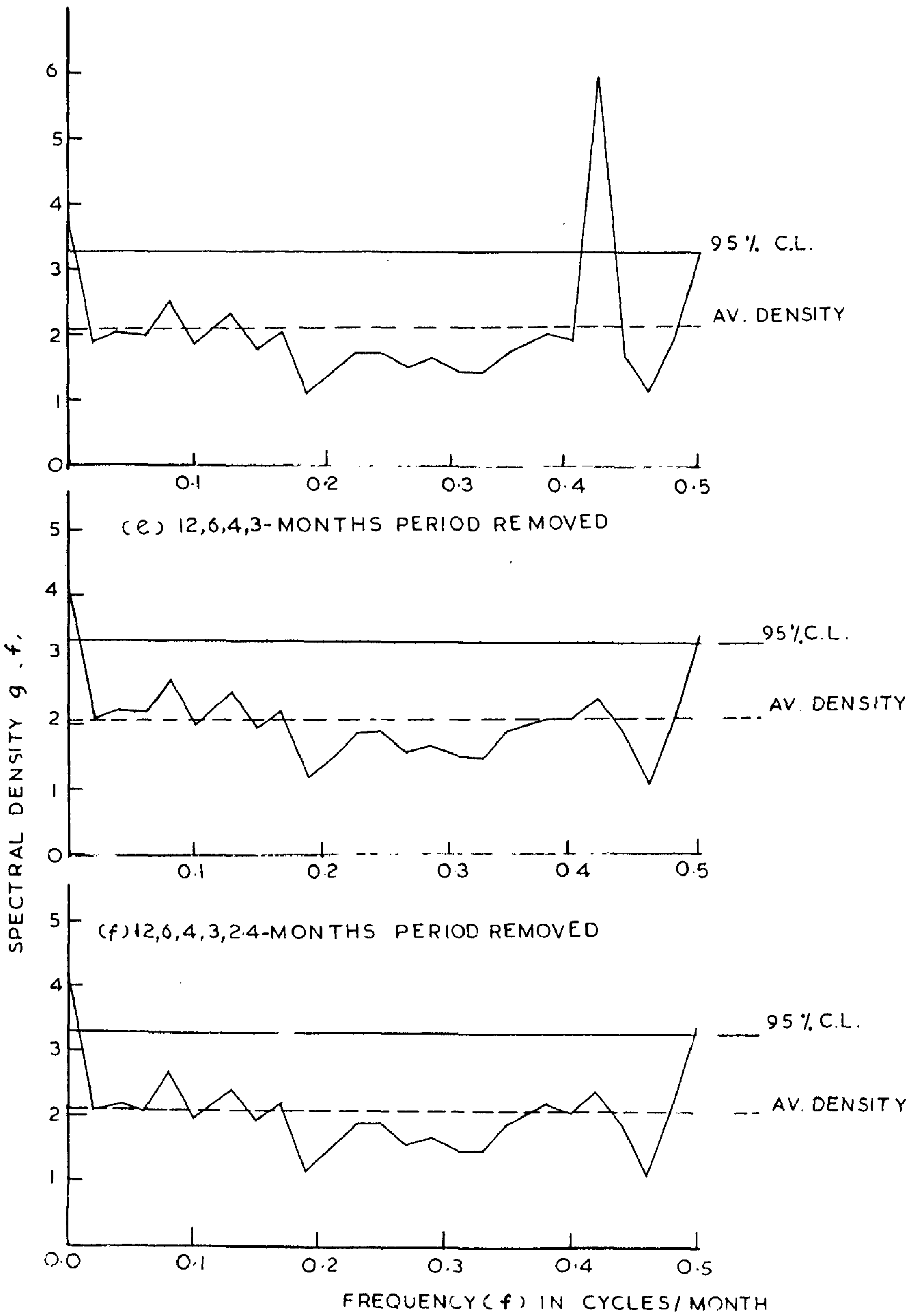
(C) 12,6-MONTHS PERIOD REMOVED



(d) 12,6,4 - MONTHS PERIOD REMOVED

MAX. LAGS = 48; HANNING SMOOTHING FN. USED

FIG. 8(ii) VARIANCE SPECTRUM ANALYSIS FOR RAINFALL SERIES



(g) RANDOM COMPONENT- FIRST ORDER A.R. REMOVED FROM (f)

MAX. LAGS = 48; HANNING SMOOTHING FN. USED

FIG.8(iii) VARIANCE SPECTRUM ANALYSIS FOR RAINFALL SERIES

7.3.5. Classical Approach of Analysis of Monthly Rainfall Sequence:

The rainfall in inches of Depth over the catchment given in table 3 was used in the analysis.

In the composition model of equation (3.4) i.e. $X_{p,\tau}=\mu_{\tau}+\sigma_{\tau}\epsilon_{p,\tau}$ it is observed, from the statistics of monthly rainfall sequence given in table No. 3, that neither σ_{τ} can be assumed to be constant nor μ_{τ} and σ_{τ} are proportional to each other.

However the classical approach involving no transformation of the data is followed in the analysis of monthly rainfall sequence and compared with the non-parametric method of separating periodic and stochastic components which is discussed in the next section.

Computer program 3 was run to obtain the Fourier coefficients from periodogram analysis. The same program was used for the correlogram and spectrum analysis of observed rainfall sequence and the series with each of six harmonics, corresponding to 12,6, 4,3,2.4 and 2- months period removed in turn.

The correlogram and spectrum analysis was done for a maximum lag equal to 96 and also 40 and the results were compared. The use of four spectral windows i.e. Hanning, Bartlet, Parzen, and Tukey, for obtaining smoothed spectrum density functions, was similar to the analysis of monthly runoff sequences. The significant periods for the observed sequence were 12 and 3 months period as shown in figure 8. After the removal of 12 - month period 6,4,3,2.4 - month periods became significant. After the removal of 12 and 6- month period 4,3,2.4 - month periods were left as significant.

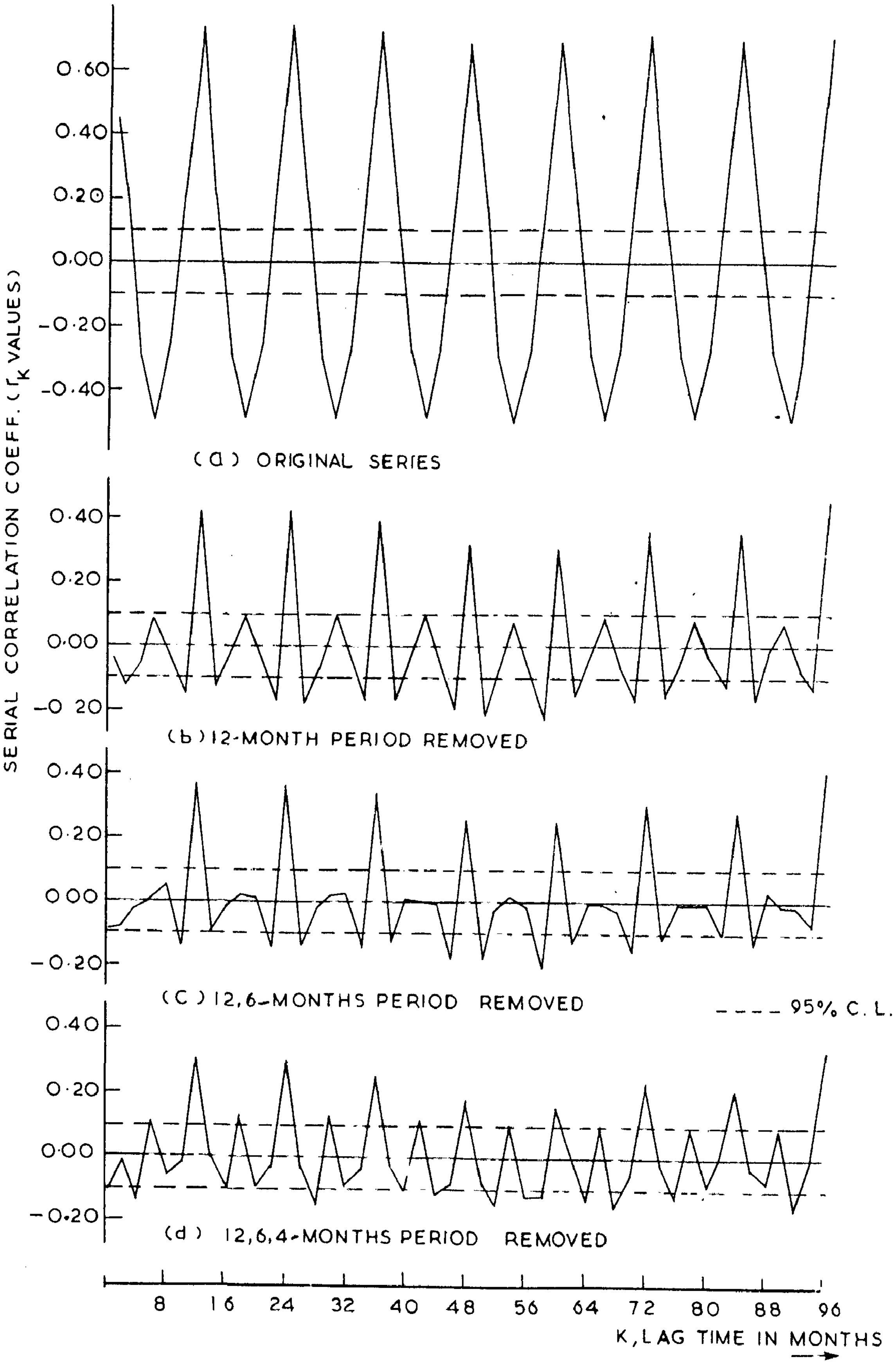


FIG. 7(i): CORRELOGRAM ANALYSIS-RAINFALL SERIES

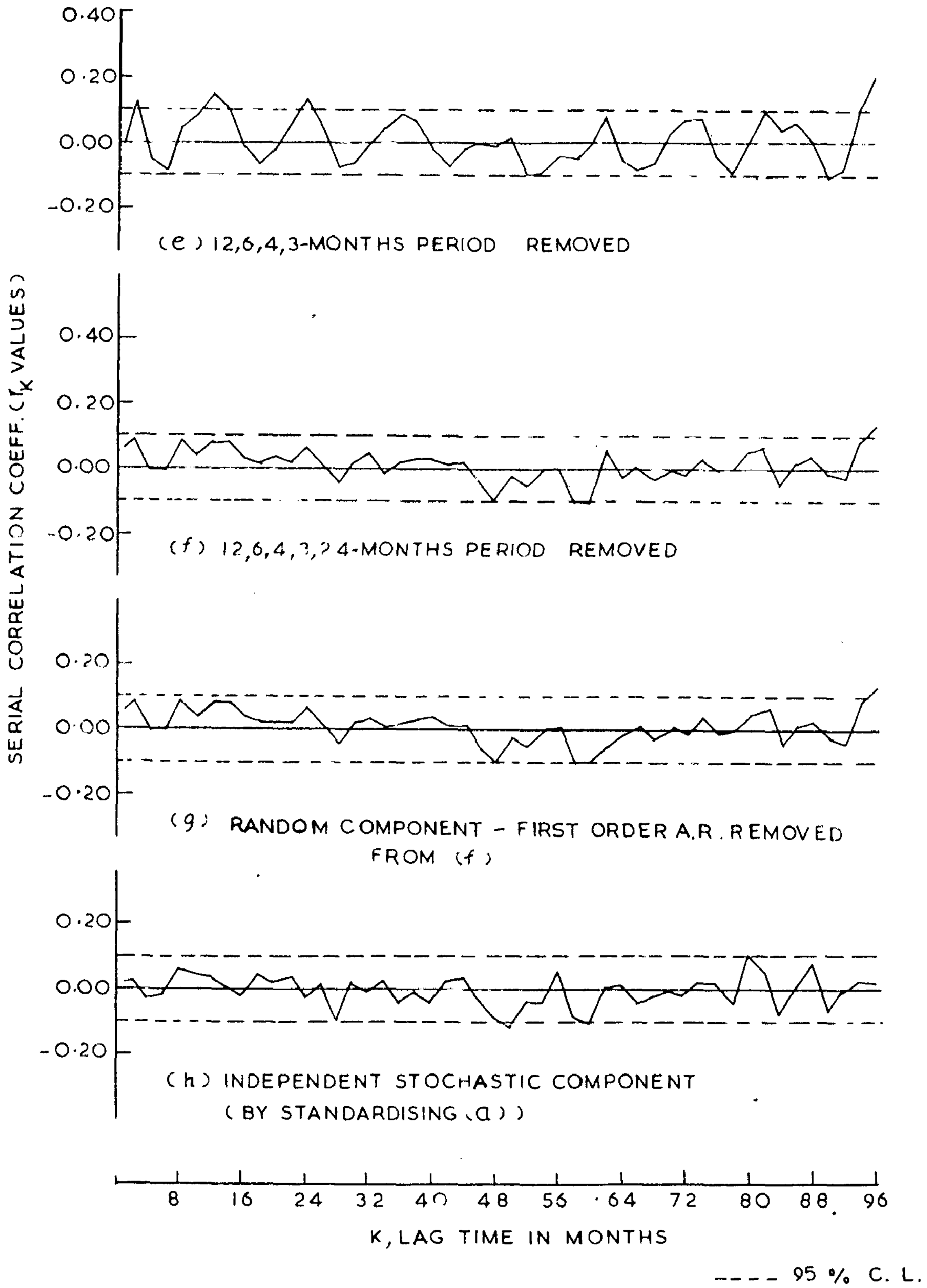


FIG.7(ii) CORRELOGRAM ANALYSIS — RAINFALL SERIES

The removal of 12,6,4 month period left 3 and 2.4 month period significant. After the removal of 12,6,4,3-months period 2.4-month period remained significant. The removal of 12,6,4,3,2.4-month period left X_t series with the stochastic component .

The correlogram of observed sequence (Fig. 7) indicate 12,6,4-month periods. The 12-month removed series reveal 6,3-months periods significant. The removal of 12,6-months period indicate a significance of 4,3-months period. The removal of 4-month period left the series with 3 months period significant. After the removal of 12,6,4,3-months period, 2.4-month period remained significant. Thus the removal of 12,6,4,3,2.4-months period left the series with the stochastic component.

It was observed (Table No. 8) that the five significant harmonics explain about 70 pct. of the total variance of X_t series , out of which about 53 pct. explained by the 12 - month cycle alone. The variance explained by each harmonic is computed using equation (3.9) and computer program 4.

The Fourier coefficients A_j and B_j using equation (3.6) and (3.7) were computed as

- 4.9736 , - 3.2069 for 12 month period.

1.0471 , 1.0311 for 6 month period.

- 0.9530 , - 1.4633 for 4 month period.

- 1.5350 , 1.9799 for 3 month period.

+ 1.1640 , - 0.1089 for 2.4 month period respectively.

The above values of A_j and B_j were substituted in equation (3.13) to obtain the periodic function in mean as

$$\mu_t = \mu_x + \sum_{j=1}^{5} \left[A_j \cos(2\pi K j/12) + B_j \sin(2\pi K j/12) \right] \quad \ldots (7.13)$$

where μ_x is the mean of trend removed series.

The Y_t series obtained after the removal of significant harmonics had the mean = 0.0001 and variance = 0.8806 (Std. deviation = 0.9381) and hence from equation (5.15)

$$\epsilon_t = (Y_t - 0.0001) / 0.9381$$

The auto-correlation coefficients of Y_t and ϵ_t series were same, as observed from the results of computer program no.9. Since the correlogram and spectrum of Y_t and ϵ_t are similar (fig. 7 and fig. 8)

The difference between the variance of time series Y_t and the total of explained variance of the significant harmonics which is attributed to the stochastic component, works out to 90 pct. (table no.8)

The correlogram (fig. 7) and spectrum (fig.8) of the stochastic component do not indicate any significant dependence nature and hence it can be assumed as an independent stochastic component. However for the present study an auto-regressive model of 1st order was fitted on the basis of determination coefficient approach and the dependence structure was removed from the stochastic component, and the remaining component was taken as independent stochastic component.

The determination coefficient approach was used for the analysis of stochastic component of rainfall sequence similar to that in the analysis of stochastic component of runoff sequence. R_1^2, R_2^2 and R_3^2 of equations (5.42) to (5.44) were found to be 0.0026, 0.0103, and 0.0103, since $R_2^2 - R_1^2 = 0.0077 < 0.01$ and $R_3^2 - R_2^2 = 0 < 0.02$ first order a.r. model was adopted as per equation (9.93) giving stochastic component as below:

$$\epsilon_t = a_1 \epsilon_{t-1} + z_t$$

where $a_1 = r_1 = 0.0507$ of X_t series.

The dependent stochastic component was removed by using equation (9.94) and the computer program no. 9, and z_t the independent stochastic component (random component) was obtained. Mean and standard deviation of z_t series were zero and one, respectively. The correlogram (fig. 7) and spectrum (fig. 8) of the residual series reveal that this series can be considered as a random component.

The z_t series was tested for fitting normal distribution function by chi-square test. The z_t series was divided into 60 class intervals and the chi-square value for (45-3) = 42 degrees of freedom was found to be 31.34, against the tabulated value of 56.96 at 95 pct. confidence limit. Hence the Normal distribution fits z_t series.

The model: Having identified all the additive components of the time series model

$$X_t = T_t + P_t + \epsilon_t$$

the individual components are described below:

T_t = series mean, 4.0620

$$P_t = A_1 \cos(2\pi t/12) + B_1 \sin(2\pi t/12)$$
$$+ A_2 \cos(4\pi t/12) + B_2 \sin(4\pi t/12)$$
$$+ A_3 \cos(6\pi t/12) + B_3 \sin(6\pi t/12)$$
$$+ A_4 \cos(8\pi t/12) + B_4 \sin(8\pi t/12)$$

where $A_1 = -4.9790$, $B_1 = -5.0237$

$A_2 = 1.0471$, $B_2 = 1.0010$

$A_3 = -0.3290$, $B_3 = -1.4692$

$A_4 = -1.5053$, $B_4 = 1.5755$ and $A_5 = 1.1640$, $B_5 = -0.1020$

$$\epsilon_t = \left(Z_t \sqrt{1-r_1^2}\right) + a_1 \epsilon_{t-1}$$

Where $r_1 = a_1 = 0.0507$

Z_t is generated as $Z_t = \mu + \sigma R_t$ where $\mu = 0$ and $\sigma = 1.0$

R_t = normally distributed random numbers with mean zero and variance unity.

The Y_t of equation (5.19) is given by

$Y_t = AV + STD\ (\epsilon_t)$ and $S_t = Y_t \times \sigma_x$

$AV = 0.0001$, $STD = 0.9901$, $\sigma_x = 5.2969$

X_t can be obtained by adding T_t, P_t and S_t components as

$X_t = (T_t + P_t + S_t)$ . . . . (7.19)

7.5.6. Non-Parametric Method of Separating Periodic and Stochastic Components in the Monthly Rainfall Sequences:

The non-parametric method of standardisation of rainfall sequence was done similar to that of runoff sequence discussed at section (7.55) by the transformation $\epsilon_{p,\tau} = \dfrac{x_{p,\tau} - m_\tau}{s_\tau}$

Where m_τ and s_τ are the monthly means and standard deviations of observed sequence and given in table no. 5

The correlogram and spectral analysis (fig. 7 and 9) of the stochastic component did not reveal presence of any dependent stochastic component. Thus the stochastic component (mean zero and variance unity) was inferred as independent stochastic component and the normal distribution tested for E_t series, by chi-square test. The series was divided with 45 class intervals with equal probability similar to the earlier cases. The Chi-square value for (45-3) = 42 degrees of freedom was found to be 61.12 where as the critical value at 95 pct. level of confidence is 58.96; and the critical value at 99 pct. confidence level as for 42 degrees of freedom is 66.00. Hence the normal distribution fits E_t series at 99 pct. confidence level, though not at 95 pct. confidence limit.

The model: Having identified the two additive components of the time series model, the model is described below

E_t, the stochastic component is described as

$$E_t = \mu + \sigma (z_t)$$

where $\mu = 0$ and $\sigma = 1.0$

z_t = Normally distributed random number with zero mean and unit variance.

$$X_{p,\tau} = E_t \times STD(\tau) + AV(\tau) \ldots\ldots\ldots \quad (7.16)$$

where $AV(\tau)$ and $SD(\tau)$, of τ-month are the monthly means, and monthly standard deviations of observed rainfall sequence, given in table no. 9.

7.9.7. Comparison of the analysis of monthly rainfall sequence by the two approaches (classical and non parametric):

The findings of the analysis of monthly rainfall sequence from the two approaches i.e. i) classical involving no transformation of the data , ii) non-parametric method, are as follows:-

The monthly means and standard deviations of monthly rainfall sequence (table 5) were not proportional, unlike the monthly runoff sequence. However the classical approach assuming monthly standard deviation as constant, and non-parametric method were adopted in the present study.

The analysis of monthly rainfall sequence was done similar to the analysis of untransformed monthly runoff sequence. 12-month cycle and its 6,4,3,2.4-months subharmonics were required for the description of periodic component by Fourier series analysis. The explained variance of the periodic component was about 70 pct. of which 53 pct. was accounted by 12- month cycle alone. The stochastic component of rainfall sequence did not indicate presence of dependence component by correlogram and/or spectrum analysis (fig. 7 and fig. 8). However use of determination coefficient approach suggested first order a.r. model. The stochastic component of non-parametric method also similarly displayed a completely independent stochastic process as seen from the correlogram in fig. 7, and spectrum in fig. 9. The removal of first order a.r. process was not indicated by correlogram and spectrum analysis, however for the sake of comparative study a first order auto-regressive process was separated from stochastic component of parametric approach. However the spectrum (fig. 8) and correlogram (fig. 7) after removal of this dependence model did not show any significant difference from those of non-parametric approach (fig. 7 and fig.9) where

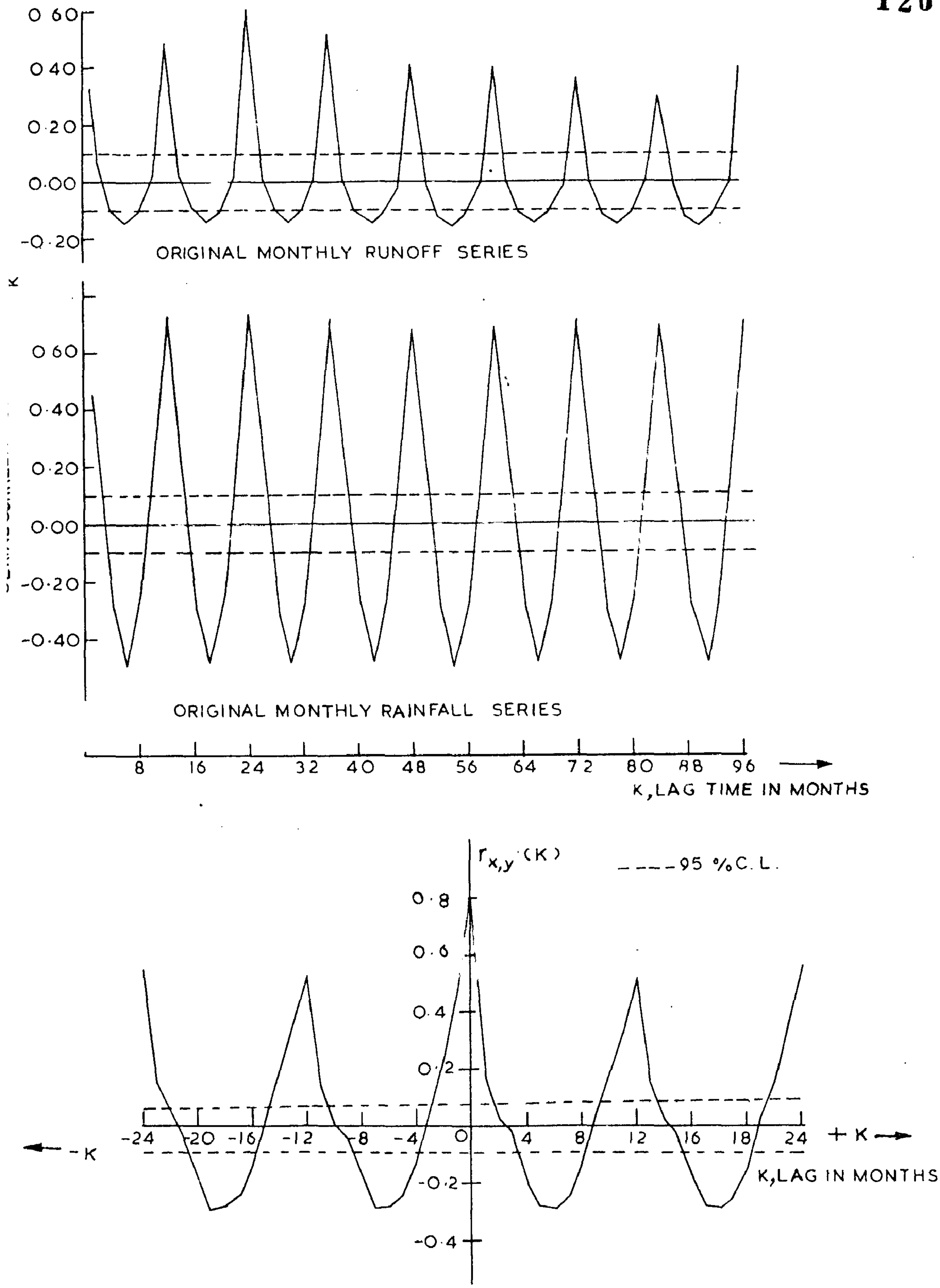


FIG. 10 CROSS-CORRELOGRAM BETWEEN MONTHLY RAINFALL & MONTHLY FLOW

stochastic component was taken as an independent process. Hence on the basis of these results it can be concluded that after the removal of deterministic component for an observed monthly rainfall sequence, the remaining stochastic component can be taken as independent (random) process.

The normal distribution fitted the independent stochastic component for both parametric and non parametric approaches.

7.3.6. Cross-Correlation between series of monthly rainfall and monthly flow:

Lag cross-correlation can be used for investigating relationships between monthly rainfall and monthly runoff. For the calculation of r_{xy} (K) values use of equation (3.19) was made in which $x_{i,t}$ was replaced by $y_{i,t}$ where $y_{i,t}$ indicate the monthly rainfall series. Computer program 9 was used for computing r_{xy}(K) values. The correlogram of observed rainfall, observed runoff sequences and the cross-correlogram between rainfall and runoff are shown in figure no. 10.

The cross-correlogram is not exactly symmetrical for K = 0, because K and -K give different values of r(K) and r(-K). Maximum absolute value of r_{xy} was 0.794 at K = 0, indicating the time lag of maximum correlation. The confidence limits were drawn at 95 pct. probability level using equation (3.20). A large no. of values lie outside the 95 pct. confidence limit which is the case of correlated time series. The cross-correlogram for -K, corresponds to rainfall preceding runoff, and has a physical meaning and justification. On the other hand the cross-correlofrom fig. 9 K corresponds to runoff preceding rainfall and the results of a formalistic approach has no physical meaning. It is observed from fig. 10 that the monthly rainfall is correlated to monthly runoff for a maximum of K = -2 i.e.(2 months) i.e. to say that flow in any month is influenced by upto previous two months.

7.4. STATISTICS OF MONTHLY RUNOFF GENERATED BY TIME SERIES MODELS:

Monthly Runoff sequences were generated by the Time Series Models derived from the different approaches of analysis, namely i) by the logarithmic transformation of the observed and runoff sequences (equation 5.9), ii) by the classical approach (where no transformation is needed to the observed sequence) as given by equation (5.4), iii) by the non-parametric method (equation 5.5)

The normally distributed random numbers used in the above Time Series Models, were generated by using the computer program no. 9 (10), in which the uniformly distributed numbers were first generated and transformed as normally distributed random numbers by Box and Muller model of equation (5.9).

Computer program no. 10 was used for generating the monthly runoff sequences for a period of 50 years and also 100 years, to facilitate comparative study of the statistics of generated sequences, for different periods.

The monthly runoff sequences were generated for both 50 and 100 years period, for all the three approaches. Figures 11

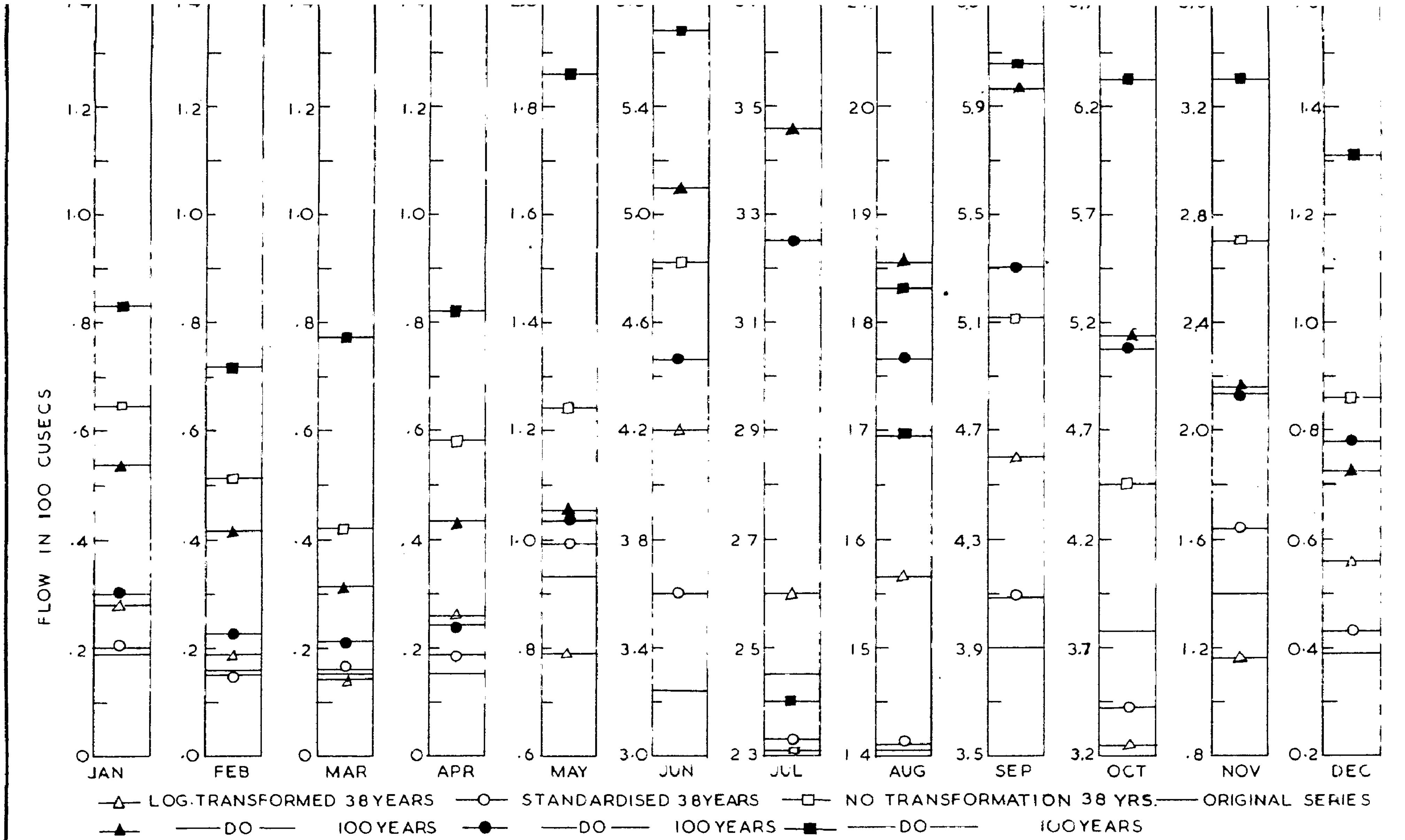


FIG. 11 MONTHLY MEAN OF GENERATED RUNOFF SEQUENCE BY TIME SERIES MODELS

than zero, K corresponds to runoff preceding rainfall and the results of a formulistic approach has no physical meaning. It is observed from fig. 10 that the monthly rainfall is correlated to monthly runoff for a maximum of K = -2 i.e.(2 months) i.e. to say that flow in any month is influenced by upto previous two months.

7.4. GENERATION OF MONTHLY RUNOFF SEQUENCES BY TIME SERIES MODELS:

Monthly Runoff sequences were generated by the Time Series Models derived from the different approaches of analysis, namely i) by the logarithmic transformation of the observed and runoff sequences (equation 5.9), ii) by the classical approach (where no transformation is needed to the observed sequence) as given by equation (5.4), iii) by the non-parametric method (equation 5.5)

The normally distributed random numbers used in the above Time Series Models, were generated by using the computer program no. 8 (14), in which the uniformly distributed numbers were first generated and transformed as normally distributed random numbers by Box and Muller model of equation (5.3).

Computer program no. 10 was used for generating the monthly runoff sequences for a period of 50 years and also 100 years, to facilitate comparative study of the statistics of generated sequences, for different periods.

The monthly runoff sequences were generated for both 50 and 100 years period, for all the three approaches. Figures 11

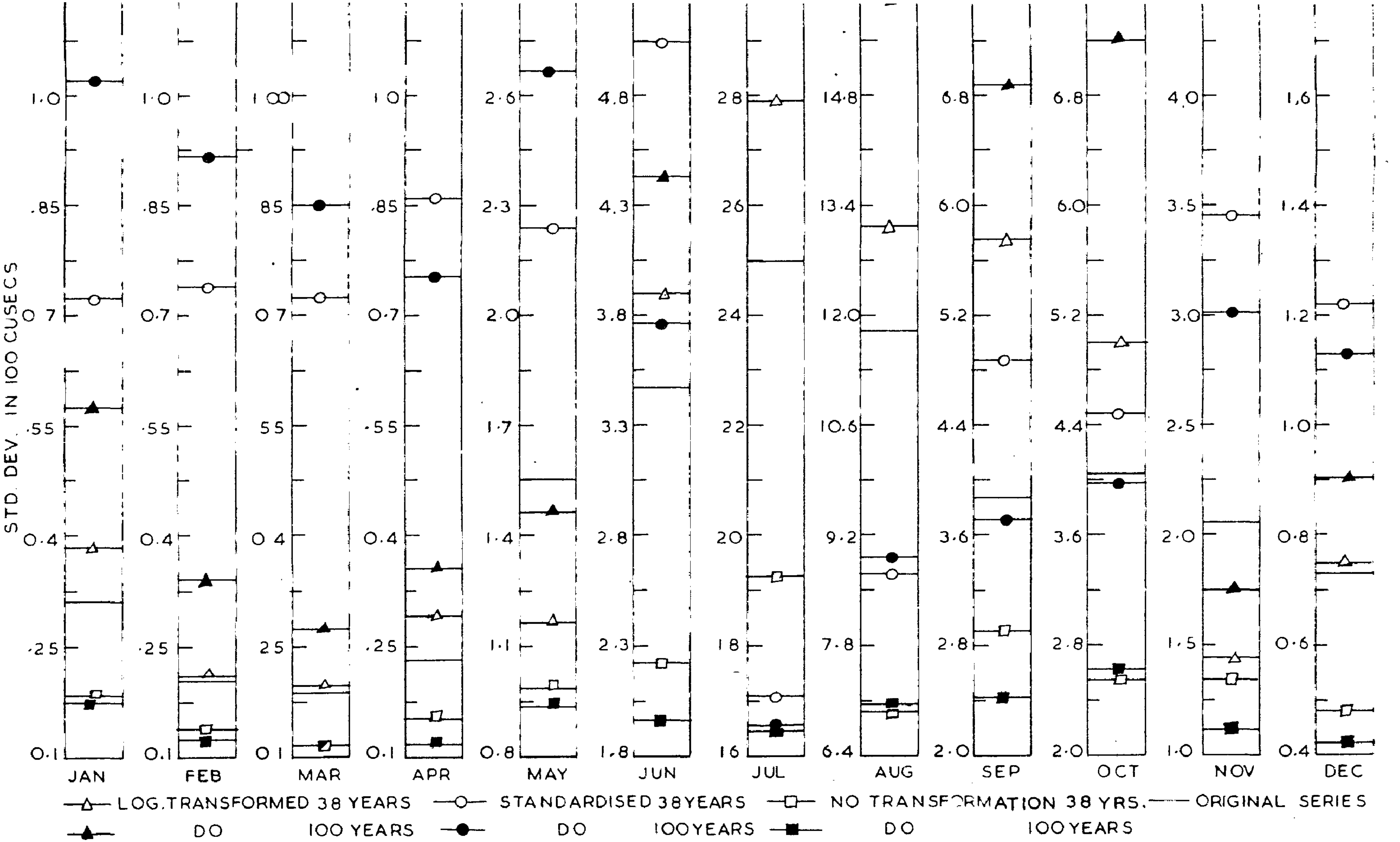


FIG. 12 MONTHLY STANDARD DEVIATION OF GENERATED SEQUENCE BY TIME SERIES MODELS

and 12 show the monthly mean and monthly standard deviation of generated runoff sequences for 53 years and 100 years from the three time series models with respect to the monthly mean and monthly standard deviation of the observed sequence.

Monthly means of generated data for 53 years period, using non-parametric approach compare quite reasonably with that for original sequence though they are somewhat higher. Whereas when 100 years data was generated they were however still higher than those for 53 years generated sequence.

The log. transformation approach results, for monthly means and monthly standard deviations of generated sequence for both 53 years and 100 years period, did not compare that well with that of original data. However they are nearer to the original monthly mean and monthly standard deviations as compared to the results from no transformation case.

The monthly standard deviation of the generated data using non-parametric approach are generally closer to those for original data, though they are somewhat lower, when 53 years data is generated. When 100 years data is generated these values are somewhat lower than those for 53 years generated sequence.

The statistics of generated data depends to some extent on the best fit distribution found for random component, and the method of generation and quality of random numbers of that particular distribution used for generating synthetic sequences. In addition various assumptions regarding trend and periodicity in parameters such as mean, standard deviation, serial correlation

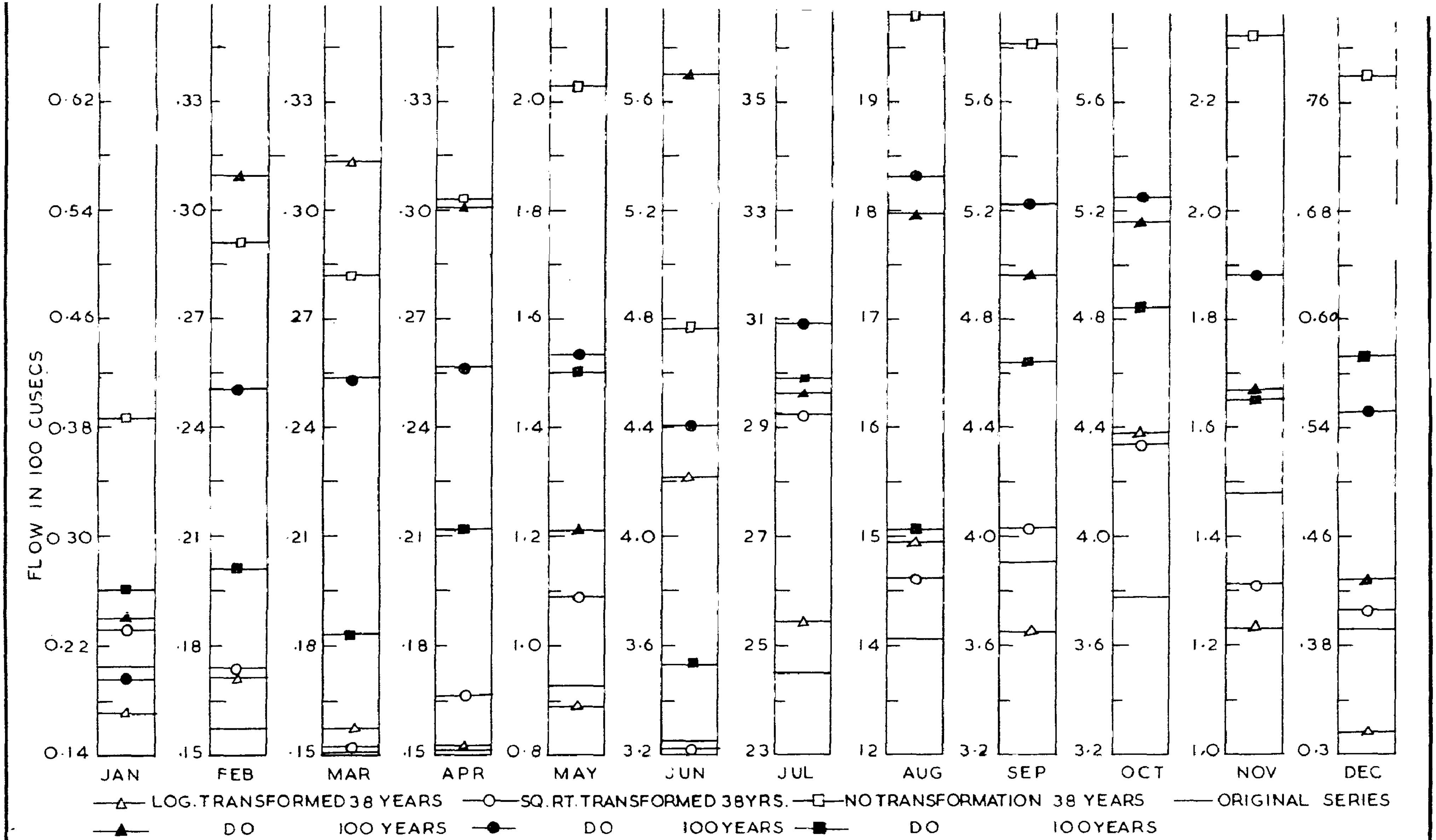


FIG. 13 MONTHLY MEAN OF GENERATED RUNOFF SEQUENCE BY THOMAS-FIERING MODELS

coefficient etc., and structure of dependence model selected also affect the statistics of generated sequence.

It is not likely that any one sequence generated by using a particular model may give almost similar statistics such as mean, standard deviation etc. as the original data. However if a number of such samples are generated their combined statistics should tend to the statistics of original data with a reasonable probability.

7.5. ANALYSIS OF MONTHLY RUNOFF SEQUENCE BY NON-STATIONARY APPROACH

The monthly runoff sequences are basically non-stationary, and this is implied in the Thomas and Fiering (55) approach. The non stationarity in monthly means, standard deviations, and the correlation between successive months, is evident from table 4. The Thomas Fiering Model (equation 4.1) uses the statistics of the observed monthly runoff sequences given in table no. 4. Monthly runoff sequences were generated for a period of 50 years and also 100 years by using equation (4.1) (computer program no. 11), statistics of which is shown in figure 13 and 14. It was found that some values in the generated sequence were negative. These negative values are encountered they are retained and used to derive the subsequent values in the sequence. Once generated sequence is complete, however, all negative values in the generated sequence were replaced by zero.

The logarithmic transformation of the observed sequence and fitting Thomas Fiering Model (computer program 12) to the

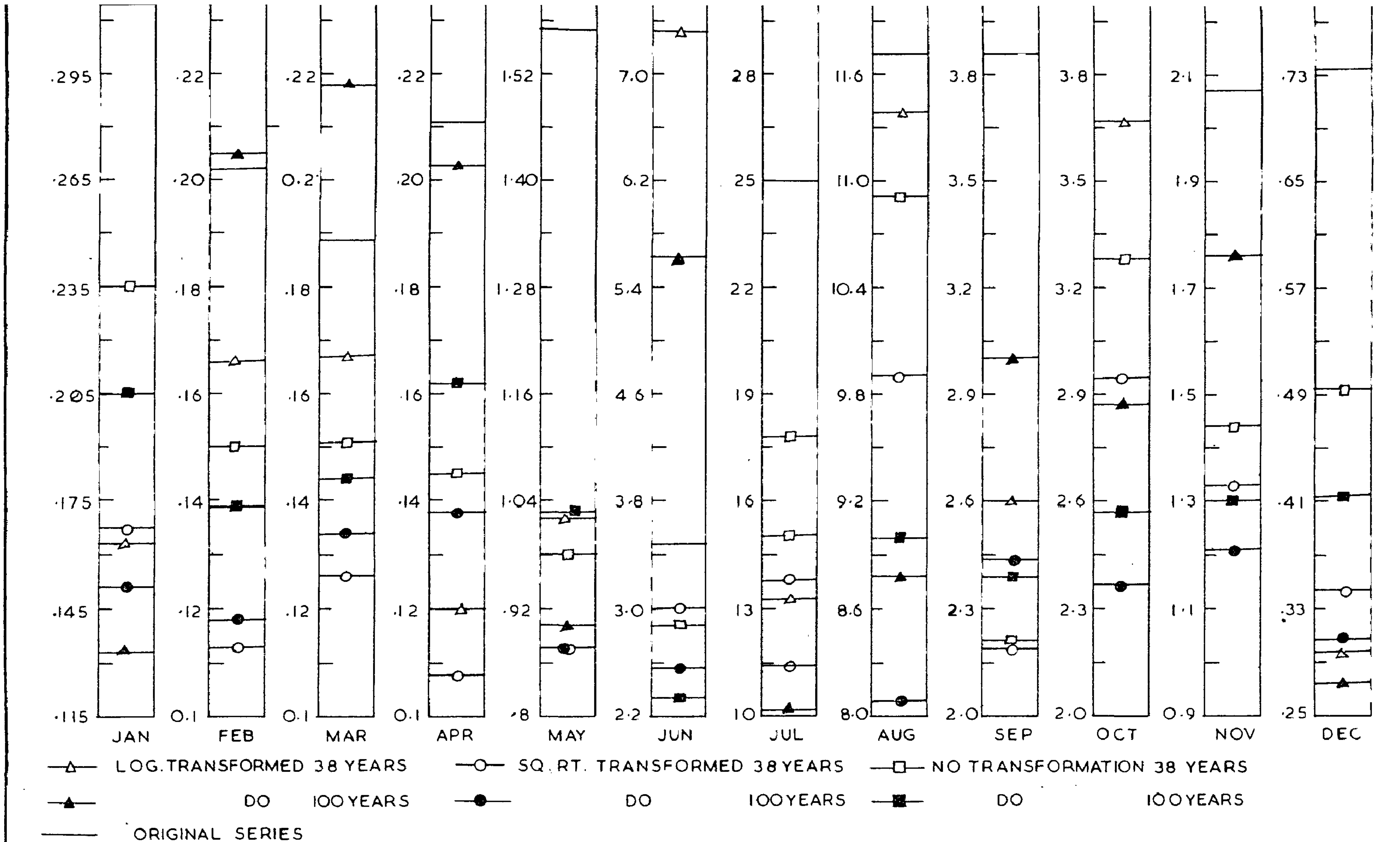


FIG. 14 MONTHLY STANDARD DEVIATIONS OF GENERATED SEQUENCE BY THOMAS FIERING MODELS

logarithms of flow, has the advantage of eliminating the negative flows that occur when untransformed flows are used in the model. To avoid infinite logarithms when flow is zero, as suggested by Rodriguez et.al. (29) an increment of 0.012 times the average individual monthly flow was added to each historical flow before computing the logarithm. The square root transformation of observed flows before fitting the Thomas Fiering Model (computer program 13) has also an advantage of eliminating the negative values in generated sequence when untransformed flows were used. Monthly runoff sequences were generated for both 50 years and 100 years by fitting a Thomas Fiering model for the untransformed flows and also their square root and logarithmic values. The monthly means and standard deviations of the generated sequences are shown in fig. 13 and fig. 14.

Computer program no. 14 incorporates the skewness of the monthly flows. The transformation suggested in equation (4.2) and (4.9) allow the normal random variate to be replaced by a random variate that has approximately a gamma distribution. Z'_1 from equation (4.9) was calculated, which replaced the normally distributed random numbers of equation (4.1). In order to have a comparative study the sequences were also generated both for 50 and 100 years period using the approach which incorporates skewness parameter also. However the monthly means and monthly standard deviations of generated data were very high in comparison to those for original data, probably due to some statistical instability. Fiering and Jackson (19) have also indicated the possible presence of extreme instability of higher moments when this transformation approach is used.

For the three approaches viz. no transformation, Log. transformation, and square root transformation using non-stationarity approach of Thomas Fiering Model, the results for 50 years and 100 years generated sequence indicate that both square root transformation and Log. transformation are comparatively better than no transform as shown in fig. 13 and 14. The monthly means during low flow season are nearly of same order for both these techniques, however during high flow season Log. transformation method gave comparatively better results. The monthly standard deviation for generated data using both square root and Log. transformation did not compare that well with those for original data (fig. 13 and fig. 14), the values are generally lower for generated data.

CHAPTER - 8

CONCLUSIONS AND SCOPE FOR FURTHER STUDY

8.1. I. CONCLUSIONS:

i) Long term trend was not significant in 38 years of the observed sequences of monthly rainfall and runoff analysed in this study.

ii) For the separation of periodic and stochastic components parametric methods of logarithmic transformation and notransformation of observed sequences and also non parametric method were studied. From the comparative study of the three approaches, the logarithmic transformation in which periodicity in monthly mean and standard deviations are accounted (with the assumption that monthly means are proportional to monthly standard deviations), gave comparatively better results. 12 - month main harmonic and 6-month subharmonic were found to be sufficient for the description of periodic component by Fourier series, in the case of analysis of logarithmically transformed data. The order of the dependence model of the stochastic component in each of the three cases were found to be different on the basis of coefficient of determination approach. 3 - parameter log. normal distribution fitted the independent stochastic component for no transformation case and non-parametric case. For the logarithmic transformation, normal distribution fitted the random component. Thus it can be concluded that random component in all the three cases fitted a log normal type of probability distribution.

iii) The generation of synthetic data for 50 years and 100 years period using log. transformation model gave reasonably satisfactory results as judged by comparison of monthly means and monthly standard deviations of observed and generated data. However if a number of such samples are generated then only these combined statistical characteristics will tend to the statistics of original sequence with certain probability. When only one or two samples are generated the statistical characteristics of the generated data depends somewhat on best fit distribution for the random component and method used to generate random numbers of that particular distribution.

iv) However the results from generation of data using logarithmic transformation was not that good compared to the original sequences. Therefore the generation of larger number of samples are required to arrive at a conclusion as to the better approach. The statistics of generated sequence of one sample depends upon the probability distribution fitted to the random component.

v) Both the correlograms and variance spectra of monthly rainfall and runoff sequences were useful and should be used simultaneously; while the correlograms show the physical cycles detectable in these series, the variance spectra show the number and significance of various harmonics to be used in the Fourier series description of periodic component of time series. Both of these techniques show the types of dependence models of stochastic component for both rainfall and runoff sequences. However for the calculation of the order of the dependence model, determination coefficient approach is required.

vi) This study indicated that when 12 month cycle and the significant subharmonics are removed from the monthly series, the remaining component may be considered approximately an independent stochastic process for monthly rainfall sequence and as an approximately a linear dependent stationary process for monthly runoff sequence. The subsurface water carryover in river basins from month to month is mainly responsible for the sequential dependence in monthly runoff sequences.

vii) Generally speaking the ratio of variance explained by the periodic component, to the total variance is higher for monthly rainfall sequence than that for monthly runoff sequence for the Lakshmanathirtha catchment.

IX. i) In the non-stationarity approach of Thomas Fiering Model, the generation of sequences by the logarithmic and square root transformation of the observed runoff sequences, gave better results when compared to no transformation case. However generation of more samples is necessary to arrive at a conclusion in this regard.

ii) The non-stationarity approach requires a larger number of parameters of monthly statistics, in comparison to stationarity approach.

III. i) In the stationarity approach bias is introduced as the information for all the months is used to generate each month flow values in the case of homogeneous process.

ii) The logarithmic transformation of the data record, gave better results in generation of sequences both from the stationarity and non-stationarity approaches. However evaluation of the statistics of two samples generated could not give a definite clue of superiority of either technique, this would require generation of large number of samples.

IV. From this study the following results are significant.

i) an improved understanding and mathematical description of hydrologic stochastic processes.

ii) development of improved methodology for the generation of new samples of hydrologic sequences and their use to practical problems.

8.2. SCOPE FOR FURTHER WORK.

i) the inclusion of the trend of the sampling variation of the data, in the generating process required further study.

ii) The structural analysis of monthly hydrologic sequence by fitting a periodic function for both monthly means and monthly standard deviations requires further investigation.

iii) The possible effect of removal of periodicity in parameters on the order of dependence model of the stochastic component needs further investigation.

iv) Further study is needed for using Thomas Fiering Model including skewness parameter also.

v) The stochastic simulation of monthly stream flows by multiple regression process utilising rainfall and runoff data together. needs study.

REFERENCES

1. Alven W.H.V. — 'Reliability Engineering' ARINC Research Corporation.

2. Anderson — 'Distribution of serial correlation coefficient' Ann. Math. Statist. 8 (1934).

3. Bartlett M.S. — 'Periodogram analysis and continuous spectra' Biometrika 37, 1950.

4. Benjamin J.R. and Cornell C.A. — 'Probability, statistics, and Decision for Civil Engineers' Mc Graw - Hill Book Company.

5. Blake L.S. — 'Civil Engineers Reference Book.

6. Chin W.Q. and Yevjevich V. — 'Almost-periodic, stochastic process of long-term climatic changes'. paper No. 65, Hydrology papers of Colorado State University U.S.A.

7. Chow V.T. — 'Stochastic Hydrologic systems' - proceedings of the first U.S. Japan seminar in Hydrology, Water Resources Publications, Colorado 1971.

8. Chow V.T. — 'Hand book of applied Hydrology'.

9. Chow V.T. and Kareliotis S.J. — 'Analysis of stochastic hydrologic systems' - Water Resources Research vol. 6 no. 6 Dec, 1970.

10. Clarke R.T. — 'Mathematical models in Hydrology' Food and Agricultural organization Rome, 1973.

11. Colston N.V. and Wiggert J.M. — ' A technique for generating synthetic flow record to estimate the variability of dependable flows for a fixed reservoir capacity', Water Resources Research vol. 6 no. 1 Feb, 1970.

12. Dandy F.E. — Comment on 'Analysis of stochastic hydrologic systems' by Chow and Kareliotis (Ref. No. 9 above) Water Resources Research vol. 8 no. 1 Feb. 72.

13. Fiering M.B. and Jackson B.B. — 'Synthetic streams flows' Monograph of American Geophysical Union. 1971.

14. Gafoor J.U. — 'Generation of random numbers of various distributions.' M.E. (Hydrology) Special Problem.

15. Cong S.P. 'Fortran IV and Engineering applications.'

16. Hurst H.E. 'Methods of using long term storage in reservior's Proceedings of Institution of Civil Engineering, 1956.

17. Julian P.R. 'Variance spectrum Analysis' Water Resources Research Vol. 3 No. 3, 1967

18. Kisiel C.C. 'Time series and analysis of Hydrologic Data' Advances in Hydroscience vol. 5 Ed. Chow, V.T.

19. Lehmer D.H. 'Mathematical methods in large-scale computing units. Ann. Comp. Lab. Harvard University, 1951

20. Markovic R.D. 'Probability functions of best fit to distributions of annual precipitation and runoff' Paper no. 8, Hydrology papers of Colorado State University.

21. Matalas N.C. 'Time series Analysis' Water Resources Research, Vol. 3, No. 3, 1967.

22. Matalas N.C. 'Mathematical assessment of synthetic Hydrology' Water Resources Research vol. 3, No. 4, 1967.

23. Milton A. and Stegun I.A. ' Handbook of Mathematical functions'.

24. Moreau D.H. and Pyatt E.E. 'Weekly and Monthly flows in Synthetic Hydrology' Water Resources Research vol6, No. 1, Feb.'70.

25. Pahuja S.L. 'A study of the generation of monthly stream flow sequences' M.E. (Hydrology) dissertation.

26. Quimpo R.G. and Yevjevich V. 'Stochastic description of daily river flows' - proceedings of first International Hydrology Symposium, Colorado State University 1967.

27. Quimpo R.G. 'Autocorrelation and spectral analysis in Hydrology' Journal of Hydraulics division ASCE, vol. 94 March 1968.

28. Raudkivi A.J. and Lawgun N. 'Simulation of Rainfall sequences' - Journal of Hydrology' vol. 22, No. 3/4, August 1974.

29. Rodriguez [illegible] — 'Adequacy of Markovian models with cyclic components for stochastic stream flow simulation' Water Resources Research Vol. 7 no. 5 Oct. 71.

30. Roesner L.A. and Yevjevich V. — 'Mathematical models for time series of monthly precipitation and monthly runoff' paper 15 Hydrology paper of CSU, U.S.A.

31. Sangal B.P. and Biswas A.K. — 'The 3 - parameters log-normal distribution and its application in Hydrology'- Water Resources Research, Vol. 6, No. 2, April 70.

32. Satish Chandra — Lecture notes on 'Stochastic Hydrology' at University of Roorkee, Roorkee.

33. Subin Pinkayan and [illegible] — ' A mathematical model of monthly river flow at Bhumiphol dam' - proceedings of symposium on water Resources, 1971. Indian Institute of Science, Bangalore.

34. Sun - Fu - Shih — ' Time series analysis and its applications to water Resources', proceedings of second world congress on water resources, New Delhi, 1975.

35. Thomas and Fiering — 'Mathematical Synthesis of stream flow sequences for the analysis of river basins by simulation' - Design of water resources system by Maass et. all Harvard University Press, Cambridge, 1962.

36. Wallis and Matalas — 'Correlogram analysis revisited' Water resources research vol. 7 No. 6, Dec. 71.

37. Yevjevich V. — 'Structural analysis of hydrologic time series' paper No. 56. Hydrology paper of Colorado State University, U.S.A.

38. Yevjevich, V. — 'Stochastic Processes in Hydrology' Water Resources Publications, Colorado, U.S.A.

39. Yevjevich, V. — 'Generation of Hydrologic samples' paper No. 72, Hydrology Paper of Colorado State University.

40. Yevjevich V. and Jeng R.I. — 'Effect of inconsistency and non-homogeneity on hydrologic time series'- proceedings of the International Hydrology Symposium, Colorado State University, 1967.

A N N E X U R E

COMPUTER PROGRAMS

INDEX TO COMPUTER PROGRAMS

```
      PROGRAM.NO.1
C     H.S.V.MURTHY STRUCTURAL ANALYSIS OF HYDROLOGIC TIME SERIES
      DIMENSION X(480),Y(480),SUM(10),A(10,10),AC(6),YC(500)
      DIMENSION GAMA (500)
      READ997,N1,M
      READ997,K1,K2
      READ998,(Y(J),J=1,M)
997   FORMAT(2I3)
998   FORMAT(6F10.1)
      DO 800 J=1,M
      X(J)=J
 800  CONTINUE
      DO99IK=2,N1
      N=IK
      NN=2*N-1
      DO6K=1,NN
      KK=K-1
      XK=KK
      SUM(K)=0.0
      DO6J=K1,K2
6     SUM(K)=SUM(K)+(X(J)**(XK))
      L=N+1
      DO18I=1,N
      A(I,L)=0.0
      XI=I-1
      DO9J=K1,K2
9     A(I,L)=A(I,L)+Y(J)*X(J)**XI
      DO18J=1,N
      JK=I+J-1
18    A(I,J)=SUM(JK)
      CALL SOLEQN (A,N,L)
      DO100I=1,N
      AC(I)=A(I,L)
100   CONTINUE
      SUM1=0.0
      SUM2=0.0
      SUM3=0.0
      GOTO(102,102,103,104,105),N
102   DO106J=K1,K2
      YC(J)=AC(1)+AC(2)*X(J)
106   CONTINUE
      GOTO110
103   DO107J=K1,K2
      YC(J)=AC(1)+(AC(2)*X (J))+(AC(3)*X(J)*X(J))
107   CONTINUE
      GOTO110
104   DO108J=K1,K2
      YC(J)=AC(1)+(AC(2)*X (J))+(AC(3)*X(J)*X(J))  +(AC(4)*(X(J)**3)  )
108   CONTINUE
      GOTO110
105   DO109J=K1,K2
      YC(J)=AC(1)+(AC(2)*X (J))+(AC(3)*X(J)*X(J))  +(AC(4)*(X(J)**3) )
     1+(AC(5)*(X(J)**4))
109   CONTINUE
110   DO112J=K1,K2
      GAMA(J)=((Y(J)-YC(J))**2)
      SUM1=SUM1+GAMA(J)
```

```
      SUM2=SUM2+Y(J)
      SUM3=SUM3+Y(J)*Y(J)
112   CONTINUE
      T=K2-K1+1
      VAR=SUM3-SUM2*SUM2/T
      A1=K2-K1
      VAR=VAR/A1
      CORI=SQRTF(1.0-SUM1/(T*VAR))
      PUNCH113,CORI
113   FORMAT(4HCORI,F8.4)
99    CONTINUE
      STOP
      END
      SUB PROGRAM.NO.1.1
C     H.S.V.MURTHY STRUCTURAL ANALYSIS OF HYDROLOGIC TIME SERIES
      SUBROUTINE SOLEQN(A,N,L)
      DIMENSION A(10,10)
      DO25 K=1,N
      IF(A(K,K))20,50,20
20    DO25I=1,N
      IF(I-K)8,25,8
8     FACT=A(I,K)/A(K,K)
      DO30J=1,L
30    A(I,J)=A(I,J)-A(K,J)*FACT
25    CONTINUE
      DO60I=1,N
60    A(I,L)=A(I,L)/A(I,I)
      PUNCH200,(A(I,L),I=1,N)
200   FORMAT(5F10.5)
50    RETURN
      END


      PROGRAM.NO.2
C     CALCULATION OF WEIGHTED AVERAGE RAINFALL BY THIESSEN POLYGON
C     H.S.V.MURTHY STRUCTURAL ANALYSIS OF HYDROLOGIC TIME SERIES
      DIMENSIONQ(342,12),X(38,12),AC(10)
      READ101,N
101   FORMAT(I3)
      READ 102,((Q(I,J),J=1,12),I=1,N)
102   FORMAT(12F6.2)
      DO802J=1,12
      DO802I=1,38
      KK=0
      KK=KK+I
      DO100K=1,9
      AC(K)=Q(KK,J)
      KK=KK+38
100   CONTINUE
      X(I,J)=0.19830*AC(1)+0.10170*AC(2)+0.04225*AC(3)+
     10.07759*AC(4)+0.07114*AC(5)+0.09507*AC(6)+0.05535*AC(7)+
     20.14330*AC(8)+0.20930*AC(9)
 802  CONTINUE
      PUNCH 103,((X(I,J),J=1,12),I=1,38)
103   FORMAT(6F12.4)
      STOP
      END
```

```
      PROGRAM.NO. 3
      DIMENSION Q(40,13),AV(13),STD(13),B(13),R(13),CV(13),CS(13),CK(13)
      DIMENSION Y(500),SCC(100),CC (100),AP(8),BP(8),C(8),AA(8)
      DIMENSION BB(8),F(100),D(100,,G(100),Y1(500),Y2(500)
      DIMENSION AZ(15),BZ(15)
C     H.S.V.MURTHY  DISSERTATION WORK  TIME SERIES ANALYSIS HYDROLOGY
1     MIT=0
636   MIT=MIT+1
      READ101,N,N1,K1,K2,M,NLOG
      READ101,KOT,MK,MMK,NST,IIT
101   FORMAT(6I3)
      READ102,(Y(K),K=1,N1)
102   FORMAT(6F12.4)
      PRINT145,N,N1,K1,K2,M,NLOG
      PRINT145,KOT,MK,MMK,NST
145   FORMAT(1X,6I3)
      PRINT103,(Y(K),K=1,N1)
103   FORMAT(1X,6F12.4)
      GOTO(378,379,120),IIT
378   CALL  CORG(Y,N1,K1,K2,M,SCC)
      CALL CAPR(SCC,ARC1,ARC2,AR31,AR32,AR33)
      CALL  SPEM(SCC,KOT,MK,MMK)
379   LL=1
      CALL  STAT(Y,LL,N1,AVV,STDS,CSS)
      CALL SSDS(Y,N1,Q,N)
      CALL      STCS (Q,N,AV,STD,B,R,CV,CS,CK)
      CALL  PERD(Q,N,AP,BP)
      DO114JK=1,6
      AA(JK)=AP(JK)
      BB(JK)=BP(JK)
114   CONTINUE
      DO83K=1,N1
      Y1(K)=K
83    CONTINUE
      SUMX=0.0
      SUMY=0.0
      SUMXX=0.0
      SUMXY=0.0
      DO84K=1,N1
      SUMX=SUMX+Y1(K)
      SUMY=SUMY+Y(K)
      X8=Y1(K)*Y1(K)
      XY=Y1(K)*Y(K)
      SUMXY=SUMXY+XY
      SUMXX=SUMXX+X8
 84   CONTINUE
      AN=N1
      AM=AN
      DENOM=AN*SUMXX-SUMX*SUMX
      A6=(SUMY*SUMXX-SUMX*SUMXY)/DENOM
      B6=(AN*SUMXY-SUMX*SUMY)/DENOM
      PRINT85,A6,B6,DENOM
85    FORMAT(1X,3F20.4)
      AV1=SUMX/AM
      AV2=SUMY/AM
      SUM=0.0
      SUMB=0.0
      DO36K=1,N1
      Y2(K)=Y(K)-(A6+B6*Y1(K))
      Y1(K)=Y1(K)-AV1
```

```
      SUM8=SUM8+Y1(K)*Y1(K)
      SUM=SUM+Y2(K)*Y2(K)
86    CONTINUE
      PRINT103,(Y2(K),K=1,N1)
      SYSQ=SUM/(AM-2.0)
      SB=SQRT (SYSQ/SUM8)
      T6=B6/SB
      PRINT65,SB,T6,SYSQ
      SYBR=SQRT (SYSQ/AM)
      CL1=AV2+1.96*SYBR
      CL2=AV2-1.96*SYBR
      PRINT85,CL1,CL2,SYBR
      IF(T6-1.96)87,87,88
88    PRINT94
94    FORMAT(20HLINEAR TREND REMOVED)
      IF(NST-1)130,87,130
130   IF(NLOG-1)96,87,96
87    DO95K=1,N1
      Y2(K)=Y(K)
95    CONTINUE
      IF(NST-1)96,125,96
125   CALL SSDS(Y2,N1,Q,N)
      CALL       STCS (Q,N,AV,STD,R,R,CV,CSJCK)
      DO127I=1,N
      DO127J=1,12
      Q(I,J)=(Q(I,J)-AV(J))/STD(J)
127   CONTINUE
      PRINT160
160   FORMAT(5X,19HSTANDARDISED SERIES)
      PRINT103,((Q(I,J),J=1,12),I=1,N)
      IF(NST-1)129,128,129
96    IF(NLOG-1)120,120,121
120   DO122K=1,N1
      Y2(K)=Y(K)
      Y2(K)=ALOG(Y2(K))
122   CONTINUE
      PRINT161
161   FORMAT(5X,22HLOG TRANSFORMED SERIES)
      PRINT103,(Y2(K),K=1,N1)
      CALL  SSDS(Y2,N1,Q,N)
      CALL  PERD (Q,N,AP,BP)
      CALL CORG(Y2,K1,K2,M,SCC)
      CALL CAPR(SCC,ARC1,ARC2,AR31,AR32,AR33)
      CALL  SPEM(SCC,KOT,MK,MMK)
      DO1500JK=1,6
      AA(JK)=AP(JK)
      BB(JK)=BP(JK)
1500  CONTINUE
121   LL=2
      CALL STAT(Y2,LL,N1,AVV,STDS,SS)
      AXV=AVV
      DO124K=1,N1
      Y2(K)=Y2(K)-AXV
124   CONTINUE
```

```
      PRINT162
162   FORMAT(5X,22HMEAN SUBTRACTED SERIES)
      PRINT103,(Y2(K),K=1,N1)
126   LL=2
      CALL  SSDS(Y2,N1,Q,N)
      CALL STAT(Y2,LL,N1,AVV,STDS,CSS)
      CALL PERD(Q,N,AP,BP)
      CALL  CORG(Y2,N1,K1,K2,M,SCC)
      CALL CAPR(SCC,ARC1,ARC2,AR31,AR32,AR33)
      CALL  SPEM(SCC,KOT,MK,MMK)
      AM=12.0
      DO20JK=1,6
      AJ=JK
      AXX=2.0*3.1416*AJ/AM
      DO93IK=1,12
      AIK=IK
      AZ(IK)=AA(JK)*COS(AXX*AIK)
      BZ(IK)=BB(JK)*SIN(AXX*AIK)
93    CONTINUE
      DO302I=1,N
      DO302J=1,12
302   Q(I,J)=Q(I,J)-AZ(J)-BZ(J)
      PRINT145,JK
      PRINT103,((Q(I,J),J=1,12),I=1,N)
128   CALLDSSS(Q,N,Y2,N1)
      LL=2
      CALL STAT(Y2,LL,N1,AVV,STDS,CSS)
      CALL PERD(Q,N,AP,BP)
      CALL  CORG(Y2,N1,K1,K2,M,SCC)
      CALL CAPR(SCC,ARC1,ARC2,AR31,AR32,AR33)
      CALL  SPEM(SCC,KOT,MK,MMK)
20    CONTINUE
129   CONTINUE
      GOTO(605,605,605,608),NIT
605   GOTO636
608   STOP
      END
```

SUB PROGRAM NO. 3.1

```
      SUBROUTINE  CORR(Y,N1,K1,K2,M,SCC)
      DIMENSION Y(1250),COV(100),SCC(100)
C     H.S.V.MURTHY  DISSERTATION WORK  TIME SERIES ANALYSIS HYDROLOGY
      DO10K=K1,M,K2
      XZ=0.0
      X1=0.0
      X2=0.0
      X3=0.0
      X4=0.0
      NK=N1-K
      DO28I=1,NK
      J=I+K
      XZ=XZ+Y(I)
      X1=X1+Y(J)
      X2=X2+Y(I)*Y(I)
      X3=X3+Y(J)*Y(J)
      X4=X4+Y(I)*Y(J)
28    CONTINUE
      ANK=NK
      A3=SQRT (ANK*X2-XZ*XZ)
      B3=SQRT (ANK*X3-X1*X1)
      COV(K)=(ANK*X4-XZ*X1)
      SCC(K)=COV(K)/(A3*B3)
      COV(K)=COV(K)/(ANK*ANK)
10    CONTINUE
      PRINT850
850   FORMAT(1X,4HLAGK,15X,7HSERCOEF,7X,5HCOVAR)
      DO3000K=K1,M,K2
      PRINT901,K,SCC(K),COV(K)
901   FORMAT(1X,I5,10X,F12.6,F20.6)
3000  CONTINUE
      RETURN
      END
```

SUB PROGRAM NO. 3.2

```
      SUBROUTINESSDS(Y,N1,Q,N)
      DIMENSION  Y(500),Q(40,13)
C     H.S.V.MURTHY  DISSERTATION WORK  TIME SERIES ANALYSIS HYDROLOGY
      K=0
      DO105I=1,N
      DO105J=1,12
      K=K+1
      Q(I,J)=Y(K)
105   CONTINUE
      RETURN
      END
```

```
      SUB PROGRAM NO. 3.3
      SUBROUTINE  CAPR(SCC,ARC1,ARC2,AR31,AR32,AR33)
      DIMENSION SCC(100)
C     H.S.V.MURTHY  DISSERTATION WORK  TIME SERIES ANALYSIS HYDROLOGY
      U=SCC(1)
      V=SCC(2)
      W=SCC(3)
      U2=U*U
      U3=U2*U
      U4=U3*U
      V2=V*V
      V3=V2*V
      V4=V3*V
      W2=W*W
      RSQ1=U2
      RSQ2=U2+V2-2.0*U2*V
      RSQ2=RSQ2/(1.0-U2)
      X1=2.0*U3*W+2.0*U2*V2+2.0*U*V2*W-2.0*U2*V-4.0*U*V*W-U4-V4
      X1=X1-U2*W2+U2+V2+W2
      X2=1.0-2.0*U2-V2+2.0*U2*V
      RSQ3=X1/X2
      X3=U-U*V
      X4=1.0-U2
      X5=V-U2
      ARC1=X3/X4
      ARC2=X5/X4
      PRINT145
145   FORMAT(2X,24HRSQ1*RSQ2*RSQ3*ARC1*ARC2)
      PRINT146,RSQ1,RSQ2,RSQ3,ARC1,ARC2
146   FORMAT(2X,5F15.4)
      X6=(1.0-U2)*(U-W)
      X7=(1.0-V)*(U*V-W)
      X8=(1.0-V)*(1.0-2.0*U2+V)
      AR31=(X6-X7)/X8
      AR32=(1.0-V)*(V+V2-U2-U*W)/X8
      AR33=((U-W)*(U2-V)-(1.0-V)*(U*V-W))/X8
      PRINT147
147   FORMAT(2X,14HAR31*AR32*AR33)
      PRINT146,AR31,AR32,AR33
      RETURN
      END
```

```
      SUB PROGRAM NO. 3.4
      SUBROUTINE SPEM (SCC,KOT,MK,MMK)
      DIMENSION SCC(100),F(100),D(100),G(100)
C     H.S.V.MURTHY  DISSERTATION WORK  TIME SERIES ANALYSIS HYDROLOGY
70    READ5,K1,M,K2
5     FORMAT(3I3)
      PRINT145,K1,M,K2
145   FORMAT(5X,3I3)
      KT=0
      JK=0
      KOT=KOT+1
      MX=M+1
      DO10I=K1,MX,K2
      JK=JK+1
      AM=M
      AI=I-1
  10  F(JK)=AI/(2.0*AM)
50    KT=KT+1
      GOTO(51,22,23,71),KT
51    NVJ=2
      DO493K=1,M
      AK=K
493   D(K)=(1.0+COS(3.1416*AK/AM)/2.0
      GOTO100
22    NVJ=2
      DO12K=1,M
      AK=K
12    D(K)=1.0-(AK/AM)
      GOTO100
23    NVJ=2
      DO13K=1,M
      AK=K
      JM=M/2
      AAK=AK/AM
      IF(K-JM)24,24,25
24    D(K)=1.0-6.0*(AAK**2.0)+6.0*(AAK**3.0)
      GOTO26
25    D(K)=2.0*((1.0-AAK)**3.0)
26    CONTINUE
13    CONTINUE
      GOTO100
71    NVJ=2
      DO72K=1,M
      AK=K
72    D(K)=0.54+0.46*COS (3.1416*AK/AM)
100   JK=0
      PRINT40,KT
      SUM33=0.0
      DO31I=K1,MX,K2
      JK=JK+1
      SUM=0.0
      FR=F(JK)
      DO32K=1,M
      AK=K
      COST=COS(2.0*3.1416*FR*AK)
      IF(NVJ-1)400,300,400
300   SUM=SUM+((1.0+COST)/2.0)*SCC(K)*COST
      GOTO32
400   SUM=SUM+D(K)*SCC(K)*COST
32    CONTINUE
```

```
      G(JK)=2.0*(1.0+2.0*SUM)
      SUM33=SUM33+G(JK)
31    CONTINUE
40    FORMAT(1X,I5)
      JJK=0.0
      PRINT110
110   FORMAT(2X,9HFREQUENCY,16X,10HSMOOTHSPFN)
      DO301J=K1,MX,K2
      JJK=JJK+1
      PRINT41,F(JJK),G(JJK)
301   CONTINUE
      PRINT112,SUM33
112   FORMAT(2X,6HSUM33=,F16.6)
      PRINT61,KT,MK,K1,MX,K2
61    FORMAT(5X,5I4)
41    FORMAT(2X,2F16.8)
      IF(KT-MK)50,50,60
 60   IF(KOT-MMK)70,70,80
80    KOT=0
      RETURN
      END
```

```
      SUB PROGRAM NO. 3.5
      SUBROUTINE STAT(Y,LL,N1,AVV,STDS,CSS)
      DIMENSION Y(500)
C     H.S.V.MURTHY  DISSERTATION WORK  TIME SERIES ANALYSIS HYDROLOGY
      AM=N1
      T=AM/((AM-1.0)*(AM-2.0))
      AVV=0.0
      STDS=0.0
      VAR=0.0
      RAN=0.0
      CVV=0.0
      CSS=0.0
      DO500K=1,N1
      AVV=AVV+Y(K)
      STDS=STDS+Y(K)*Y(K)
      CSS=CSS+(Y(K)**3)
500   CONTINUE
      SUM=STDS
      AVV=AVV/AM
      STDS=SQRT ((STDS-AM*AVV*AVV)/AM)
      CSS=(CSS-(3.*AVV*SUM)+(AVV**3)*2.*AM)*T/(STDS**3)
      CVV=STDS*100./AVV
      VAR=STDS**2
      GOTO(524,525),LL
524   SUM6=0.0
      RL1=0.0
      RL2=0.0
      DO522K=1,N1
      SUM6=SUM6+(Y(K)-AVV)
      IF(SUM6-RL1)520,521,521
520   RL1=SUM6
      GOTO522
521   IF(SUM6-RL2)522,522,523
523   RL2=SUM6
522   CONTINUE
      RANGE=RL2-RL1
      PRINT510,RANGE
510   FORMAT(1X,5HRANGE,F20.5)
525   PRINT576
576   FORMAT(10X,20HAVV*STDS*VAR*CVV*CSS)
      PRINT575,AVV,STDS,VAR,CVV,CSS
575   FORMAT(1X,3F18.4,2F9.4)
      RETURN
      END
```

```
      SUB PROGRAM NO. 3.6
      SUBROUTINE STCS (Q,N,AV,STD,B,R,CV,CS,CK)
      DIMENSION Q(40,13),AV(13),STD(13),B(13),R(13),CV(13),CS(13),CK(13)
C     H.S.V.MURTHY  DISSERTATION WORK  TIME SERIES ANALYSIS HYDROLOGY
      AM=N
      T=AM/((AM-1.0)*(AM-2.0))
      DO 501 J=1,12
      AV(J)=0.0
      CS(J)=0.0
      CK(J)=0.0
      STD(J)=0.0
      DO 502 I=1,N
      AV(J)=AV(J)+Q(I,J)
      SUM4=AV(J)
      CS(J)=CS(J)+(Q(I,J)**3.0)
      SUM3=CS(J)
      CK(J)=CK(J)+(Q(I,J)**4.0)
  502 STD(J)=STD(J)+Q(I,J)*Q(I,J)
      AV(J)=AV(J)/AM
      SUM2=STD(J)
      SUM1=AV(J)
      STD(J)=SQRT ((STD(J)-AM*AV(J)*AV(J))/AM)
      CS(J)=(CS(J)-(3.*AV(J)*SUM2)+(AV(J)**3)*(3.*AM- AM))*T/(STD(J)**3)
      CK(J)=(CK(J)-(4.0*SUM1*SUM3)+(6.0*(SUM1**2)*SUM2)
     1-(4.0*(SUM1**3)*SUM4)+(AM*(SUM1**4)))/(AM*(STD(J)**4))
501   CV(J)=(STD(J)*100.0)/AV(J)
      AV(13)=AV(1)
      STD(13)=STD(1)
      NM=N-1
      DO504I=1,NM
      I1=I+1
  504 Q(I,13)=Q(I1,1)
      DO505J=1,11
      R(J)=0.0
      DO 506 I=1,N
  506 R(J)=R(J)+Q(I,J)*Q(I,J+1)
505   R(J)=((R(J)/AM)-AV(J)*AV(J+1 ))/(STD(J)*STD(J+1))
      R(13)=R(1)
      R(12)=0.00
      ANM=NM
      DO508I=1,NM
      I1=I+1
  508 R(12)=R(12)+Q(I,12)*Q(I1,1)
      R(12)=((R(12)/ANM)-(AV(12)*AV(13)) )/(STD(12)*STD(13))
      DO 507 J=1,12
  507 B(J) =R(J)*STD(J+1)/STD(J)
      PRINT600
600   FORMAT(10X,2HAV,8X,2HSD,10X,1HR,9X,2HBJ,9X,2HCS,5X,2HCK,5X,2HCV)
      PRINT601,(J,AV(J),STD(J),R(J),B(J),CS(J),CK(J),CV(J) ,J=1,12)
601   FORMAT(1X,I2,2F12.5,4F10.5,F8.1)
      RETURN
      END
```

```
      SUB PROGRAM NO. 3.7 AND 3.8
      SUBROUTINE PERD(Q,N,AP,BP)
      DIMENSION Q(40,13),AP(8),BP(8),C(8)
C     H.S.V.MURTHY  DISSERTATION WORK  TIME SERIES ANALYSIS HYDROLOGY
      AM=12.0
      DO48JK=1,6
      AJ=JK
      AX=2.0*3.1415*AJ/AM
      SUM11=0.0
      SUM22=0.0
      DO21K=1,N
      DO21IK=1,12
      AIK=IK
      SUM11=SUM11+Q(K,IK)*COS(AX*AIK)
      SUM22=SUM22+Q(K,IK)*SIN(AX*AIK)
21    CONTINUE
      AN=N
      AP(JK)=SUM11/(6.0*AN)
      BP(JK)=SUM22/(6.0*AN)
      C(JK)=(AP(JK)*AP(JK)+BP(JK)*BP(JK))/2.0
48    CONTINUE
      PRINT49
49    FORMAT(5X,6HAP(JK),14X,6HBP(JK),14X,5HC(JK))
      PRINT30,(AP(JK),BP(JK),C(JK),JK,JK=1,6)
30    FORMAT(1X,3F20.8,5X,I5)
      RETURN
      END
      SUBROUTINE DS35(Q,N,Y,N1)
      DIMENSION Q(40,13),Y(500)
C     H.S.V.MURTHY  DISSERTATION WORK  TIME SERIES ANALYSIS HYDROLOGY
      K=0
      DO106I=1,N
      DO106J=1,12
      K=K+1
      Y(K)=Q(I,J)
106   CONTINUE
      RETURN
      END
```

PROGRAM.NO. 4

```
C   C   SUBROUTINE PERO(Q,N,AP,DP)
        DIMENSION Q(10,13),AP(8),BP(8),C(8) ,DP(8)
C       H.S.V.MURTHY  DISSERTATION WORK  TIME SERIES ANALYSIS HYDROLOGY
        READ101,N
101     FORMAT(I2)
10      READ900,((Q(I,J),J=1,12),I=1,N)
900     FORMAT(3X,F12.5)
1012    PUNCH900,((Q(I,J),J=1,12),I=1,N)
1000    AM=12.0
        DO48JK=1,6
        AJ=JK
        AX=2.0*3.1416*AJ/AM
        SUM11=0.0      $  SUM22=0.0
        DO21K=1,N
        DO21IK=1,12
        AIK=IK
        SUM11=SUM11+Q(K,IK)*COSF(AX*AIK)
        SUM22=SUM22+Q(K,IK)*SINF(AX*AIK)
21      CONTINUE
        AN=N
        AP(JK)=SUM11/(6.0*AN)
        BP(JK)=SUM22/(6.0*AN)
        C(JK)=(AP(JK)*AP(JK)+BP(JK)*BP(JK))/2.0
48      CONTINUE
        PUNCH49
49      FORMAT(5X,6HAP(JK),14X,6HBP(JK),14X,5HC(JK))
        PUNCH30,(AP(JK),BP(JK),C(JK),JK,JK=1,6)
30      FORMAT(1X,3F20.8,5X,I5)
        AV=0.0   $  VAR=0.0
        DO302I=1,N
        DO302J=1,12
        AV=AV+Q(I,J)
        VAR=VAR+((Q(I,J)**2))
302     CONTINUE
        AV=AV/12.
        VAR=(VAR-12.0*AV*AV)/11.0
        DO402JK=1,6
        DP(JK)=C(JK)/VAR
402     CONTINUE
        PUNCH30,(DP(JK),BP(JK),C(JK),JK,JK=1,6)
        PUNCH32,AV,VAR
32      FORMAT(2X,3HAV=,F16.3,4HVAR=,F20.3)
        READ999,((Q(I,J),J=1,12),I=1,N)
999     FORMAT(15X,F12.5)
        GOTO1012
        STOP
        END
```

```
      PROGRAM.NO.5
      TO OBTAIN RANDOM COMPONENT
C     H.S.V.MURTHY STRUCTURAL ANALYSIS OF HYDROLOGIC TIME SERIES
      DIMENSION Y(500),SCC(100),COV(100),Y1(500)
      READ101,N1,K1,K2,M,MIT,MT
      READ101,MII
101   FORMAT(6I3)
      READ125,AV1,STAD,AV2,STAND
      READ125,AK1,AK2,AK3,AK4
      READ125,AK5,AK6,AK7,AK8
      READ125,AK9,AK10,AK11,AK12
125   FORMAT(4F15.6)
C     STAND=ORIGINAL SERIES STD.DEV.
      READ102,(Y(K),K=1,N1)
102   FORMAT(6F12.4)
      GOTO(132,133),MT
132   DO83K=1,N1
      AK=K
      PT1=AK1*COS(2.0*3.1416*AK/12.0)+AK2*SIN(2.0*3.1416*AK/12.0)
      PT2=AK3*COS(4.0*3.1416*AK/12.0)+AK4*SIN(4.0*3.1416*AK/12.0)
      GOTO(120,83),MII
120   PT3=AK5*COS(6.0*3.1416*AK/12.0)+AK6*SIN(6.0*3.1416*AK/12.0)
      PT4=AK7*COS(8.0*3.1416*AK/12.0)+AK8*SIN(8.0*3.1416*AK/12.0)
      PT5=AK9*COS(10.0*3.1416*AK/12.0)+AK10*SIN(10.0*3.1416*AK/12.0)
      COST=AV1+PT1+PT2+PT3+PT4+PT5
      Y(K)=(Y(K)-COST)/STAND
83    CONTINUE
      PUNCH102,(Y(K),K=1,N1)
      DO84K=1,N1
      Y(K)=(Y(K)-AV2)/STAD
84    CONTINUE
      PUNCH102,(Y(K),K=1,N1)
133   CALL CORC(Y,N1,K1,K2,M,SCC)
      CALL  CAPR (SCC,ARC1,ARC2,AR31,AR32,AR33)
      U1=ARC1  $  U2=ARC2
      U=AR31     $   V=AR32
      W=AR33
      S1=SCC(1)      $   S2=SCC(2)
86    MIT=MIT+1
      GOTO(70,71,72,73),MIT
70    DO90K=2,N1
      KK=K-1
      AN1=Y(K)-S1*Y(KK)
      AN2=SQRT(1.0-S1*S1)
      Y1(K)=AN1/AN2
90    CONTINUE
      Y1(1)=Y(1)/AN2
      GOTO98
71    DO91K=3,N1
      KK=K-1
      IK=K-2
      AN1=Y(K)-U1*Y(KK)-U2*Y(IK)
      AN2=U1*U1+U2*U2+2.0*U1*U2*S1
      AN2=SQRT(1.0-AN2)
      Y1(K)=AN1/AN2
```

```
91 CONTINUE
   Y1(1)=Y(1)/AN2
   Y1(2)=Y(2)/AN2
   GOTO98
72 DO85K=4,N1
   KK=K-1
   IK=K-2
   KI=K-3
   AN1=Y(K)-U*Y(KK)-V*Y(IK)-W*Y(KI)
   AN2=U*U+V*V+W*W+2.0*(U*V*S1+U*W*S2+V*W*S1)
   AN2=SQRT(1.0-AN2)
   Y1(K)=AN1/AN2
85 CONTINUE
   Y1(1)=Y(1)/AN2
   Y1(2)=Y(2)/AN2
   Y1(3)=Y(3)/AN2
98 PUNCH102,(Y1(K),I=1,N1)
   LL=2
   CALL  STAT(Y1,LL,N1,AVV,SDS,CSS)
   N5=N1-1
   DO10K=1,N5
   N2=K+1
   DO10J=N2,N1
   IF(Y1(K)-Y1(J))4,4,10
4  SUM=Y1(K)
   Y1(K)=Y1(J)
   Y1(J)=SUM
10 CONTINUE
   PUNCH102,(Y1(K),K=1,N1)
   GOTO86
75 STOP
   END
```

```
      PROGRAM.NO.6
      PARAMETERS FOR LN-3 DIST.
C     H.S.V.MURTHY STRUCTURAL ANALYSIS OF HYDROLOGIC TIME SERIES
      DIMENSION Y(500), UJ(65),UJJ(65)
      READ5,N1
5     FORMAT(I3)
      READ1012,(UJ(I),I=1,60)
1012  FORMAT(16F5.2)
1     READ35,PMEN,PSTD
35    FORMAT(2F15.4)
      READ900,(Y(I),I=1,N1)
900   FORMAT(6F12.4)
      AM=N1
      AVV=0.0
      DO45K=182,272
      AVV=AVV+Y(K)
45    CONTINUE
      AVV=AVV/91.
      PUNCH35,AVV
      ALL=PSTD*PSTD/(2.0*(PMEN-AVV))
      ALL=AVV-ALL
      PUNCH35,ALL
      DO1312K=1,N1
      Y(K)=Y(K)-ALL
 1312 CONTINUE
      LL=2
      CALL  STAT(Y,LL,N1,AVV,STDS,CSS)
      DO1412K=1,N1
      Y(K)=LOGF(Y(K))
1412  CONTINUE
      PUNCH900,(Y(I),I=1,N1)
      CALL  STAT(Y,LL,N1,AVV,STDS,CSS)
      DO44I=1,59
      UJJ(I)=AVV+STDS*UJ(I)
 44   CONTINUE
      PUNCH35,(UJ(I),UJJ(I),I=1,59)
      STOP
      END
```

```
      PROGRAM.NO.7
C     H.S.V. MURTHY STRUCTURAL ANALYSIS OF HYDROLOGIC TIME SERIES
      K-S TEST FOR GOODNESS OF FIT
      DIMENSION Y(500),F1(500),Y1(500),Y2(500)
      READ101,N1,K1,K2
101   FORMAT(4I3)
      READ103,ALL,PMEN,PSTD
      READ103,(Y(K),K=1,N1)
103   FORMAT(6F12.4)
      N11=N1/2
      DO4K=1,N11
      II=K-1
      II=N1-II
      C=Y(K)
      Y(K)=Y(II)
      Y(II)=C
4     CONTINUE
      DO500K=1,N1
      Y(K)=Y(K)-ALL
      Y2(K)=LOGF(Y(K))
500   CONTINUE
      LL=2
      CALL  STAT(Y,LL,N1,AVV,STDS,CSS)
      CALL STAT(Y2,LL,N1,AVV,STDS,CSS)
      A5=1.0/SQRTF(2.0*3.1416)
      DO1312K=K1,K2
      Y2(K)=(Y2(K)-AVV)/STDS
      X=Y2(K)
      FIRST=X
      N=1
      NF1=1
      C=-1.0
6     NN=2*N+1
      NF1=NF1*N
      NDEN=(2**N)*NF1*NN
      XX=X**NN
      DEN=NDEN
      TERM=(C*XX)/DEN
      TERM=TERM*A5
      FIRST=FIRST+TERM
      TERM=ABSF(TERM)
      C=-C
      IF(TERM-0.0001)61,60,60
60    N=N+1
      GOTO6
61    F1(K)=FIRST+0.5
1312  CONTINUE
      DO25K=K1,K2
      Y1(K)=K/N1
      Y(K)=Y1(K)-F1(K)
25    CONTINUE
      PUNCH103,(Y(K),K=K1,K2)
      PUNCH110,(K,Y2(K),F1(K),Y(K),K=K1,K2)
110   FORMAT(2(I3,3F10.4))
      STOP
      END
```

```
      PROGRAM.NO.8
      DIMENSION T(12),ZITA(2), ZT(695)
      READ10,M,NG
10    FORMAT(2I10)
      READ38,XO
38    FORMAT(F10.2)
      A=2.0** 9+3.0
      AM=2.** 18
      AMR=1.00/AM
      AAM=A*AM
      AXO=A*XO
      NGM=NG*M
      NG2=NGM/2
      PI=3.1415936
      NG1=0
      K2=0
108   NG1=NG1+1
      DO107I=1,M
      DO106J=1,2
      AAN=AXO*AMR
C     NAN=AAN
C     ANA=NAN
C     ZITA(J)=AAN-ANA
      AAN1=AAN*.0001
      NAN1=AAN1
      ANA1=NAN1
      ANA1=ANA1*10000.0
      DIF=AAN-ANA1
      IDIF=DIF     $AIDIF=IDIF
      ZITA(J)=DIF-AIDIF
      AXO=ZITA(J)*AAM
106   CONTINUE
      IF (K2-NGM)110,110,111
110   K1=K2+1
      K2=K1+1
      ZT(K1)=ZITA(1)
      ZT(K2)=ZITA(2)
C     TRANSORMS INTO NORMAL RANDOM NOS. BY BOS AND MULLER METHOD
111   U1=ZITA(1)
      U2=ZITA(2)
      ARGNT=2.0*PI*U2
      T(I)=(-2.0*LOGF(U1))**0.5*COSF(ARGNT)
107   CONTINUE
      PUNCH20,(T(I),I=1,M)
20    FORMAT(6F12.6)
      NGM1=NG1*M
      IF(NGM1-NGM)108,109,109
109   PUNCH30,AXO
30    FORMAT(E16.8)
      PUNCH 20,(ZT(I), I=1,NGM  )
      STOP
      END
```

```
      PROGRAM.NO.9
C     H.S.V.MURTHY STRUCTURAL ANALYSIS OF HYDROLOGIC TIME SERIES
      DIMENSION Y(550),SCC(100),COV(100) ,Y1(550)
C     H.S.V.MURTHY  DISSERTATION WORK  TIME SERIES ANALYSIS HYDROLOGY
      READ101,N1,K1,K2,M
101   FORMAT(4I3)
      READ103,(Y(K),K=1,N1)
      READ103,(Y1(K),K=1,N1)
103   FORMAT(6F12.4)
      PUNCH850
850   FORMAT(1X,4HLAGK,15X,7HSERCOFF,7X,5HCOVAR)
      DO10K=K1,M,K2
      KA=K-25
      XZ=0.0
      X1=0.0
      X2=0.0
      X3=0.0
      X4=0.0
      NK=N1-48
      DO28I=1,NK
      IA=I+24
      JA=24+I+KA
      XZ=XZ+Y(IA)
      X1=X1+Y1(JA)
      X2=X2+Y(IA)*Y(IA)
      X3=X3+Y1(JA)*Y1(JA)
      X4=X4+Y(IA)*Y1(JA)
28    CONTINUE
      ANK=NK
      A3=SQRT (ANK*X2-XZ*XZ)
      B3=SQRT (ANK*X3-X1*X1)
      COV(K)=(ANK*X4-XZ*X1)
      SCC(K)=COV(K)/(A3*B3)
      COV(K)=COV(K)/(ANK*ANK)
      PUNCH901,K,SCC(K),COV(K),IA
901   FORMAT(1X,I5,10X,F12.6,F20.6,I5)
10    CONTINUE
      STOP
      END
```

```
      PROGRAM.NO.10
C     TIME SERIES MODEL FOR GENERATION OF DATA
C     H.S.V.MURTHY STRUCTURAL ANALYSIS OF HYDROLOGIC TIME SERIES
      DIMENSIONQ(100,13),AV(13),STD(13),B(13),R(13),CV(13),CS(13),CK(13)
      DIMENSION EPSL(1250),Y(1250),RAND(1250),PCOM(1250)
      DIMENSION ST(1250),AT(13),GS(13),SCC(100),COV(100)
C     H.S.V.MURTHY  DISSERTATION WORK  TIME SERIES ANALYSIS HYDROLOGY
      READ101,M1,K1,K2,M,N,N1
101   FORMAT(6I3)
      L=M1*12
      READ103,(EPSL(I),I=1,L)
103   FORMAT (6F12.6)
      LL=2
      CALL  STAT(EPSL,LL,L,AVV,STDS,CSS)
      READ104,(AT(J),J=1,12)
      READ104,(GS(J),J=1,12)
104   FORMAT(6F10.4)
108   READ101,IIT,MIT,MMT,MII,IMT,NVJ
      READ104,ZHI,ZMU,RHON,AV2,STPS,AV1
      READ104,AA1,T1,T2,T3,TR,TRR
      READ104,A21,A22,R21,AK1,AK2,AK3
      READ104,AK4,AK5,AK6,AK7,AK8
      PRINT101,IIT,MIT,MII,MMT,IMT,L
      GOTO(106,107),IIT
106   DO84K=1,L
      RAND(K)=ZHI+EXP(ZMU+EPSL(K)*RHON)
84    CONTINUE
      GOTO110
107   DO109K=1,L
      RAND(K)=ZMU+RHON*EPSL(K)
109   CONTINUE
110   GOTO(112,113),MIT
112   DO83K=1,L
      AK=K
      PT1=AK1*COS(2.0*3.1416*AK/12.0)+AK2*SIN(2.0*3.1416*AK/12.0)
      PT2=AK3*COS(4.0*3.1416*AK/12.0)+AK4*SIN(4.0*3.1416*AK/12.0)
      PCOM(K)=AV1+PT1+PT2
      GOTO(120,83),MII
120   PT3=AK5*COS(6.0*3.1416*AK/12.0)+AK6*SIN(6.0*3.1416*AK/12.0)
      PT4=AK7*COS(8.0*3.1416*AK/12.0)+AK8*SIN(8.0*3.1416*AK/12.0)
 9    PCOM(K)=PCOM(K)+PT3+PT4
83    CONTINUE
113   PMEAN=0.0
114   GOTO(115,116,117),MMT
115   PRINT118
118   FORMAT(24HFIRST ORDER ARMODEL USED)
      PT=SQRT(1.0-AA1*AA1)
      DO119K=1,L
      IF(K-1)10,10,11
10    ST(K)=PT*RAND(K)
      GO TO 119
11    ST(K)=PT*RAND(K)+AA1*ST(K-1)
119   CONTINUE
      GO TO 40
116   PRINT 125
```

```
125   FORMAT(24HSECOD ORDER ARMODEL USED)
      PT=SQRT(1.0-(A21*A21+A22*A22+2.0*A21*A22*R21))
      DO126K=1,L
      IF(K-1)20,20,21
20    ST(K)=PT*RAND(K)
      GOTO126
21    IF(K-2)24,24,25
24    ST(K)=A21*ST(K-1)+PT*RAND(K)
      GOTO126
25    ST(K)=A21*ST(K-1)+A22*ST(K-2)+PT*RAND(K)
126   CONTINUE
      GOTO40
117   PRINT128
128   FORMAT(24HTHIRD ORDER ARMODEL USED)
      PT=SQRT(1.0-(T1*T1+T2*T2+T3*T3+2.0*T1*T2*TR+2.0*T1*T3*TRR
     1+2.0*T2*T3*TR))
      DO129K=1,L
      IF(K-1)30,30,31
30    ST(K)=PT*RAND(K)
      GOTO129
31    IF(K-2)34,34,35
34    ST(K)=T1*ST(K-1)+PT*RAND(K)
      GOTO129
35    IF(K-3)36,36,38
36    ST(K)=T1*ST(K-1)+T2*ST(K-2)+PT*RAND(K)
      GO TO 129
38    ST(K)=T1*ST(K-1)+T2*ST(K-2)+T3*ST(K-3)+PT*RAND(K)
129   CONTINUE
40    GO TO (91,92),MIT
91    DO130K=1,L
      ST(K)=AV2+STPS*ST(K)
130   CONTINUE
      DO131K=1,L
      Y(K)=ST(K)+PCOM(K)
131   CONTINUE
      GOTO191
92    K=0
      DO132I=1,M1
      DO132J=1,12
      K=K+1
      Q(I,J)=AT(J)+GS(J)*ST(K)
132   CONTINUE
      CALL  DSSS(Q,M1,Y,L)
191   GOTO(97,98),IMT
97    DO99K=1,L
      Y(K)=EXP(Y(K))
99    CONTINUE
98    CALL  SSDS(Y,L,Q,M1)
      DO802I=1,M1
      DO802J=1,12
      IF(Q(I,J))71,71,72
71    PRINT78,I,J,Q(I,J)
78    FORMAT(1X,2I5,F12.4)
      Q(I,J)=0.0
72    G(I,J)=Q(I,J)
802   CONTINUE
      CALL    STCS(Q,N,AV,STD,B,R,CV,CS,CK)
      CALL  STCS(Q,M1,AV,STD,B,R,CV,CS,CK)
```

```
      CALL  DSSS(Q,M1,Y,L)
      LL=1
      CALL  STAT(Y,LL,L,AVV,STDS,CSS)
      CALL  STAT(Y,LL,M1,AVV,STDS,CSS)
      CALL  CORG(Y,L,K1,K2,M,SCC)
      CALL CORG(Y,M1,K1,K2,M,SCC)
      GOTO108
1000  STOP
      END
```

```
      PROGRAM.NO.11
      THOMAS FIERING MODEL-NO TRANSFORMATION OF ORIGINAL  DATA
C     H.S.V.MURTHY STRUCTURAL ANALYSIS OF HYDROLOGIC TIME SERIES
      DIMENSIONQ(100,13),AV(13),STD(13),B(13),R(13),CV(13),CS(13),CK(13)
      DIMENSIONEPSL(1200)
      READ101,N,M
  101 FORMAT(2I3)
      READ102,((Q(I,J),J=1,12),I=1,N)
  102 FORMAT(6F12.4)
      PRINT301
  301 FORMAT(//27HRESULTS FOR ORIGINAL SERIES//)
      L=M*12
      READ103,(EPSL(I),I=1,L)
103   FORMAT (6F12.6)
      CALL        STCS (Q,N,AV,STD,B,R,CV,CS,CK)
      Q(1,1)=AV(1)
      IJ=0
      DO 902 I=1,M
      DO 902 J=1,12
      KK=I
      K=I
      JJ=J+1
      IJ=IJ+1
      IF(JJ-12)202,202,203
  203 JJ=1
      KK=KK+1
202   Q(KK,JJ)=AV(JJ)+B(J    )*(Q(K,J)-AV(J))+STD(JJ)*(SQRT(1.-R(J)
     1*R(J)))*EPSL(IJ)
902   CONTINUE
      PRINT303
  303 FORMAT(//40HRESULTS FOR GENERATED              SERIES//)
      PRINT 102,((Q(I,J),J=1,12),I=1,M)
      DO802 I=1,M
      DO802 J=1,12
      IF(Q(I,J)) 71,71,72
71    PRINT35,I,J,Q(I,J)
35    FORMAT(2I5,F12.4)
      Q(I,J)=0.0
72    Q(I,J)=Q(I,J)
802   CONTINUE
      CALL        STCS (Q,M,AV,STD,B,R,CV,CS,CK)
      CALL        STCS (Q,N,AV,STD,B,R,CV,CS,CK)
      STOP
      END
```

```
      PROGRAM.NO.12
      THOMAS FIERING MODEL-LOG. TRANSFORMATION OF ORIGINAL DATA
C     M.S.V.MURTHY STRUCTURAL ANALYSIS OF HYDROLOGIC TIME SERIES
      DIMENSIONQ(100,13),AV(13),STD(13),B(13),R(13),CV(13),CS(13),CK(13)
    . DIMENSIONEPSL(1200),AT(12)
      READ101,N,M
  101 FORMAT(2I3)
      READ102,((Q(I,J),J=1,12),I=1,N)
  102 FORMAT(6F12.4)
      PRINT301
  301 FORMAT(//27HRESULTS FOR ORIGINAL SERIES//)
      L=M*12
      READ103,(EPSL(I),I=1,L)
103   FORMAT (6F12.6)
      CALL        STCS (Q,N,AV,STD,B,R,CV,CS,CK)
      DO201I=1,N
      DO201J=1,12
      AT(J)=AV(J)
      Q(I,J)=Q(I,J)+0.012*AT(J)
201   Q(I,J)=ALOG(Q(I,J))
      CALL        STCS (Q,N,AV,STD,B,R,CV,CS,CK)
      Q(1,1)=AV(1)
      IJ=0
      DO 902 I=1,M
      DO 902 J=1,12
      KK=I
      K=I
      JJ=J+1
      IJ=IJ+1
      IF(JJ-12)202,202,203
  203 JJ=1
      KK=KK+1
202   Q(KK,JJ)=AV(JJ)+B(J   )*(Q(K,J)-AV(J))+STD(JJ)*(SQRT(1.-R(J)
     1*R(J)))*EPSL(IJ)
902   CONTINUE
      PRINT303
  303 FORMAT(//40HRESULTS FOR GENERATED LOG ARITHM  SERIES//)
      PRINT 102,((Q(I,J),J=1,12),I=1,M)
      DO802 I=1,M
      DO802 J=1,12
      Q(I,J)=EXP(Q(I,J))
      Q(I,J)=Q(I,J)-0.012*AT(J)
802   CONTINUE
      CALL        STCS (Q,M,AV,STD,B,R,CV,CS,CK)
      CALL        STCS (Q,N,AV,STD,B,R,CV,CS,CK)
      STOP
      END
```

PROGRAM.NO. 13

```
C     GENERATION  OF  DATA  BYTFM
      DIMENSION Q(40,13),AV(13),STD(13),B(13),R(13),CV(13),CS(13),CK(13)
      DIMENSION EPSL(600)
      READ101,N,M
  101 FORMAT(2I2)
      READ102,((Q(I,J),J=1,12),I=1,N)
  102 FORMAT(6F12.4)
      PUNCH301
  301 FORMAT(//27HRESULTS FOR ORIGINAL SERIES//)
      L=M*12
      DO 201 I=1,N
      DO 201 J=1,12
201   Q(I,J)=SQRT(Q(I,J))
      CALL        STCS (Q,N,AV,STD,B,R,CV,CS,CK)
      READ103,(EPSL(I),I =1,L)
103   FORMAT (6F12.6)
      Q(1,1)=AV(1)
      IJ=0
      DO 202 I=1,M
      DO 202 J=1,12
      KK=1
      K=I
      JJ=J+1
      IJ=IJ+1
      IF(JJ-12)202,202,203
  203 JJ=1
      KK=KK+1
202   Q(KK,JJ)=AV(JJ)+B(J   )*(Q(K,J)-AV(J))+STD(JJ)*(SQRTF(1.-R(J )
     1*R(J)))*EPSL(IJ)
      PUNCH 303
  303 FORMAT(//40HRESULTS FOR GENERATED SQUARE ROOT SERIES//)
      DO802 I=1,M
      DO802 J=1,12
      IF(Q(I,J)) 71,71,72
71    Q(I,J)=0.0
72    Q(I,J)=Q(I,J)*Q(I,J)
802   CONTINUE
      PUNCH 102,((Q(I,J),J=1,12),I=1,M)
      CALL        STCS (Q,M,AV,STD,B,R,CV,CS,CK)
      STOP
      END
```

```
   PROGRAM.NO.14
   THOMAS FIERING MODEL-INCORPORATING SKEWNESS PARAMETER ALSO
   H.S.V.MURTHY STRUCTURAL ANALYSIS OF HYDROLOGIC TIME SERIES
   DIMENSIONQ(100,13),AV(13),STD(13),B(13),R(13),CV(13),CS(13),CK(13)
   DIMENSIONEPSL(1200),GA(13)
   READ101,N,M
01 FORMAT(2I3)
   READ102,(GA(J),J=1,12)
02 FORMAT(6F12.4)
   READ102,((Q(I,J),J=1,12),I=1,N)
   PRINT301
01 FORMAT(//27HRESULTS FOR ORIGINAL SERIES//)
   CALL        STCS (Q,N,AV,STD,B,R,CV,CS,CK)
   L=M*12
   READ103,(EPSL(I),I=1,L)
   FORMAT (6F12.6)
   Q(1,1)=AV(1)
   IJ=0
   DO 902 I=1,M
   DO 902 J=1,12
   KK=I
   K=I
   JJ=J+1
   IJ=IJ+1
   IF(JJ-12)202,202,203
03 JJ=1
   KK=KK+1
   EL=2.0/GA(J)
   ELL=1.0+(GA(J)*EPSL(IJ)/6.0)-(GA(J)/36.0)
   EPSL(IJ)=EL*(ELL**3)-EL
   Q(KK,JJ)=AV(JJ)+B(J   )*(Q(K,J)-AV(J))+STD(JJ)*(SQRT(1.-R(J)
  1*R(J)))*EPSL(IJ)
   CONTINUE
   PRINT303
03 FORMAT(//40HRESULTS FOR GENERATED SQUARE ROOT SERIES//)
   PRINT 102,((Q(I,J),J=1,12),I=1,M)
   DO802 I=1,M
   DO802 J=1,12
   IF(Q(I,J)) 71,71,72
   PRINT35,I,J,Q(I,J)
   FORMAT(2I5,F12.4)
   Q(I,J)=0.0
   Q(I,J)=Q(I,J)
   CONTINUE
   CALL        STCS (Q,M,AV,STD,B,R,CV,CS,CK)
   CALL        STCS (Q,N,AV,STD,B,R,CV,CS,CK)
   STOP
   END
```