"ON THE IDENTIFICATION OF AN OPTIMUM FLOOD FREQUENCY MODEL"

FRANCIS M. MUTUA

A thesis submitted in fulfilment for the degree of Doctor of Philosophy in Meteorology in the University of Nairobi.

March 1986

This thesis is my original work and has not been presented for a degree in any other University.

Signature 10/4/1986

FRANCIS MUSYOKA MUTUA

This thesis has been submitted for examination with our approval as University supervisors.

Signature ....Dr. L. J. Ogallo....................

Signature ....

Prof. J.D. SALAS

Department of Meteorology

University of Nairobi

P.O. Box 30197

Nairobi, Kenya

TABLE OF CONTENTS

PAGE

PAGE

PAGE

ABSTRACT

Although a number of nonparameteric analysis have appeared in the literature, traditional flood frequency analysis has been approached primarily as a problem in parametric statistical inference. Peak annual streamflow data are assumed to come from a parent population whose distribution function is known, is analytically expressable and contains a finite number of parameters. A large number of peak flow distributions have been studied, for example, the normal, the lognormal, the Gumbel, the Gamma and Weibull distributions. Goodness of fit procedures then test whether or not the data do indeed fit the assumed distribution with a specified degree of confidence. However, the use of the convetional goodness of fit procedures in flood frequency analysis has several disadvantages. Firstly, is the lack of power of these goodness of fit tests with respect to the typically skewed flood peak distributions. This generates considerable variability in the estimation of design events. Secondly, the conventional goodness of fit tests are subjective in that the final results drawn from such tests depend very much on the level of confidence utilized. This means that different levels of confidence can often lead to conflicting results. Lastly, and probably the most serious disadvantage is

that these conventional goodness of fit tests can pass more than just one flood peak distribution model in a given situation. Hence, these disadvantages introduce limitations to the usefulness and power of such goodness of fit procedures in the identification process of the most optimum flood peak probability model. However, Kite (1977) shows that these conventional goodness of fit tests can nevertheless be modified to be useful in model identification.

The Akaike Information theory is used in this study to introduce another approach of goodness of fit which can be used to identify, more positively, the most optimum flood peak probability model. The Akaike Information Criterion gives an objective measure of the goodness of fit of a given flood probability model. This measure has no relevance to the commonly used confidence limits. The minimum of these measures for the various competing flood peak distribution models is used to identify the most optimum model. The Akaike Information Criterion is also used, in this study, to detect the outliers, which due to the nature of flood peak data can exist. The goodness of fit of nine probability distributions was investigated in this study. These comprised of the lognormal type 3, the Pearson, the log-Pearson, the Fisher Tippet and log-Fisher Tippet distributions; and four relatively new probability distributions, namely, the Walter Boughton, the log-Walter Boughton (Boughton, 1980), the Wakeby and the log-Wakeby (Boughton, 1978a) distributions. The annual peak flow data used in this study was collected from sixty river catchments with areas from a few square kilometers to several thousand square kilometers, and distributed randomly in all the major basins in Kenya. The lengths of the records varied from at least twenty five years to about fifty years with the majority having between thirty to forty years.

The modified Smirnov-Kolmogorov and the chi-square tests identify the Wakeby distribution as the best for flood frequency analysis, compared to the rest. This is further confirmed by the more objective and reliable AIC test. Thus, the results show that the Wakeby distribution is generally the most suitable for flood frequency analysis in Kenya. One of the worst fitting distribution is found to be the log-Pearson Type 3, despite its popularity worldwide. The results also show that the peak flow data used does not have any true outliers. However, some of the poorly fitting distributions occasionally show some outlying observations when the more flexible distributions indicate no such outlying values.

LIST OF FIGURES

LIST OF TABLES

TABLE | PAGE

CHAPTER I

1.0 INTRODUCTION

The magnitude and frequency of occurrence of extreme hydrologic events is of great importance in all parts of the world. Since man has for reasons of communication, water supply, agriculture and others, built most of his communities on the flood plains of large rivers, his lifestyle is extremely susceptible to flood disasters. Today's pressure of population increases the density of development along the river banks.

Forecasts of flood events can generally reduce the damages caused by floods. There are two ways in which this beneficial effect can be achieved. Firstly and most obviously, warning of flood event which enables people to evacuate the danger zone. If a sufficient lead-time is provided, vulnerable possessions can be removed from the danger zone. Preparations can also be made such as sand-bagging, to minimize the flood damage.

The second method of achieving benefits through forecasting is to use flood frequency analysis in the design of the hydraulic structures within the flood plain, and for flood plain zoning. As examples, knowledge of the magnitude-frequency relationships can be used in the design of dams, highway bridges, railway bridges, culverts, water supply systems and flood control structures (Kite, 1977). The American Water Works Association (1966) reported that about 20% of the dam failures all over the world are mainly due to faulty spillway designs. Flood plain zoning ensures that housing and industrial developments are not located in the high risk zones; or at least, if they are located in such areas, there can be no justification for compensation in the inevitable event of flood damage. High risk zones along the banks of rivers susceptible to floods should be used for activities compatible with this risk such as parkland, recreation areas, some type of grazing or crop production (Kite, 1977).

Frequency analysis can, not only be used as a means of preventing disaster but also as a means of introducing efficient designs. When a hydraulic design is underdesigned through inadequate or inaccurate data or methods, the results, are regretably obvious. The dam may fail, the highway may flood or the bridge may collapse. These extreme consequences do not happen very often and so the hydrologist, equating non-failures with successes, is satisfied with his design techniques. Frequently, structures are overdesigned and hence very safe, but also very expensive. A truly efficient design can only be achieved as a result of studies relating cost of risk and frequency analyis.

The United States Water Resources Council (1967) noted that because of the range of uncertainity in design flood analysis, there is a need for continued research and development to investigate the many unsolved problems in the field. Current methods providing design floods for hydraulic structures included the use of the flood formulae (Dalrymple, 1964), the deterministic use of the meteorological data in techniques such as the dynamic flow equations and the so called Probable Maximum Flood method and the stochastic use of the frequency analysis. The Probable Maximum Flood and other similar methods suffer from the major disadvantages of being entirely subjective and of having no associated probability level. The latter is particularly important since to non-technical people, it implies that no risk is involved and that the maximum flood cannot exceed this certain limit. This of course is untrue and can at times have disastrous consequences. In order to be able to define the objectives of this research more fully, the following section gives a general review of the various methods which are used in flood forecasting and design.

1.1 LITERATURE REVIEW

We can divide the flood forecasting and design procedures into two broad categories:

(i) rainfall - runoff catchment models which are utilized essentially to forecast the future states of streamflow in a river, and

(ii) methods which are used for hydraulic design purposes only. These two broad categories are reviewed independently in the following subsections.

1.1.1 CATCHMENT MODELS

Most frequently, the warning of a flood event is based on a relationship between the river stage at the location of interest and the upstream hydrological meteorological, topographical and soil parameters. This type of relationship is called a catchment rainfall - runoff model.

River flows are the result of an interplay of many physical processes. The complexity of hydrological phenomena due to the great variety in the characteristics of a drainage basin and its response to precipitation together with the difficulties associated with obtaining the accurate data makes it impossible to reproduce such phenomenon accurately. Hence, it becomes necessary to construct simplified models of hydrological processes for use in flood flow forcasting. In general, catchment models can be divided into three main categories: similitude models, analogue models and mathematical models.

1.1.1.1 SIMILITUDE MODELS

The similitude models are based on the analysis of the physics of the various processes in the hydrologic cycle (Terstreip and Stall, 1974). They essentially consist of constructing a scale model of a drainage basin and then calibrating it by comparing the rainfall - runoff relations for the scale model with those of the actual drainage basin. The similitudes models have the advantage that most of their governing parameters have a direct physical interpretation and their ranges can be established reasonably well on the basis of field and laboratory investigations. These models are most useful when the streamflow data is not readily available, for example if the catchment is ungauged or if a hypothetical land use is to be investigated. A major disadvantage of the similitude models is that they are generally quite demanding in terms of computers and data requirements. Such models often require an intensive and sophisticated instrumentation in the watershed for calibration purposes. Further, the results of experiments under laboratory or field conditions are affected by additional factors not taken into account or not accurately reproduced in the model. Due to these constraints, similitude models are generally not popular for forecasting purposes.

1.1.1.2 ANALOGUE MODELS

Analogue models are based on the analogy between the movement of water in streams and electric currents in designed circuits or the flow of water in a system of vessels (Dawdy and Litchy, 1968). These models make use of the electronic analogue computers which solve problems by behaving electronically in a manner analogous with the problem solution. The analogue computer has the advantage of integrating the problem variables continuously instead of using numerical approximations as is the case with digital computers. Since the analogue computer is a parallel device in that all computations proceed simultaneously, it is a much faster computing device than the digital computer. When the size of a problem is doubled the time for solution in the analogue computer remains the same although the amount of equipment required is also approximately doubled. On the other hand the digital computer, which is a sequential machine, takes about twice as long when the problem is doubled and when the amount of equipment remains unaltered.

When used for streamflow forecasting, the result of an analogue computer is presented in a graphical form as a continuous plot of the variables involved. The operator can visualize results as being the dynamic response of the drainage basin under investigation. Also, the results of the various ways of combining the different components of the entire system can be quickly defined to determine the changes in specific processes that are necessary to meet the prototype conditions. Thus, the analogue system is very helpful during the exploratory phases of developping the component relationships and also the composite model. Since the changes in problem size affect the amount of the analogue equipment required, and also considering the economic constraints that may be involved, the use of analogue models is undertaken when all the other types of models are not feasible economically.

1.1.1.3 MATHEMATICAL MODELS

Mathematical models are the most widely used catchment models at the present time (Clark, 1973). This is mainly due to their flexibility and ability to handle large quantities of data with computers to describe complex systems. For purposes of hydrological forecasting, there are about four main categories of the mathematical models, namely, stochastic models, system approach models, dynamic models and conceptual catchment models.

Stochastic models are relatively classical procedures based on simplified hydrologic balance equations which are usually calibrated statistically. These models normally yield the volume of runoff or some other outstanding characteristic such as the peak discharge.

The systems approach models describe mathematically the relation between precipitation and streamflow without reference to the process by which they are related. They are also called lumped parameter or black-box models. The unit hydrograph, time-area curves and linear reservoir concepts, often in their more complicated systems analysis versions, are the basic examples of the approach. These models are often used in combination with others. In such cases, the systems approach components represent the actual distribution or response function.

Dynamic models, also called distributed parameter models, are based on the solution of equations for unsteady motion of water both for overland and channel flow, either in their pure hydraulic form or in a simplified hydrological form. The Kinematic wave, or other unsteady flow description by simplified differential equations, mostly solved by numerical methods represent examples of this type of model. These types of models are mostly suitable for routing channel flow from one point to another.

Conceptual catchment models are based on simplified description of water movement in the basin, both in time and space, on and under the surface of the earth, accounting for water storage and movement in discrete but short time intervals. The moisture accounting in these models can be either indices which are either implicit or explicit. The prolification of conceptual models and their increased importance for hydrological forecasting led to the World Meteorological Organization (WMO) to initiate a project on intercomparison of conceptual models which are mostly used in operational hydrologic forecasting (WMO, 1975). The choice of a model to be used depends mainly on the intended application and on how much information is available, on the behaviour of the system to be modelled.

1.1.2 METHODS USED FOR FLOOD DESIGN PURPOSES ONLY

The peak discharge that has been or may be exprienced is a pertinent item in the design of structures along or across a river. The peak discharge can be computed from the Probable Maximum flood or from a flood frequency analysis concept.

1.1.2.1 FLOOD FORMULAE

Many flood formulae have been developed for computing peak discharges (Chow, 1962), but most are considered inadequate for engineering design. Studies based on flood records are usually preferred, but flood records are scarce for most drainage areas, and often use must be made of some formula. The most commonly used flood formulae are the rational formula and the Myers-Jarvis enveloping curves.

The relation between rainfall and the peak discharge has been represented by many empirical or semi-empirical formulae. The rational formula can be taken as a representative of such formulae. Although this formula is based on a number of assumptions which cannot be readily satisified under natural conditions, its simplicity has won its popularity.

The rational formula is of the form

$$Q = CIA \qquad \text{.... (1)}$$

where Q is the peak discharge, C is a runoff coefficient depending on the characteristics of the drainage basin, I, is the rainfall intensity and A is the drainage area which is in acres if Q is in cubic feet per second and I in inches per hour. The formula is called rational because the units of the quantities involved are numerically consistent approximately.

When using the formula, one must assume that the maximum rate of flow owing to a certain rainfall intensity over the drainage area is produced by that rainfall which is maintained for a period equal to the period of concentration of the flow at the point under consideration. Theoretically, this is the *time of concentration* which is the time required for the surface runoff from the remotest part of the drainage basin to reach the point being considered. For uniform rainfall intensity, this is the *time of equilibrium* at which the rate of runoff is equal to the rate of rainfall supply. For natural drainage basins of large size and complex drainage patterns, runoff water originating in the most remote portion may arrive at the outlet too late to contribute to the peak flow. Accordingly, the time of concentration is generally greater than the lag time of the peak flow. For small drainage basins with simple drainage patterns, the time of concentration may be very close to the lag time of the peak flow. For small agricultural drainage basin, Rasmer (1972) determined the time of concentration by noting the time required for the water in the channel at the gauging station to rise from the low to the maximum stage as recorded by the water-stage recorder. An emperical formula for the time of concentration in hours is given by Kirpich (1940),

$$t_c = 0.00013 \frac{L^{0.77}}{S^{0.385}} \qquad \text{.... (2)}$$

where L is the length of the basin area in feet, measured along the water-course from the gauging station and in a direct line from the upper end of the water-course to the farthest point on the drainage basin, and S is the ratio in feet to L of the fall of the basin from the farthest point of the basin to the outlet of runoff, or approximately the average slope of the basin in dimensionless ratio.

The range of values of the runoff coefficient C, in relation to the characteristics of the drainage basin are discussed by Merrill (1932).

The assumptions involved in the rational formula are:

(i) The rate of runoff resulting from any rainfall intensity is a maximum when this rainfall intensity lasts as long or longer than the time of concentration.

(ii) The maximum runoff resulting from a rainfall intensity, with a duration equal to or greater than the time of concentration, is a simple fraction of such rainfall intensity: that is, it assumes a straight line relation between Q and I, and Q = O when I = O.

(iii) The frequency of peak discharges is the same as that of rainfall intensity for a given time of concentration.

(iv) The relationship between peak discharges and size of drainage area is the same as the relationship between duration and intensity of rainfall.

(v) The coefficient of runoff is the same for storms of various frequencies.

(vi) The coefficient of runoff is the same for storms on a given watershed.

It is believed (Chow, 1964) that these assumptions might nearly hold for paved areas with gutters and sewers of fixed dimensions and hydraulic characteristics. The formula has thus been rather popular for the design of drainage systems in urban areas and airports. The exactness and satisfaction of these assumptions in application to other drainage basins, however, have been questioned. In fact, many hydrologists have called attention to the inadequacy of the method.

Probably, the most widely used flood formula is that developed by Major E.D.T. Myers (Dalrymple, 1964). As modified by Jarvis (1942), this formula is given as

$$Q = 100p\sqrt{M} \quad .... (3)$$

where Q is the discharge in cubic feet per second, p is the numerical percentage rating on the Myers scale, and M is the size of the drainage area in square miles. In ordinary use, peak discharge Q in cubic feet per second per square mile are plotted on logarithmic paper against the drainage area M is square miles. Then a straight line, called the *Myers curve*, with a slope of 1:2 is drawn as an envelope through the upper points. This curve is intended to give an estimate of flood peaks that could occur anywhere in the region.

The Myers - Jarvis formula makes two questionable assumptions:

(i) The flood peaks vary as the 0.5 power of the drainage area, and

(ii) The flood producing characteristics of the streams in the region under consideration follow the same law as expressed by the formula.

Experience, however, has shown that these assumptions are usually not realistic (Dalrymple, 1964).

1.1.2.2 THE PROBABLE MAXIMUM DESIGN FLOOD

It is recognised that there is a physical upper limit to the amount of precipitation that can fall over a specified area in a given time (Bernard, 1944). This upper limit has become known as the Probable Maximum Precipitation, and is more precisely defined as, that depth of precipitation which for a given area and duration, can be reached but not exceed under known meteorological conditions.

A procedure to determine probable maximum precipitation for regions of little to moderate topographic variation has been developed and is widely applied (U.S. Weather Bureau, 1937-1970, Bernard, 1944). The procedure involves two steps:

(i) the preparation of probable maximum depth - area - duration curves, and

(ii) the selection by means of these curves of a pattern of storm for use in the basin.

The probable maximum depth - area - duration curves are prepared by simple enveloping of the moisture adjusted depth - area - duration values for all storms considered transposable to the region of study. The moisture adjustment factor is the ratio of the maximum total moisture content in an atmospheric column of unit cross-section in the region to the total moisture in a similar column that occurred during the storm. The maximum values for different size - areas and for different durations may and probably do come from different storms. The pattern storm depends on the size of area, and to a lesser degree, on the duration deemed hydrologically critical. For areas upto a a few hundred square miles in size and not too peculiar in shape, the pattern is not usually very critical and a single pattern may be applicable to many basins (Chow, 1964). For areas of tens of thousands of square miles or more in size and areas of unusual shape, close examination of the patterns may be necessary to determine one that is realistic. In ordinary cases, the pattern is based on one of the storms that gives the enveloping depth - area - duration values near the area size and duration of hydrologic importance. All depth - area - duration values for the storm are multiplied by a factor, which is the factor that brings the values of certain points to tangency with the enveloping curves. In case of unusual basin shape or of large topographic variation, the pattern storm may be based on combinations of different parts of actual storms.

Although the use of the probable maximum precipitation has some serious problems (as will be mentioned later), the advantages of the procedure are several. Firstly, it provides both empirical and statistical controls. Secondly, the values obtained are directly related to the largest that have occurred. Thirdly, the exprience of a basin is extended through transposition. Fourthly, the use of actual storms for patterns ensures realism in that nature's integrations are used rather than hard - to - justify synthetic ones, and lastly, the overcompounding of probabilities is minimised.

Several features of the results which are obtained through this method are worth mentioning. The highest estimates of probable maximum precipitation often exceeds the greatest value of observed precipitation, in certain areas, by only a small percent. In other basins, they may be several times as great as the maximum observed. The greatest maximising process for a given basin is storm transposition. If a precipitation value several times as large as any over a given project basin has been observed over a nearby basin, then it is considered that the observed ischyets of the actual storm can be transferred, or transposed so as to indicate the maximum amount over the project basin. Obviously, there are geographical, topographical as well as synoptic limits as to how far this transposition can be extended. A difficulty arises in areas that are small in geographical extend and not homogeneous, meteorologically and topographically, with the neighbouring areas. Then, transposition results in values that are very much lower, comparatively, than would be obtained from transposition in large homogeneous areas.

The maximum probable flood can be obtained from the probable maximum precipitation by substracting loss rates and adding the snow melt (where applicable) with due regard to the relationship with time. It is seldom economically practicable to design for the probable maximum flood. The design flood is usually selected by pertinent facts. Pertinent facts are represented by stream-flow records which are either computed from precipitation records or by application of hydrologic principles to measured physical factors or directly observed and analysed to best apply to the particular situation.

Where an area is thickly inhabited or developed industrially, and the failure of protective works can result in loss of life and great property damage, a design on the basis of the probable maximum flood may be justified. In agricultural areas where failure would result only in flooding of crops, a design for a much smaller degree of protection would be justified. Varying conditions lie between these extremes and varying design discharges would be called for in providing protective works.

1.1.2.3 FLOOD FREQUFNCY ANALYSIS

Flood frequency analyses are generally made for one of the two purposes.

(i) as a guide to judgement in determining the capacity of a structure, such as a highway bridge opening or a cofferdam where it is considered permissible to take a calculated risk, and

(ii) as a means of estimating the probable flood damage prevented by a system of flood protection works over a period of years, usually equal to the estimated economic life of the works.

In the first case, the magnitude cf the flood discharge that will be equalled or exceeded in a certain period of years is required. In the second case, it may be necessary to consider, in addition to the peak flood discharge, peaks and season of occurrence. Most often, it is for the purpose of estimating the design flood (capacity of a hydraulic structure) that flood frequency analysis is done.

One of the most common problems faced in hydrology is the estimation of a design flood from a fairly short record of stream-flow data. By plotting the magnitude of the measured events against their observed return periods on a probability paper, some kind of a pattern is generally apparent. The difficulty with this approach is how to use this pattern to extend the available data and estimate design floods of higher return periods.

If a large number of records, at least as long as the return period of the required design event, are available then the problem is simplified. In the extreme, if a large data sample were available, then the design event and its risk estimates could be derived directly from the sample data. However, this kind of data is never available and therefore the short data sample is generally used to fit a frequency distribution which in turn is used for extrapoloration to the desired design event, either graphically or by estimating the parameters of the probability distribution function.

Graphical methods have the advantage of simplicity and visual representation and also the fact that no assumption of distribution type is made. However, these advantages are outweighed by the disadvantage that the method is highly subjective and is therefore not compatible with the other phases of engineering design. Nevertheless, subjective information need not be totally elliminated since the techniques such as the Bayesian analysis may permit its use (Wood and Rodriquez-Iturbe, 1975).

Although a number of non-parameteric frequency analyses have appeared in the literature (Vicens *et al*, 1975; Wood and Rodriquez-Iturbe, 1975), traditional flood frequency analysis has been approached primarily as a problem in parameteric statistical inference (Yevjevich, 1972b). In this method, peak annual streamflow data are assumed to come from a parent population whose distribution function is known, is analytically expressable and contains a finite number of parameters. The main source of error in using the parameteric approach in flood frequency analysis is that, it is not known apriori which of the many possible probability distributions is the true distribution, that is, which distribution, if any, the flood events naturally follow. Thus, an identification process is usually necessary, since in general, the sample events available are for relatively low return periods (around the centre of the distribution) while the required estimates of the design events are of large return periods in the tail of the distribution. Many distributions have similar shapes in their centres but differ widely in their tails. It is thus possible to fit several distributions to a sample data and end up with several different estimates of the required design flood. Chi-square and similar tests of goodness - of - fit can be used to choose the distribution which best describes the sample data. However, this does not usually solve the basic problem (Kite, 1977).

Numerous different probability or frequency distributions have been used in hydrology (Chow, 1964). Discrete distributions such as the binomial and Poisson have been used to define the average intervals between flood events and to evaluate risks (Hall and Howell, 1963; Kalinin, 1960). Continuous distributions such as the normal, the lognormal, Pearson, Log-Pearson and the general extreme value distributions (in the real and logarithmic domains) have been studied with relevance to flood modelling. Furthermore, other new and more meaningful distributions such as the Walter Boughton and Wakeby distributions have been discovered (Houghton, 1978a; Boughton, 1980).

Once a probability distribution has been chosen, the second source of error becomes apparent. The statistical parameters of the probability distribution must be estimated from the sample data. Since the sample data is subject to error (more so with flood data), the method of fitting must be able to minimise these errors and must therefore be as efficient as possible. In the following sections, we will review the estimation methods which are commonly used in hydrology.

1.1.2.3.1 PARAMETER ESTIMATION METHODS

The population properties of random variables are usually characterised by parameters, percentiles or other similar numbers. Because populations are rarely known in sciences like hydrology, the population properties of random variables must be estimated from the available sample data. Any estimate obtained from a sample is a statistic. The statistic also may be defined as a function of the sample values. This function of the sample values also refers to some property of the population of random variable.

If a probability distribution has three parameters α, β and γ the estimators of these values denoted here as a, b, and c should be as close to the population values as is practically possible. The measure of goodness of the estimation methods are the variances of the differences (α- a), (β- b) and (γ-c). The smaller these variances are, the better is the estimation method. The estimated values a,b, and c, are called the estimators of α,γ and γ. There is a large number of possible estimates of any parameter Some of them may be considered as best estimates. The concept of best estimates must be clearly defined. Though the best estimates, designated by $\hat{\alpha}$, $\hat{\beta}$ and $\hat{\gamma}$ there may be a large number of these estimates (Yevjevich, 1972a).

Because the various samples of the same size drawn from a given population have different estimates a,b and c of α, β and γ respectively, the sampling distributions of the estimates a,b, and c about their population values α, β and γ are important properties of estimation methods. The bias and mean square errors are the main criteria used for determining the best estimates (Raynal and Salas, 1985). The smaller the variance of an estimate, the better the estimate is in comparison with the others.

An estimator is often classified as biased or unbiased, regular or irregular, efficient or inefficient, sufficient, or insufficient and consistent or inconsistent (Yevjevich, 1972a). To define these five concepts of estimates, assume that the estimation is needed for α in the one parameter probability density function, $f(x;\alpha)$, of the random variable, x; or for α and β in the two parameter probability density function $f(x;\alpha,\beta)$. The variance of partial derivatives of the logarithms of these functions, designated by D(a) is given by

$$D(a) = \int_{-\infty}^{\infty} \left[\frac{\partial \ln f(x;\alpha)}{\partial \alpha} \right]^2 f(x;\alpha)dx - E\left[\frac{\partial \ln f(x;\alpha)}{\partial \alpha} \right]^2 =$$

$$\int_{-\infty}^{\infty} \frac{1}{f(x;\alpha)} \cdot (\underline{\dot{f}(x;\alpha})^2 \, dx \quad \ldots. \ (4)$$

In the case $f(x;\alpha,\beta)$, the corresponding expressions of (4) give D(a) and D(b).

If the expected value of an estimator, a, of α is

$$E(a) = \alpha+\nu(\alpha) \quad .... (5)$$

in which $\nu(\alpha) \neq 0$, the estimator, a, is unbiased.

The difference between the sample and population values is $(a-\alpha)$. If the variance of this difference satisfies the condition that it is greater than a certain value (given in (6)) the estimator is called regular. This inequality is (Carmer, 1958):

$$Var(a-\alpha) = E(a-\alpha)^2 \geq \frac{\left[1+ \frac{d\nu(\alpha)}{d\alpha}\right]^2}{N.D(a)} \quad .... (6)$$

in which N is the sample size, D(a) is defined by (4) and $\nu(\alpha)$ is defined by (5). For an unbiased estimator, $\hat{\alpha}$, (6) becomes

$$Var(\hat{\alpha}-\alpha)^2 \geq 1/N.D(\infty) \quad .... (7)$$

If an estimator is unbiased and regular, and if

$$Var(\hat{\alpha}) = E\ (\hat{\alpha}-\alpha) = 1/N.D(\alpha) \quad .... (8)$$

the estimator is efficient (Yevjevich, 1972a). It has the smallest possible variance of deviation $(\hat{\alpha}-\alpha)$ for a given N and $f(x;\alpha)$, and is usually understood as the best estimator. For any other estimator a, different from α, the variance is $E(a-\alpha)^2 > E(\hat{\alpha}-\alpha)^2$, and the ratio

$$e_a = \frac{E(\hat{\alpha}-\alpha)^2}{E(a-\alpha)^2} \leq 1 \qquad \ldots . (9)$$

in which the numerator is given by (8) is called the efficiency of an estimation method which produces the variance given by the denominator of (9). Only an efficient estimator has a $e_a = 1$. Two efficient estimators α_1 and α_2, of the parameter α have the same mean, α, the same variance given by (8), and their correlation coefficient, $\rho = 1$. When an efficient estimator exists, it can always be determined by the method of maximum likelihood (Mood *et al*, 1974).

If α_1 and α_2 are two independent estimates of α, then α_1 is considered a sufficient estimate if the joint proabbility distribution of α_1 and α_2 has the property (Yevjerich, 1972a)

$$F(\alpha_1,\alpha_2) = F(\alpha_1).F(\alpha_2/\alpha_1) = F(\alpha_1)K(x_1,x_2, \ldots .,x_N) \qquad \ldots . (10)$$

in which $F(\alpha_1)$ is the distribution of α_1, $F(\alpha_2/\alpha_1)$ is conditional probability distribution of α_2 given α_1, and $K(x_1,x_2,\ldots .,x_N)$ is not a function of α. When (10) holds α_2 does not produce any new information about α, which is not already contained in α_1, regardless of the estimation method used to produce the estimate α_2. In this case, α_1 is optimum, or it is also a sufficient estimate.

Lastly, if the sample size N is large, $\hat{\alpha}$ is called a consistent estimate of α if it converges with a probability unity to α as N tends to infinity. Because many unbiased estimates have the variances of the type $Var(\hat{\alpha}) \approx C/\sqrt{N}$, where C is a constant, then in most cases the condition of consistency is usually satisfied. If this is the case, and if the efficiency e_a tends to unity as N goes to infinity, the estimate $\hat{\alpha}$ is called the asymptotically efficient estimate.

The above five properties of estimators have different significances when applied in statistics. In the majority of cases of estimation, the most important properties are to have unbiased (and therefor regular), and efficient (and usually sufficient and consistent) estimates (Yevjevich, 1972a). These two properties of unbiased and efficient estimates are usually required in hydrology when the requirement is to extract the maximum information from the same data.

Hydrologic data are often scarce. In arid and semi-arid regions, where the hydrological variables have the highest variation, and often have distributions which are highly skewed, the irony is that sample data are shortest in length, and the number of observational points is usually small. It is therefore a mistaken attitude of many practitioners in hydrology that simple estimation methods, like the graphical method or method of moments, are sufficiently accurate because of the scarce data. This approach should be quite the opposite. For these variables in hydrology with great variation, a highly skewed distribution, and a short series of observations, the unbiased and efficient estimates should be always searched for (Yevjevich, 1972a). The main problem is selecting the estimation method in each particular case that gives the unbiased and efficient estimates with the least computational cost.

Generally, the methods of distribution parameter estimation currently in use in hydrology given in their ascending order of efficiency, from the least efficient to the most efficient, are the graphical method, the least squares method, the method of moments and the maximum likelihood method. However, for large samples (of size greater than 20), and for distributions which have distribution functions expressable in an inverse form the method of probability weighted moments is generally most efficient. The following sections give a brief review of these parameter estimation methods.

1.1.2.3.1.1 THE GRAPHICAL ESTIMATION METHOD

If one draws the normal distribution in the form of a straight line in cartesian probability scales to the plotted empirical distribution, the 50% probability gives the estimate of the mean, $\bar{x}$, and the difference of the value of the variate x for 84.13 percent and 15.87 per-cent gives 2S, where S is the standard deviation. These values, $\bar{x}$ and S, are said to be graphically estimated, because the straight line fit of the normal distribution in these scales is usually made visually in drawing a straight line. The accuracy however, is very limited.

Generally, any distribution other than the normal distribution, can be used. However, the basic condition for use of this method is for the coordinate scales to be transformed in such a way that a given probability distribution function, $F(x;\alpha,\beta, \ldots)$, plots as a straight line. The graphical fit may, however, be used for the fitting of the functions.

$$y = f(x;\alpha,\beta,\ldots) \qquad \ldots\ (11)$$

without a transformation of coordinate scales. This requires, a visual tracing of the curve, selecting m points for m parameters $\alpha,\beta,\ldots$, in the various parts of the curve, through which (11) should pass. The coordinates y_i and $F_i(x;\alpha,\infty,\ldots)$ of m points give m equations from which the parameters may be estimated. The accuracy of this method is highly subjective, and depends on experience and good judgment.

1.1.2.3.1.2 LEAST SQUARES ESTIMATION METHOD

This method consists of the estimation of parameters by fitting a theoretical function to an empirical distribution, or any other empirical curve. The sum of all squares of deviations of observed points from the fitted function is minimised to produce least squares.

In the case a function $y = f(x;\alpha,\beta,)$ should be fitted to data by determining the best estimates a,b,...., of α,β, ..., the analytical method of least - squares minimizes the sum.

$$S = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (y_i - y)^2 = \sum_{i=1}^{n} \left| y_i - f(x_i;\alpha,\beta,....) \right|^2 \qquad .... (12)$$

in which x_i and y_i are coordinates of observed points, α,β, ... are replaced by their estimates a,b,...., and n is the sample size. Instead of minimizing $\sum(y_i-y)^2$, often $\sum e_n^2$, which are perpendicular deviations from the function $f(x;\alpha,\beta,....)$ also may be minimized. So that the line $f(x;\alpha,\beta,....)$ may have the minimum of (12), all partial derivatives of the sum with respect to parameters a,b,.... should be zero so that

$$\frac{\partial}{\partial a}\{\sum_{i=1}^{n}(y_i-y)^2\} = 0; \quad \frac{\partial}{\partial b}\{\sum_{i=1}^{n}(y_i-y)^2\} = 0; \quad .... (13)$$

These partical derivatives give m equations for the determination of m parameters. The simplest functions $f(x;\alpha,\beta,....)$ are those which are linear with respect to parameters. Among such functions, the polynomial function, of x is a general case, since other functions can be approximated by polynomials if they are developed in power series form. There must be more points than parameters, or $n > m$. Generally, n should be much greater than m, if the derived line is to be used for purposes of prediction.

Three conditions should be satisfied for the least squares method to be an efficient estimation method:

(a) that the deviations e_i of (12) are normally, or at least symmetrically, distributed;

(b) that the population variance of e_i is independent of the position of a group of values e_i, or the deviations are mutually independent along the line, and

(c) that the population variance of e_i along the least-squares curve is constant.

These conditions are rarely satisfied in hydrology, especially the latter two, because often the deviations e_i depend on y_i; usually they increase with an increase in y. In this case, by minimizing S of (12) through (13), greater weights are given to larger y - values than to smaller ones. Therefore, a bias is introduced. To alleviate this, various transformations of coordinates are used in an attempt to make the above conditions satisfied as closely as practically feasible. The logarithmic transformation is most commonly used in hydrology for this purpose, though it has the bias of giving larger weights to smaller y - values than to larger ones.

1.1.2.3.1.3 METHOD OF MOMENTS

The expected value E(x) of a random variable x is called the first population moment of x. In general, the expected value $E(x^r)$ is called the r-th population moment of x. Similarly, when dealing with a sample $x_1, x_2, \ldots, x_N$, the r-th sample moment is defined by

$$\mu_r = \frac{1}{N} \sum_{i=1}^{N} x_i^r \qquad \ldots. \ (14)$$

If the random variable x represents a probability distribution model with parameters $\alpha, \beta, \ldots,$ the population moments are functions of these parameters. Therefore, the moment parameter estimates are obtained by equating population moments and sample moments, and solving for the parameters. If m is the number of parameters to estimate, then the first m population and sample moments must be equated and solved simultaneously.

For instance, consider a probability distribution model given by

$$f(x;\alpha,\beta) = \frac{1}{\beta\sqrt{2\pi}} \exp\left| -(x-\alpha)^2/2\beta^2 \right|, -\infty \leq x \leq \infty \qquad \ldots. \ (15)$$

We would like to estimate the parameters α and β based on the observed sample $x_1, x_2, \ldots x_N$ which is believed to have a probability density function of (15). Since we have two parameters, then the first and the second population and sample moments must be equated. The first two population moments of (15) are

$$E(x) = \alpha \qquad \ldots. \ (16)$$

and

$$E(x^2) = \alpha^2+\beta^2 \qquad \ldots. \ (17)$$

Similarly, the first two sample moments of the sample are

$$\mu_1 = \frac{1}{N}\sum_{i=1}^{N} x_i \qquad \ldots. \ (18)$$

and

$$\mu_2 = \frac{1}{N}\sum_{i=1}^{N} x_i^2 \qquad \ldots. \ (19)$$

Equating the first population and sample moments of (16) and (18) respectively, we obtain the estimate $\hat{\alpha}$ of the parameter α as

$$\hat{\alpha} = \frac{1}{N}\sum_{i=1}^{N} x_i \qquad \ldots. \ (20)$$

Equationg the second population and sample moments of (17) and (19) respectively, we obtain the estimate $\hat{\beta}$ of β as

$$\hat{\alpha} + \hat{\beta}^2 = \frac{1}{N}\sum_{i=1}^{N} x_i^2$$

or

$$\hat{\beta} = \left| \frac{1}{N}\sum_{i=1}^{N} x_i^2 - \alpha^2 \right|^{\frac{1}{2}} \qquad \text{.... (21)}$$

The estimation of parameters by the method of moments is usually not difficult to obtain and it is simpler than the estimation by other methods. Often, the moment estimates are used as first approximations for the estimation by other methods. Except for the estimate of the mean, the moment estimates of other parameters are usually biased, although adjustment can be applied to make them unbiased. Moment estimates are asymptotically efficient when the underlying distribution is normal. For skewed variables though, the moment estimates generally are not asymptotically efficient.

1.1.2.3.1.4 METHOD OF MAXIMUM LIKELIHOOD

Let $f(x;\alpha,\beta,....)$ be a probability density function of x, with $\alpha,\beta,....$, parameters to be estimated. The product

$$L = \prod_{i=1}^{N} f(x_i;\alpha,\beta,....) \qquad \text{.....(22)}$$

is called the likelihood function of a sample of N observations from a population of the continuous variable x. In the case of a discrete variable with probability $p_i(x;\alpha,\beta,\ldots.)$, the likelihood function is

$$L = \prod_{i=1}^{N} p_i(x_i;\alpha,\beta,\ldots.), \qquad \ldots.. (23)$$

in which N is the sample size and $\alpha,\beta,\ldots.,$ are replaced by the estimates a,b,..... The estimate of $\alpha,\beta,\ldots.$ consists of determining a,b,.... from the sample data in such a way that the value of L in (23) is maximised. Since the natural logarithm of L, usually denoted as $\ell n(L)$, attains the maximum values at given a,b,.... as L does, the likelihood equation is

$$\ell n(L) = \ell n\left|\prod_{i=1}^{N} f(x_i;a,b,\ldots)\right| = \sum_{i=1}^{N} \ln\left|f(x_i;a,b,\ldots.)\right| \qquad \ldots.. (24)$$

while its partial derivatives in a,b,...., equated to zero and called the maximum likelihood equations, are

$$\frac{\partial}{\partial a}\ell n(L) = 0; \; \frac{\partial}{\partial b}\ell n(L) = 0; \qquad \ldots. (25)$$

These equations, represented by the system of equations in (25), which are of the same number as the number of parameters allow the computation of the estimates a,b,..... Therefore, the maximum likelihood equations are functions of the parameters to be estimated for a given sample. In this case, the estimates a,b,...., are the efficient estimates of α,β,.....,if the efficient estimate exist.

The solution of (25) must exclude all estimates which are constants, and retain only those roots of these equations that depend on the sample $x_1, x_2, \ldots x_N$. The estimates by the maximum likelihood method have the following properties.

(i) they are asymptotically unbiased

(ii) if efficient estimates exist for the parameters α,β,...., the maximum likelihood method produces them, and

(iii) if sufficient estimates exist, the solutions of the likelihood equations are functions of these sufficient estimates.

Often, the likelihood equation (23), gives solutions for estimates that converge to the true population values as the number of indepenent observations tends to infinity, with probability unity, so that the estimates are consistent.

The use of the least squares method, the graphical method and the probability weighted method which is reviewed in section 2.8.2, for parameter estimation can introduce further errors when an improper plotting position formula is used. The plotting position for the observed floods has been a matter of considerable controversy (Langbein, 1960). A general plotting position formula is of the form (Harter, 1971):

$$p = \frac{m-a}{n-a-b+1} \quad .... (26)$$

where m is the rank order of the sample values arranged in ascending order of magnitude, p is the plotting position which is also the probability of non-exceedance of the corresponding event, and a and b are constants which depend on the shape of the assumed probability density function. When a = b = 0, we get the commonly used Weibull plotting position formula, which is appropariate for symmetrical distributions. However, a value of a = b = 0.38 has been shown to be the best for asymmetrical distributions and for sample sizes of 15 or larger (Singh, 1980).

Having selected a distribution function and estimated its parameters, the question is how to use such distribution to estimate the design flood events. Chow (1951) proposed a general equation of the form,

$$x_T = \mu + K\sigma \quad .... (27)$$

where x_T is the required estimate of the design flood with return period T, and μ and σ population mean and standard deviation respectively, which can be replaced by the corresponding sample estimates, and K is the frequency factor. The frequency factor is a function of T and other population parameters. In most cases, a relationship can be derived between the return period T, the distribution paramters and the frequency factor K (Kite, 1977). Therefore, by use of (27), with sample parameter estimates, it is possible to extrapolate beyond the recorded flood magnitudes to obtain design floods of higher return periods.

1.2 OBJECTIVES OF THE STUDY

This project is entirely devoted to the advancement of the flood frequency approach using data from the major river basins in Kenya. Shen *et al* (1980) showed that at least three parameters are necessary for a probability distribution to model peak flows adequately. Hence, only those probability distributions with at least three parameter will be investigated herein.

The three parameter log-normal, Pearson Type 3, log-Pearson Type 3 and the Fisher Tippet-Type distribution in the real and logarithmic domains have been used very extensively all over the world for flood modelling. However, it has never been shown which of these distributions is in general better than the others, although some countries have adopted the use of one of them in preference to the others.

Recently, the Walter Boughton and the Wakeby distributions were proposed to have more advantages over the traditional distributions (Houghton, 1978b; Boughton, 1980). Nonetheless, no work has been done to give a convincing evidence to support such proposals. In this project, we wish to find which of the nine probability distributions, namely, the three parameter lognormal, the Pearson Type 3, the log-Pearson Type 3, the Fisher Tippet Type, the log-Fisher Tippet Type, the Walter Boughton, the log-Walter Boughton, the Wakeby and log-Wakeby distributions, is most favourable in general than the others using annual peak discharge data from the major drainage basins in Kenya. This will be achieved, mainly by finding the model which most frequently gives the minimum Akaike Information Criterion in the context of the existence of low and high outliers. The conventional goodness-of-fit tests, namely, the chi-square and the Smirnov-Kolmogorov tests will also be used to supplement the Akaike Information Criterion test.

In order to understand the optimum model identification process better, the properties and fitting procedures of both the traditional three parameter probability distributions and the new distributions will be discussed. Since the outlier problem in flood frequency model identification is a new approach in that field, a discussion of outlier detection and accommodation, with emphasis on the Akaike information theory, will also be given. The data used together with the climatology of the region of study will however, be reviewed first.

1.3 DATA USED

Kenya can be divided into five distinct drainage areas, namely the Lake Victoria basin, the Rift Valley drainage basin, the Tana River drainage basin, the Athi River drainage basin and the Uaso Nyiro drainage basin. Table 1 gives a summary of the characteristics of these drainage basins.

1.3.1 DATA COLLECTION

The Kenya Ministry of Water Development operates a fairly good network of river gauging stations (RGS) on rivers, streams and lakes of Kenya. There is also a good network of rainfall stations, some of which are maintained by the Ministry of Water Development

TABLE 1: CHARACTERISTICS OF THE DRAINAGE BASINS IN KENYA

Drainage basin	Area (km^2)	Mean annual rain-fall	Mean annual runoff	Runoff percent-	Predominant soil category
1. Lake Victoria	49,000	1245	149	12	poorly drained
2. Rift Valley	127,000	535	6	1	well drained soil
3. Athi River	70,000	585	19	3	well drained soil
4. Tana River	132,000	535	36	7	impeded drained soils
5. Uaso Nyiro	205,000	255	4	2	well drained soils
Kenya	583,000	500	25	5	well drained soils

but the majority are maintained by the Kenya Meteorological Department. The quality of the data on river flows varies a great deal from station to station. At some of the major rivers, continuous records have been obtained from autographic water level recorders for a considerable number of years. However, at the majority of the river gauging stations, the water level is recorded manually by taking one or two readings daily on a staff gauge , and at some remote stations, the readings are taken even less frequently.

With the permission of the Director, Ministry of Water Development, the annual peak gauge heights were extracted from the River Gauging (RGS) files which are routinely updated by the Water Resources Department of the Ministry. In all, about 400 RGS files were scrutinized. However, only 60 river gauging stations were chosen. These were the stations which had the most reliable data (from the remarks on the return sheets); had at least 25 years of complete records and had well established river stage-discharge relationships (rating curves). Twenty eight stations were chosen from the Lake Victoria basin, seven from the Rift Valley basin, seven from Athi River basin, ten from Tana River basin and eight from the Uaso Nyiro basin. Figure 1 shows the location of these gauging stations in the five drainage basins in Kenya. Table 2 shows the years

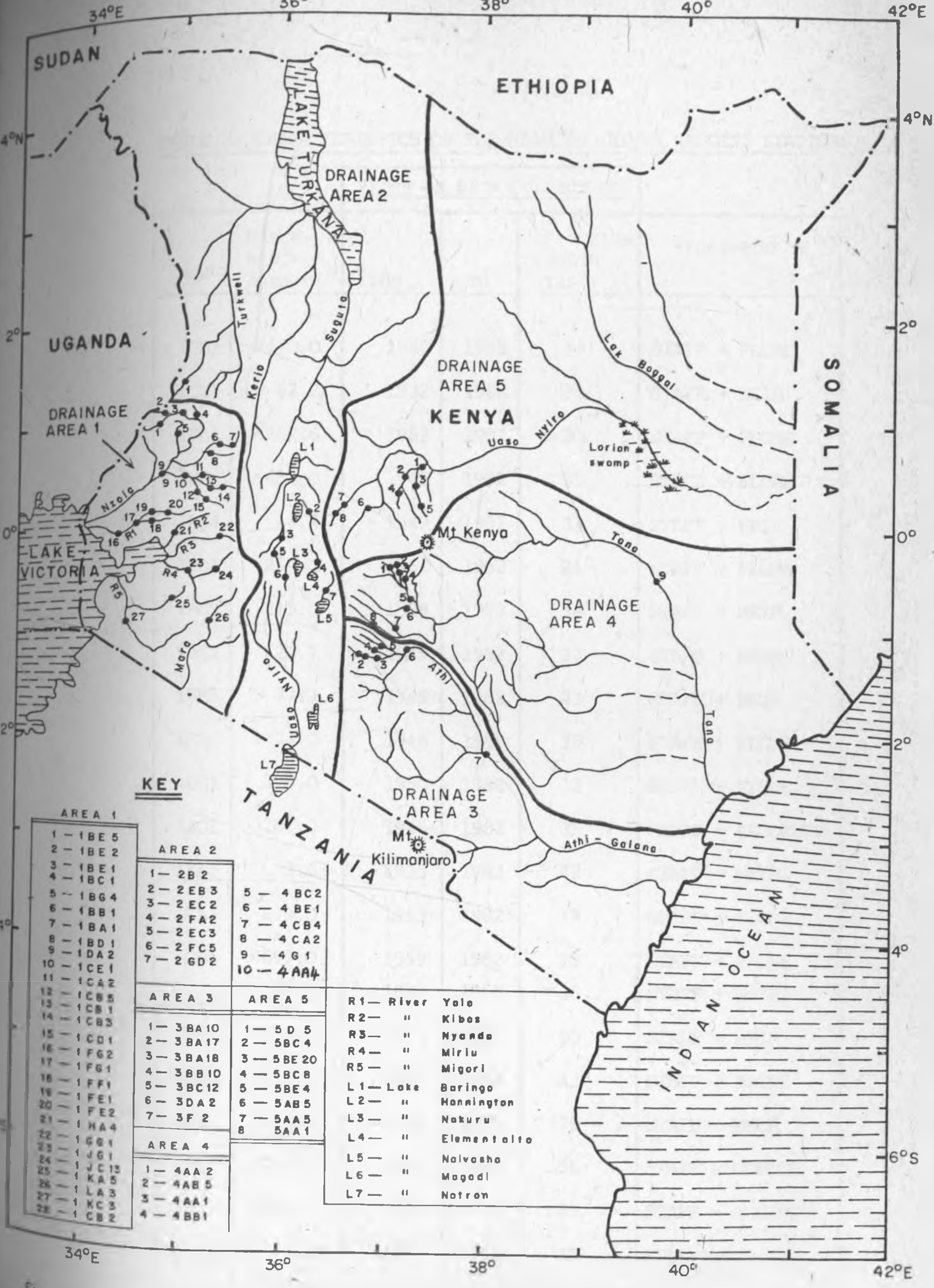

Fig 1 Drainage map of Kenya.

TABLE 2: CHARACTERISTICS OF THE SELECTED RIVER GAUGING STATIONS.

RGS	CATCH-MENT AREA Km^2	YEARS OF DATA COLLECTION FROM	TO	EFFECTIVE LENGTH (YRS.)	HYDROMETRY
1CE1	2440.0	1948	1982	34	STAFF + FLUME
1CB2	62.2	1932	1968	36	STAFF + WEIR
1BA1	262.6	1953	1982	30	STAFF + FLUME
1BE1	684.0	1948	1982	35	STAFF + FLUME
1BG4	54.4	1949	1982	33	STAFF + WEIR
1BD1	254.0	1961	1982	21	STAFF + FLUME
1BE1	715.0	1939	1973	34	STAFF + WEIR
1BE2	20.7	1949	1982	33	STAFF + WEIR
1BE5	77.7	1949	1982	33	STAFF + WEIR
1BB1	1474.0	1948	1982	35	STAFF + FLUME
1GG1	298.0	1950	1982	32	STAFF + FLUME
1JG1	3287.0	1946	1982	36	STAFF + RECORDER
1KA5	3.63	1935	1982	42	STAFF + WEIR
1CA3	679.0	1963	1982	18	STAFF + FLUME
1CB5	697.0	1959	1982	23	STAFF + FLUME
1CB3	51.8	1932	1968	36	STAFF + WEIR
1CD1	67.3	1948	1982	50	STAFF + WEIR
1CA2	717.0	1959	1982	23	STAFF + FLUME
1CB1	298.0	1933	1963	29	STAFF + WEIR
1DA2	8417.0	1947	1982	36	STAFF RECORDER
1FG1	2388.0	1947	1982	35	STAFF + RECORDER
1JC13	7.77	1957	1982	25	STAFF+WEIR + RECORDER

TABLE 2: CHARACTERISTICS OF THE SELECTED RIVER GAUGING STATIONS.

RGS	CATCHMENT AREA Km2)	YEARS OF DATA COLLECTION FROM	TO	EFFECTIVE LENGTHS (YRS)	HYDROMETRY
1KC3	3046.0	1951	1982	32	STAFF + FLUME
1FG2	2864.0	1958	1982	24	STAFF + RECORDER
1HA4	117.0	1924	1982	50	STAFF + WEIR
1FF	46.6	1959	1982	23	STAFF + FLUME
1FE2	1577.0	1961	1982	21	STAFF + FLUME
1FE1	1896.0	1960	1982	21	STAFF + FLUME
2B2	58.0	1946	1982	34	STAFF + WEIR
2EB3	331.0	1948	1982	33	STAFF + WEIR
2EC2	288.0	1931	1982	43	STAFF + WEIR
2FA2	155.0	1929	1982	40	STAFF + WEIR
2FC5	125.0	1941	1982	28	STAFF + WEIR
2EC3	48.0	1931	1982	41	STAFF + WEIR
2GD2	142.0	1935	1982	39	STAFF + FLUME
3BA10	64.7	1921	1982	41	STAFF + WEIR
3BA17	16.2	1931	1982	35	STAFF + WEIR
3BC12	357.0	1946	1982	36	STAFF + WEIR
3F2	4521.0	1952	1982	29	STAFF + FLUME
3DA2	5724.0	1956	1982	26	STAFF + RECORDER
3BA18	51.4	1939	1982	30	STAFF + WEIR
3BB10	41.4	1949	1982	27	STAFF + WEIR
4AA2	127.0	1948	1982	33	STAFF + WEIR
4CA2	518.0	1922	1982	50	WEIR + RECORDER
4BE1	414.0	1948	1982	34	STAFF + FLUME
4AA4	127.0	1947	1982	34	STAFF + WEIR

RGS = River Gauging Station number

TABLE 2: CHARACTERISTICS OF THE SELECTED RIVER GAUGING STATIONS.

	RGS	CATCHMENT AREA (Km^2)	YEARS OF DATA COLLECTION FROM	TO	EFFECTIVE LENGTHS (YRS)	HYDROMETRY
47	4BC2	2365.0	1950	1933	30	STAFF + RECORDER
48	4AB5	420.0	1950	1983	32	STAFF + WEIR
49	4G1	32892.0	1940	1983	45	STAFF + RECORDER
50	4CB4	316.0	1946	1983	37	WEIF + RECORDER
51	4AA1	254.0	1998	1983	35	STAFF + FLUME
52	4AA4	96.0	1948	1983	36	WEIF + RECORDER
53	5AA1	577.0	1945	1982	37	STAFF + WEIR + RECORDER
54	5AA5	157.0	1959	1982	23	STAFF + FLUME
55	5AB2	412.0	1960	1982	21	STAFF + FLUME
56	5BC4	1870.0	1962	1982	21	STAFF + RECORDER
57	5BC8	256.0	1960	1982	23	STAFF + WEIR + RECORDER
58	5BE4	62.0	1950	1982	21	STAFF + WEIR + RECORDER
59	5BE20	860.0	1960	1982	22	STAFF + RECORDER
60	5D5	4561.0	1950	1982	26	STAFF + RECORDER

RGS = River Gaunging Station

of data collection used in the project for each of the selected stations and also the mode of hydrometry utilized in each of the chosen stations.

The annual maximum gauge heights which were extracted as explained above were then converted into annual peak discharges by the use of the corresponding updated rating curves.

1.4 CLIMATOLOGY OF THE PROJECT AREA

In this section, the climatology of the project area will be briefly reviewed. The parameter discussed here include the rainfall and soils and their distribution over Kenya in general.

1.4.1 RAINFALL DISTRIBUTION IN KENYA

Kenya experiences a very wide range of climates, varying from semi-arid in the northern Kenya near Lake Turkana to wet over the highlands east of the Rift Valley and around Kisii and Kericho areas. As such, the average annual rainfall varies considerably from one place to the other. Figure 2 shows the distribution of average annual rainfall which is based on the Climatological Statistics for East Africa (East African Meteorological Department, 1975). Table 3 shows the

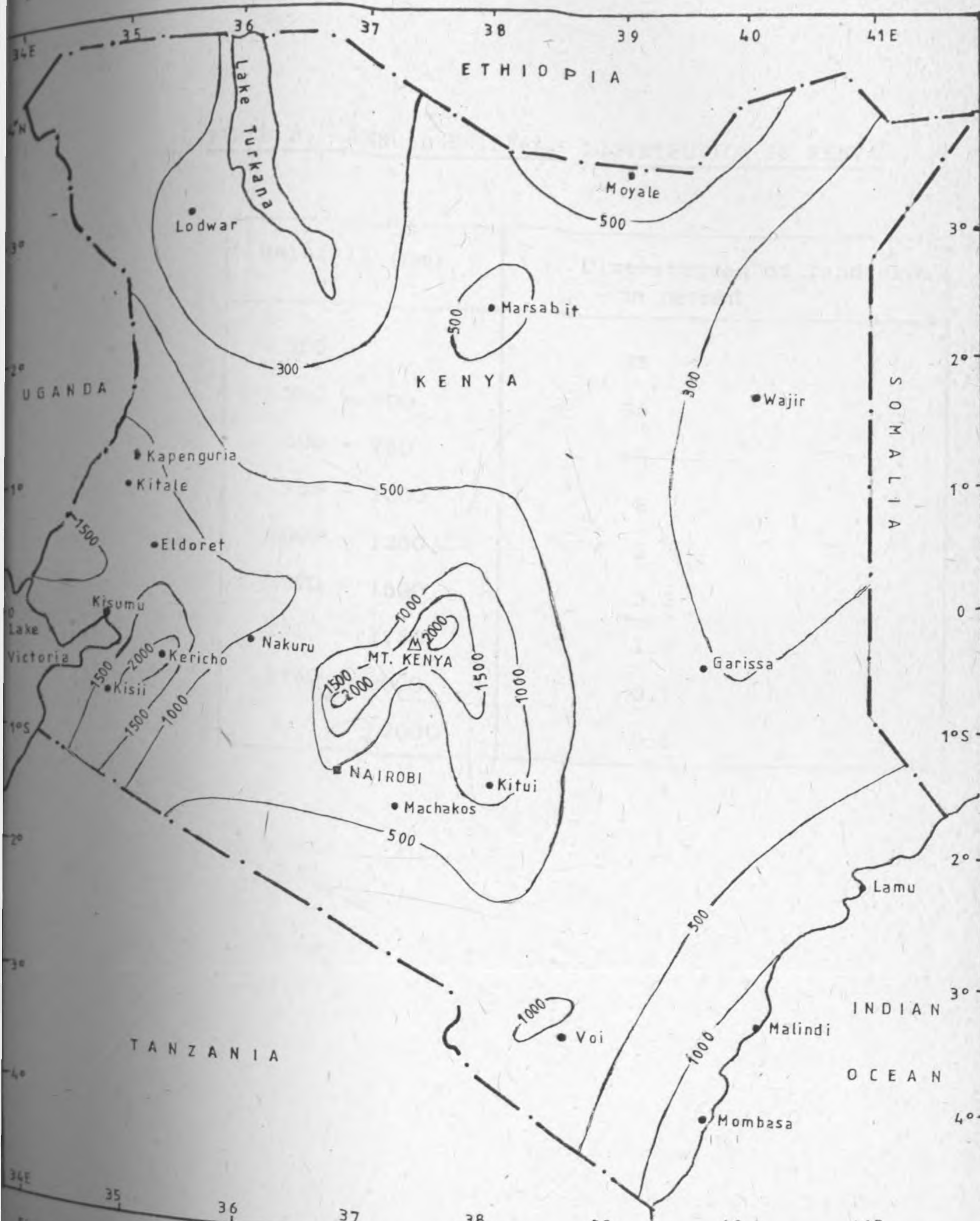

Fig. 2 Distribution of mean annual rainfall in Kenya (Data from "Climatological Statistics for East Africa" by East African Meteorological Department 1975)

TABLE 3: ANNUAL RAINFALL DISTRIBUTION IN KENYA

Rainfall (mm)	Distribution of land area in percent
< 250	27
250 - 500	35
500 - 750	18
750 - 1000	8
1000 - 1250	5
1250 - 1500	3.3
1500 - 1750	1.6
1750 - 2000	0.7
≥ 2000	0.1

relative distribution of the average annual rainfall in Kenya.

Climatological data shows that, broadly, the rains in Kenya exhibit two principal types of seasonal rainfall distribution with a broad peak between April and October which is typical in form for the Kenya highlands. Elsewhere, there are two peaks of varying relative importance; at Nairobi, the first peak (long rains) is greater than the second (short rains). This is demonstrated on Figure 3.

1.4.2. DISTRIBUTION OF SOILS IN KENYA

Potential flood runoff depends largely on the soil types. True black cotton soils being more or less impervious produce high surface runoff. Kenya consists of very many different soils as can be seen on the Explanatory Soil Map and Agro-Climatic zones map of Kenya by Sombroek and Braun (1982). However, we can group the soils of Kenya into three broad categories, namely, well drained soils, soils with impeded drainage and poorly drained soils. Figure 4 shows the distribution of these categories of soils in Kenya.

In the following chapter, we will now develop the methods of analysis which will be employed in this study.

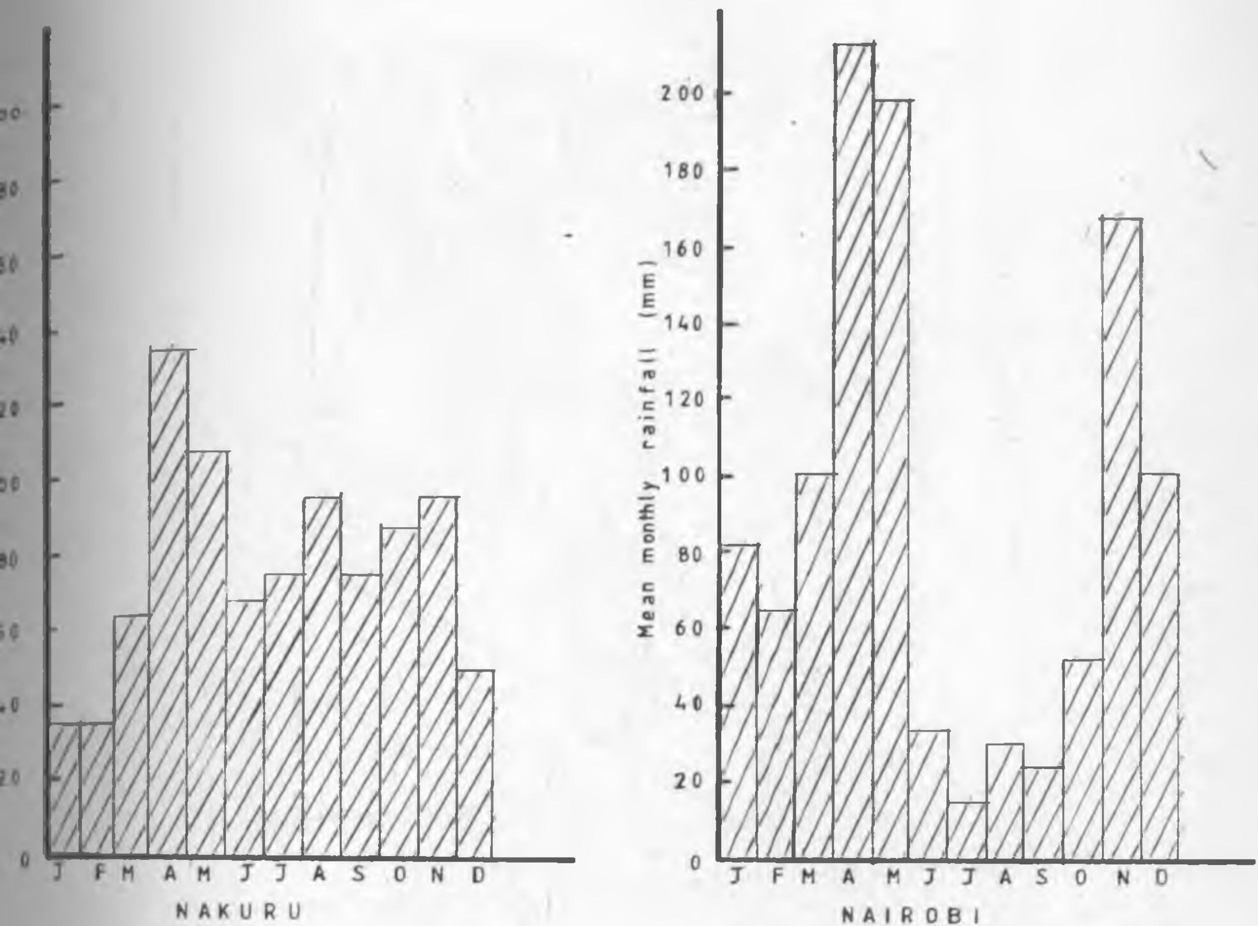

Fig. 3 Rainfall patterns in Kenya (Data from "Climatological Statistics for East Africa" by East African Meteorological Department 1975)

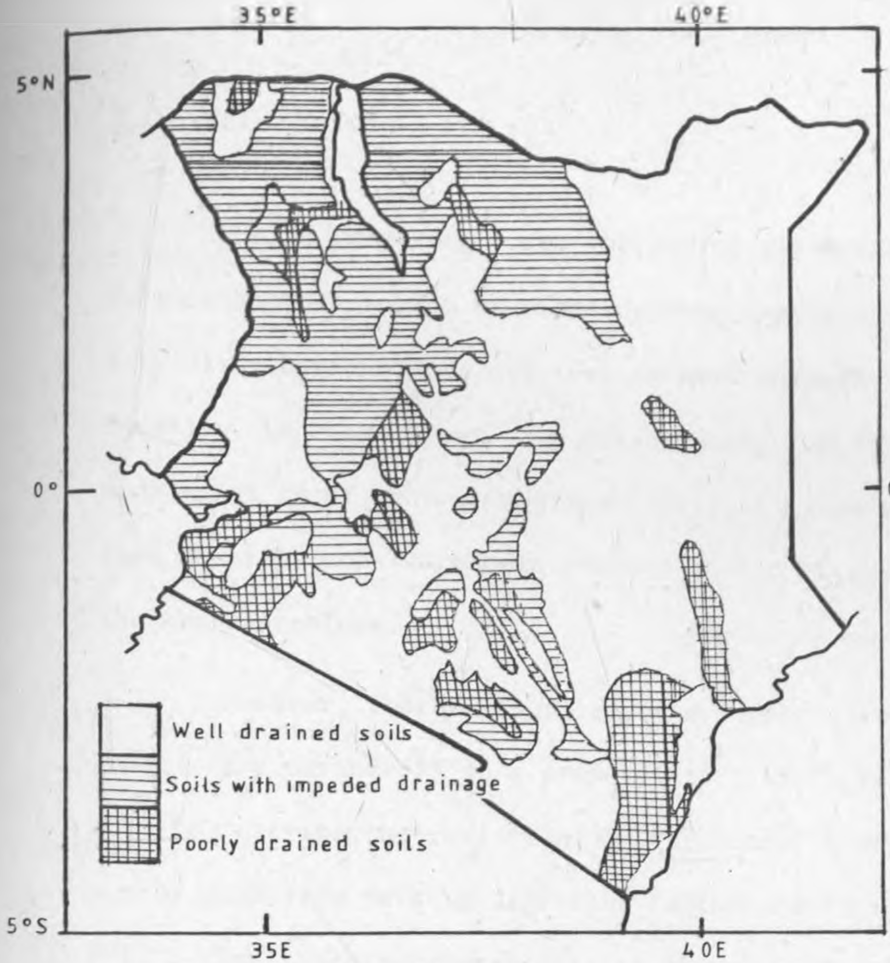

Fig. 4 Generalised soil zones of Kenya (Huddart and Woodward 1975)

CHAPTER 2

2.0 METHODOLOGY

For a very long time, the hydrologic designer has been faced with the problem of choosing from the many flood probability distributions, the one that is most suitable in a given location, and consequently in establishing the flood probability model that is of general applicability. A large number of peak flow distributions have been studied in the attempt to solve the above problems.

However, there is sufficient evidence to show that, at least three parameters in a probability distribution are necessary to describe the variations in the actual distribution of annual peak flow data in different catchments (Shen *et al*, 1981). The three parameter distributions are formed basically by adding a shift parameter (or a function of the shift parameter) to their two parameter conterparts. The location parameter, or a function of it, as the case might be, serves as a bound (lower or upper) for the random variable being fitted. The location parameter is not a constant for a given three parameter distribution, but assumes different values for each combination of the mean, variance and skewness coefficient of the random variable. Moreover, the bound becomes negative in certain instances of the three parameter lognormal, Weibull and log-Pearson distributions. This parameter has a significant physical interpretation when applied to hydrologic variables such as streamflow or precipitation. These variables cannot be negative unless they are made negative by some transformation, on one hand, and on the other, the imposition of an upper bound implying that there is a maximum certain value like flood flow, which is not physically meaningful (Rao, 1981).

Traditionally, the most commonly used three parameter flood frequency distributions are lognormal, the Pearson Type 3, the log-Pearson Type 3, the Fisher Tippet Type 2 and 3, and the log-Fisher Tippet Type 2 and 3. However, besides these three parameter traditional distributions, other relatively new probability distributions have also been proposed for use in flood frequency analysis. Two examples of such distributions are the Walter Boughton distribution (Boughton, 1980) and the Wakeby distribution (Houghton, 1978a). From these two distributions, a log-Walter Boughton distribution and a log-Wakeby distribution can be developed. The following sections give a discussion of each of these distribtuions which are relevant to our study.

2.1 THE THREE PARAMETER LOGNORMAL DISTRIBUTION

The three parameter distribution represents the normal distribution of the logarithms of the reduced variate $(x-x_o)$ where x_o is a lower or upper bound.

The probability density function, f(x), is given by

$$f(x) = \frac{1}{(x-x_o)\sigma_y\sqrt{2\pi}} e^{-\frac{\left[\ell n(x-x_o)-\mu_y\right]^2}{2\sigma_y^2}} \quad \ldots\ldots (28)$$

where μ_y and σ_y are the mean and standard deviation of the logarithms of $(x-x_o)$ respectively.

When the skewness coefficient of the actual observations (Variate x), γ_x, is positive, then $x_o < x < \infty$ and $(x-x_o)$ is distributed as the two-parameter lognormal with its origin shifted by a value x_o. When γ_x is negative, then $-\infty < x < x'_o$, and (x'_o-x) is distributed as a two parameter lognormal in the reverse direction with its origin shifted by x'_o. The three parameter lognormal distribution is unimodal and bell-shapped with a skew.

The selection of the parameter estimation procedure for the three parameter lognormal distribution is a matter of reliability (Salas and Smith, 1980). The method of moments is very simple but less accurate since the sample moments usually involve loss of information, hence, introducing bias. This occasionally results in a lower bound that is greater than the observed minimum value or in an upper bound that is smaller than the largest observation. In some rare cases, a solution is not obtainable at all. The method of maximum likelihood is more labourious if a computer is not available, but the results are more reliable.

The moment estimates of the parameters of three parameter lognormal distribution are given by

$$x_o = \hat{\mu}_x - \hat{\sigma}_x / z \qquad \text{.... (29)}$$

$$\mu_y = \ln(\hat{\sigma}_x / z) - \tfrac{1}{2}\ln(z^2+1) \qquad \text{.... (30)}$$

and

$$\hat{\sigma}_y = \left[\ln(z^2+1)\right]^{\frac{1}{2}} \qquad \text{.... (31)}$$

where $\hat{\mu}_x$ and $\hat{\sigma}_x$ are the sample estimates of the mean and the standard deviation of the actual observations, respectively, and

$$z = \frac{1-w^{2/3}}{w^{1/3}} \qquad \text{.... (32)}$$

where

$$w = \left[-\gamma_x + (\gamma_x^2+4)^{\frac{1}{2}}\right]/2 \qquad \text{.... (33)}$$

and γ_x is the skewness coefficient of the actual observations.

On the other hand, the maximum likelihood estimates of the three parameter lognormal distribution can be obtained as follows:

The log-likelihood function of the three parameters lognormal distribution is of the form (Salas and Smith, 1980)

$$LL(x;\mu_y,\sigma_y,x_o) = -\sum_{i=1}^{n} \ln(x_i - x_o) - \frac{n}{2}\ln(2\pi) - n\ln(\sigma_y) - \{\sum_{i=1}^{n} \left[\ln(x_i - x_o) - \mu_y\right]^2 / (2\sigma_y^2)\} \qquad \text{.... (34)}$$

where n is the size of the data sample. Differentiating the log-likelihood function $LL(\mathbf{x};\mu_y,\sigma_y,x_o)$ with respect to μ_y,σ_y and x_o respectively, and equating the resulting expressions to zero, we get a set of three simultaneous equations given by

$$\frac{\partial LL}{\partial \mu_y} = \{\sum_{i=1}^{n}\left[\ell n(x_i-\hat{x}_o)-\hat{\mu}_y\right]^2\}\ \hat{\sigma}_y^{2} = 0$$

$$\frac{\partial LL}{\partial \sigma_y} = -\ \frac{n}{\hat{\sigma}_y} + \{\sum_{i=1}^{n}\left[\ell n(x_i-\hat{x}_o)-\mu_y\right]^2\}/\hat{\sigma}_y = 0 \qquad \ldots. (35)$$

$$\frac{\partial LL}{\partial x_o} = \sum_{i=1}^{n}(x_i-\hat{x}_o)^{-1}(\hat{\mu}_y-\hat{\sigma}_y) - \sum_{i=1}^{n}(x_i-\hat{x}_o)^{-1}\ell n(x_i-\hat{x}_o) = 0$$

The maximum likelihood estimates $\hat{\mu}_y$, $\hat{\sigma}_y$ and $\hat{x}_o$ of μ_y, σ_y and x_o in (28) can be obtained from (35) and are given by

$$\hat{\mu}_y = \frac{1}{n}\sum_{i=1}^{n}\ell n(x_i-\hat{x}_o) \qquad \ldots. (36)$$

$$\hat{\sigma} = \frac{1}{n}\sum_{i=1}^{n}\left[\ell n(x_i-\hat{x}_o)-\hat{\mu}_y\right]^2 \qquad \ldots. (37)$$

and

$$F(x_o) = \left[\sum_{i=1}^{n}(x_i-\hat{x}_o)^{-1}\{\frac{1}{n}\sum_{i=1}^{n}\ell n^2(x_i-\hat{x}_o)\} - \frac{1}{n}\sum_{i=1}^{n}\ell n(x_i-\hat{x}_o)\right]^2$$

$$-\frac{1}{n}\sum_{i=1}^{n}\ell n(x_i-\hat{x}_o)\} + \sum_{i=1}^{n}\{\left[\ell n(x_i-\hat{x}_o)\right]/(x_i-\hat{x}_o)\} = 0$$

$$\ldots. (38)$$

Note that (38) is an implicit function, $F(\hat{x}_o)$ of $\hat{x}_o$ which is obtained by substituting (36) and (37) into the expression corresponding to $\partial LL/\partial x_o = 0$ in (35). The estimate $\hat{x}_o$ is usually obtained from (38) by an iterative procedure, such as the Newton's tangent method (Salas and Smith, 1980). This method requires the first derivative $F'(\hat{x}_o)$ of $F(\hat{x}_o)$ of (38), with respect to $\hat{x}_o$, that is

$$F'(\hat{x}_o) = F(b-c^2-c-1)+\frac{2a}{n}\left[-d+a(c+\tfrac{1}{2})\right]+e = 0 \quad \ldots.\ (39)$$

where a,b,c,d and e are given by

$$a = \sum_{i=1}^{n}(x_i-\hat{x}_o)^{-1} \quad \ldots.\ (40)$$

$$b = \frac{1}{n}\sum_{i=1}^{n}\ell n^2(x_i-\hat{x}_o) \quad \ldots.\ (41)$$

$$c = \frac{1}{n}\sum_{i=1}^{n}\ell n(x_i-\hat{x}_o) \quad \ldots.\ (42)$$

$$d = \sum_{i=1}^{n}\{\left[\ell n(x_i-x_o)\right]/(x_i-x_o) \quad \ldots.\ (43)$$

and

$$e = \sum_{i=1}^{n}\{\left[\ell n(x_i-\hat{x}_o)\right]/(x_i-\hat{x}_o)^2\} \quad \ldots.\ (44)$$

The estimate $\hat{x}_o$ is then obtained by successive updation in every iteration k according to the equation

$$\hat{x}_o^{k+1} = \hat{x}_o^k - F'(\hat{x}_o^k)/(\hat{x}_o^k) \quad \ldots. \ (45)$$

and $\hat{x}_o^{k+1}$ denotes the updated estimate of $\hat{x}_o$ after the k-th iteration. The initial estimate of $\hat{x}_o$ is usually taken to be that given by the method of moments. The stopping criterion of the iterations in (45) is constructed such that the absolute value of $(\hat{x}_o^{k+1} - \hat{x}_o^k)$ is less than the required error limit. The so obtained estimate of $\hat{x}_o$ is then used in (36) and (37) to obtain the estimates of $\hat{\mu}_y$ and $\hat{\sigma}_y$ respectively.

When the skewness coefficient of the data sample is negative, then $\hat{x}_o$ is an upper bound and (34) through (44) have to be modified accordingly. Nevertheless, the approach is similar to the one described above.

Once the estimates of the parameters of the three parameter lognormal distribution have been estimated, it is sometimes required to estimate numerically the emperical probabilities and the cumulative probabilities of the sample values. The computation of the probability density function, f(x), for a given observation involves a direct substitution of the parameters and the required observation value in (28). However, the computation of the cumulative probability or the cumulative distribution function, F(x), is a bit more difficult. By definition, the cumulative distribution, F(x), of the three parameter lognormal distribution is given by

$$F(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}\sigma_y(x-x_o)} \exp\{-\tfrac{1}{2}\left[\frac{\ell n(x-x_o)-\mu_y}{\sigma_y}\right]\}dx \quad \ldots\ldots (46)$$

An indirect way of evaluating the cumulative probabilities for the three parameter lognormal distribution is to transform the original x-values into logarithmic vaues as $y = \ell n(x-x_o)$ and use the procedure for the normal distribution. This works as follows:

Since $dx = (x-x_o)dy$, one can express (46) as

$$F(x) = \frac{1}{\sqrt{2\pi}\sigma_y} \int_{-\infty}^{x} \exp\left[-\tfrac{1}{2}(\frac{y-\mu_y}{\sigma_y})\right]^2 dy \quad \ldots. (47)$$

By using a standardized variate $\mu = (y-\sigma_y)/\sigma_y$, then (47) becomes

$$F(x) = F(u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u} \exp\left[-\tfrac{1}{2}u^2\right] du \quad \ldots. (48)$$

which is an integral of the standard normal variate. Abramowitz and Stegun (1965) gave several approximations for the cumulative distribution function, F(u), of the standard normal variate. A polynomial approximation with error less than 10^{-5} is:

$$F(u) \approx 1-f(u)(0.4318v-0.12017v^2+0.9373v^3) \quad .... \ (49)$$

where

$$f(u) = \frac{1}{\sqrt{2\pi}}\left[ex \ -\tfrac{1}{2}u^2\right] \quad .... \ (50)$$

and v is defined for $u \geq 0$ as

$$v = 1/\left[1+0.33267u\right] \quad .... \ (51)$$

A similar polynomial approximation was given by Hastings (1955) and has been used by IBM (1968) and ICL(1976). This approximation is:

$$F(u) = 1-f(u)w\{0.3193815+w\left[-0.3565638+w\left[1.781478 + w(-1821256+1.330274)\right]\right]\} \quad .... \ (52)$$

where w is defined for $u \geq 0$ as:

$$w = 1/(1+0.2316419u) \quad .... \ (53)$$

In both approximations, the cumulative probability is is 1-F(u) if u<0. Tables of cumulative probabilities for the standard normal variate are also available in many statistics , probability and other related books. In this project, the Hastings (1955) approximation is used.

2.2 THE PEARSON TYPE 3 DISTRIBUTION

The gamma distribution with three parameters or the so called Pearson Type 3 distribution, has a probability density function

$$f(x) = \frac{1}{\alpha\Gamma(\beta)}\left[\frac{x-x_o}{\alpha}\right]^{\beta-1} \exp\left[-(\frac{x-x_o}{\alpha})\right] \quad .... (54)$$

where α and β are the scale and shape parameters respectively and x_o is the location parameter. $\Gamma(\beta)$ is the gamma function given by

$$\Gamma(\beta) = \int_o^{\infty} z^{\beta-1} e^{-z} dz \quad .... (55)$$

The best method of estimating the parameters x_o, α and β of the Pearson Type 3 distribution is by the method of maximum likelihood. However, mainly due to its simplicity, the method of moments is occasionaly utilized to estimate these parameters. The moment estimates of the parameters x_o, α and β of the Pearson Type 3 distribution can be obtained as:

$$\hat{x}_o = \hat{\mu}-2\hat{\sigma}/\hat{\gamma} \quad .... (56)$$

$$\hat{\alpha} = \sigma\gamma/2 \quad .... (57)$$

and

$$\hat{\beta} = (2/\hat{\gamma})^2 \quad .... (58)$$

where $\hat{\mu}, \hat{\sigma}$ and $\hat{\gamma}$ are the sample estimates of the mean, standard deviation and the skewness coefficient respectively. Although the method of moments is relatively simple, it does not always guarantee a solution since in some instances, x_o, the location parameter, can be computed to be within the range of the actual observations. This of course is not realistic and in such cases, the moment parameter estimates are assumed to be not obtainable.

Generally, the method of maximum likelihood gives the most reliable estimates. However, this method is much more involved especially if a computer is not available.

The log-likelihood function of the Pearson Type 3 distribution is (Salas and Smith, 1980)

$$LL(x;x_o,\alpha,\beta) = n\ell n\left[\Gamma(\beta)\right] - \frac{1}{\alpha}\sum_{i=1}^{n}(x_i - x_o) +$$

$$(\beta-1)\sum_{i=1}^{n}\ell n(x_i - x_o) - n\beta\ell n(\alpha) \qquad \text{.... (59)}$$

where n is the number of observations. The partial derivatives $LL(x;x_o,\alpha,(\beta)$ with respect to α, β and x_o are:

$$\frac{\partial LL}{\partial \alpha} = \frac{1}{\alpha^2}\sum_{i=1}^{n}(x_i - x_o) - \beta n/\alpha \qquad \text{.... (60)}$$

and

$$\frac{\partial LL}{\partial \beta} = -n\left[\Gamma'(\beta)/\Gamma(\beta)\right] + \sum_{i=1}^{n} \ln(x_i - x_o) - n\ln(\alpha) \qquad \ldots\ldots (61)$$

and

$$\frac{\partial LL}{\partial x_o} = n/\alpha - (\beta-1)\sum_{i=1}^{n}\left[1/(x_i - x_o)\right] \qquad \ldots. (62)$$

Simultaneous solution of $\partial LL/\partial\alpha = 0$, $\partial LL/\partial\beta = 0$ and $\partial LL/\partial x_o = 0$ yields expressions for $\hat{\alpha}$ and $\hat{\beta}$ in terms of x_o given by

$$\hat{\alpha} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \hat{x}_o) - n\left[\sum_{i=1}^{n}(x_i - \hat{x}_o)^{-1}\right]^{-1} \qquad \ldots. (63)$$

$$\hat{\beta} = \{1 - n^2\left[\sum_{i=1}^{n}(x_i - \hat{x}_o)\sum_{i=1}^{n}\left(\frac{1}{x_i - \hat{x}_o}\right)\right]^{-1}\}^{-1} \qquad \ldots. (64)$$

and an implicit expression for $\hat{x}_o$ given by

$$F(\hat{x}_o) = n\psi(\hat{\beta}) + \sum_{i=1}^{n} \ln(x - \hat{x}_o) - n\ln(\hat{\alpha}) = 0 \qquad \ldots.(65)$$

where $\psi(\hat{\beta})$ is the so-called digamma function which can be computed using the asymptotic expansion as in Condie and Nix (1975), given by

$$\psi(\hat{\beta}) = \Gamma'(\beta) = \ln(\hat{\beta}+2) - \frac{1}{2(\hat{\beta}+2)^2} - \frac{1}{12(\hat{\beta}+2)^2} + \frac{1}{120(\hat{\beta}+2)^4} - \frac{1}{256(\hat{\beta}+2)^6} - \frac{1}{\hat{\beta}+1} - \frac{1}{\hat{\beta}} \qquad \ldots. (66)$$

where $\gamma'(\hat{\beta}) = 1/\hat{\beta}-\psi'(\beta)$

where ψ' is the derivative of $\psi(\hat{\beta})$ with respect to β and is called the trigamma function, which can be obtained from (66) as

$$\psi'(\hat{\beta}) = \frac{1}{\hat{\beta}+2} + \frac{1}{2(\hat{\beta}+2)^2} + \frac{1}{6(\hat{\beta}+2)^3} - \frac{1}{30(\hat{\beta}+2)^3} +$$

$$\frac{1}{42(\hat{\beta}+2)^7} - \frac{1}{30(\hat{\beta}+2)^9} + \frac{1}{(\hat{\beta}+1)^2} + \frac{1}{\hat{\beta}^2} \quad \ldots. (67)$$

Equation (65) is usually solved by an iterative approximation procedure starting with the moment estimate of x_o in (63), (64) and (65). Every iteration of (65) returns to (63) and (64) to update the estimates of $\hat{\alpha}$ and $\hat{\beta}$. The procedure for the Newton's tangent iterative method is as follows:

The first derivative of $F(\hat{x}_o)$ with respect to $\hat{x}_o$ is given by

$$F'(\hat{x}_o) = n\psi'(\hat{\beta})(fd')/f^2-a-nh'/h \quad \ldots. (68)$$

where

$$a = \sum_{i=1}^{n} \frac{1}{x_i-\hat{x}_o} \quad \ldots. (69)$$

$$d = \sum_{i=1}^{n} \frac{1}{(x_i-\hat{x})_o^2} \quad \ldots. (70)$$

$$f = a-n^2/b \qquad \ldots. (71)$$

$$f' = d-n^3/b^2 \qquad \ldots. (72)$$

$$h = b/n-n/a \qquad \ldots. (73)$$

$$h' = -1+nd/a^2 \qquad \ldots. (74)$$

and

$$b = \sum_{i=1}^{n} (x_i-x_o) \qquad \ldots. (75)$$

Then

$$\hat{x}_o^{(k+1)} = \hat{x}_o^{(k)} - F'(\hat{x}_o^{(k)})/(F(\hat{x}_o^{(k)}) \qquad \ldots. (76)$$

where k is the iteration number. These iterations are carried out until $\left|\hat{x}_o^{(k)} - \hat{x}_o^{(k)}\right| \leq \varepsilon$, where ε is the required error limit. Note that the estimates of $\hat{\alpha}$ and $\hat{\beta}$ have to be updated with (63) and (64) after every iteration in (76).

For a given set of parameters, the probability density function, f(x), may be obtained directly from (54). In this case $\Gamma(\beta)$ is determined by a polynomial approximation given by Hastings (1955) and which has been programmed as a computer package, in many computer installations. In the ICL 2950 computers, this package is under the name of F4GAMMA. On the other hand, the the computation of the corresponding cumulative distribution functions require a numerical integration procedure since they cannot be expressed explicitly.

In this project, the numerical integration procedure by O'Hara and Smith (1969) and programmed by Salas and Smith (1980) is used.

2.3. THE LOG-PEARSON TYPE 3 DISTRIBUTION

A random variable x has a log-Pearson Type 3 distribution if the transformed variate $y = \ell n(x)$ has a Pearson Type 3 distribution. Let $f(\cdot)$ denote a probability density function. Then

$$f(y) = \frac{1}{\alpha\Gamma(\beta)}\left[\frac{y-y_o}{\alpha}\right]^{\beta-1}\exp\left\{-\left[\frac{y-y_o}{\alpha}\right]\right\} \quad \ldots\ldots (77)$$

and

$$f(x) = \frac{1}{x\alpha\Gamma(\beta)}\left[\frac{\ell n(x)-y_o}{\alpha}\right]^{\beta-1}\exp\left\{-\left[\frac{\ell n(x)-y_o}{\alpha}\right]\right\} \quad \ldots\ldots (78)$$

where α, β and y_o are the scale, shape and location parameters and $f(y)$ and $f(x)$ are the Pearson Type 3 and log-Pearson Type 3 probability density functions respectively.

If $\alpha > 0$, then, the Pearson Type 3 distribution is positively skewed and $y_o \leq y < \infty$. In this case, the log-Pearson Type 3 distribution is also positively skewed and $\exp(y_o) \leq x < \infty$. If $\alpha < 0$, the Pearson Type 3 distribution is negatively skewed and $-\infty < y \leq y_o$. In this case, the log-Pearson distribution is either positively skewed or negatively skewed depending on the computed values of α and β.

There are about four methods available for the estimation of the parameters of the log-Pearson Type 3 distribution, namely, the direct method of moments, the method of moments applied to the Pearson Type 3 distribution, the method of mixed moments and the method of maximum likelihood.

For the direct method of moments, Bobee (1975), Kite (1977) and Salas and Smith (1980) have shown that, the application of the moment generating function technique to the log-Pearson Type 3 distribution of (78), the scale parameter α can be estimated from

$$\hat{\alpha}=1/(A+3.) \qquad \text{.... (79)}$$

where

$$A=0.23019+1.65262C+0.20911C^2-0.04557C^3 \qquad \text{.... (80)}$$

when

$3.5 \leq B \leq 6$ and

$$A = 0.45157+1.99955C \qquad \text{.... (81)}$$

when

$3.0 \leq B \leq 3.5$.

In these equations, C is computed from

$$C = 1./(B-3) \qquad \text{.... (82)}$$

and

$$B = \frac{\ell n(\mu_3)-3\ell n(\mu_1)}{\ell n(\mu_2)-2\ell n(\mu_1)} \quad .... (83)$$

where μ_1, μ_2 and μ_3 are the first three sample moments of the variate x about the origin. Similarly, the shape parameter β and the location parameter y_o are estimated by

$$\beta = \frac{n(\mu_2)-2\ell n(\mu_1)}{\ell n(1-\hat{\alpha})^2-\ell n(1-2\hat{\alpha})} \quad .... (84)$$

and

$$y_o = \ell n(\mu_1)+\ell n(1-\hat{\alpha}) \quad .... (85)$$

When

$3 \leq B \leq 6$, this

method of moments yield the most reliable estimates (Bobee,1975). However, when $B \leq 3$ or $B \geq 6$, then the above method is not accurate and therefore other estimation methods have to be sought.

We have seen that, if x is distributed as the log-Pearson Type 3 distribution, then the variate $y = \ell n(x)$ has a Pearson Type 3 distribution and the logarithmically transformed data can also be utilised to estimate the parameters α, β and y_o. As before, the moment estimates of these parameters can be obtained as:

$$\hat{\alpha} = \sigma_y \hat{\gamma}_y / z \qquad \text{.... (86)}$$

$$\hat{\beta} = (2/\hat{\gamma}_y)^2 \qquad \text{.... (87)}$$

and

$$\hat{y}_o = \hat{\mu}_y - 2\hat{\sigma}_y/\hat{\gamma}_y \qquad \text{.... (88)}$$

and $\hat{\mu}_y$, $\hat{\sigma}_y$ and $\hat{\gamma}_y$ are the sample estimates of the mean, standard deviation and skewness coefficient respectively for the log-transformed variate $y = \ell n(x)$.

Just like it is the case with the method of moments applied to the Pearson Type 3 distribution for the untransformed variate x, this method has the major disadvantage that, occasionally, $\hat{y}_o$ can be estimated to be within the range of the observations y_i, which is not realistic. Consequently, this method, though simple, is usually not used to estimate the parameters α, β and y_o.

In order to avoid the uncertainities in the parameter estimation, which are associated with the use of the skewness coefficient, Rao (1980b) showed that the parameters of the log-Pearson Type 3 distribution can be estimated by use of only the first two moments of the data by mixing the real and logarithmic moments. He described two procedures for fitting data to the log-Pearson Type 3 distribution;

(i) the method in which the mean and variance of the real data and the mean of the logarithmic data are used; a procedure which can be called MXM1, and

(ii) the method in which the mean and variance of the logarithmic data and the mean of the real data are used. He called this method MXM2.

The objective of the method of mixed moments MXM1 is to estimate the parameters α, β and y_o so that the population mean, μ_x, and variance, σ_x^2, of the log-Pearson Type 3 distribution are the same as the corresponding untransformed sample parameters $\bar{x}$ and S_x^2. Further, this method attempts to make the population mean, μ_y, of the corresponding Pearson Type 3 distribution (in log-Domain) equal to the sample mean, $\bar{y}$, of the log-transformed data. The procedure for estimating the parameters of the log-Pearson Type 3 distribution by MXM1 is as follows:-

(i) choose an initial value for the population skewness coefficient, γ_x, say CS_x - the skewness coefficient of the untransformed data.

(ii) estimate the parameters α, β and y_o by the direct method of real moments with $\mu_x = \bar{x}$, $\sigma_x^2 = S_x^2$ and γ_x as chosen

(iii) calculate μ_y by $\mu_y = \hat{\alpha}\hat{\beta} + \hat{y}_o$

(iv) if μ_y differs significantly from $\bar{y}$, repeat steps (ii) and (iii) by varying γ_x until $\mu_y \simeq \bar{y}$.

The objective of the method of mixed moments MXM2, is to estimate the parameters α, β, and y_o so that $\mu_y = \bar{y}$, $\sigma_y = S_y^2$ and $\mu_x = \bar{x}$. This method is essentially a variation of the method of moments applied to the Pearson Type 3 distribution in the log-domain with the exception that the population skewness coefficient, γ_y, of the log-transformed data is selected on the basis of the untransformed sample mean $\bar{x}$ instead of CS_y, the skewness of the log transformed data. The relationship between μ_y, σ_y and γ_x can be shown (Rao 1980b) to be;

$$\mu_y = 4\left[\ln(1-\sigma_y\gamma_y/2)+\sigma_y\gamma_y/2\right]/\gamma_y^2+\ln(\mu_x) \qquad \ldots. (89)$$

If μ_y, σ_y and μ_x are known (from the sample estimates), the only unknown variable in (89) is γ_y which can be solved for by the Newton-Raphson method. Then the values of α, β and y_o can be obtained from (86) through (88).

Rao (1980b) also showed that, the method of mixed moments MXM1, has in general, superior statistical properties than all the other moment methods.

The method of maximum likelihood gives the same results as those obtained when the method of maximum likelihood is applied to the Pearson Type 3 distribution in the log-domain. Thus, the procedure is parallel to that outlined in the discussion of the maximum likelihood method with the Pearson Type 3 distribution when x_o is replaced by y_o in (59).

The United States Water Resources Council (1977) recommended that the log-Pearson Type 3 distribution be adopted as the standard flood-frequency distribution by the United States government agencies. Being a three parameter distribution operating upon the logarithms of the flood magnitudes, the log-Pearson type 3 distribution would appear to be an extremely versatile distribution. However, its applicability in hydrology is strictly limited (Kite, 1977).

Bobee (1975) derived the density function of the log-Pearson Type 3 distribution and studied at length its mathematical and statistical properties. Bobee (1975) presented in a general fashion the various forms of the log-Pearson Type 3 distribution. Rao (1980a) extended Bobees work and presented the properties and results of the log-Pearson Type 3 distribution on the basis of its population parameters. Bobee (1975) and Rao (1980a) showed that the scale parameter, α and the shape parameter, β, govern the overall geometric shape of the distribution, which takes four basic forms, namely, J- shape, reverse J-shape, bell (unimodal shape) and U-shape. In addition, several transitional shapes varying from one basic shape to another occur. Figure 5 shows some of the various forms of the log-Pearson Type 3 distribution for some critical parameter values. The critical parameter values which produce different shapes are summarised in Table 4.

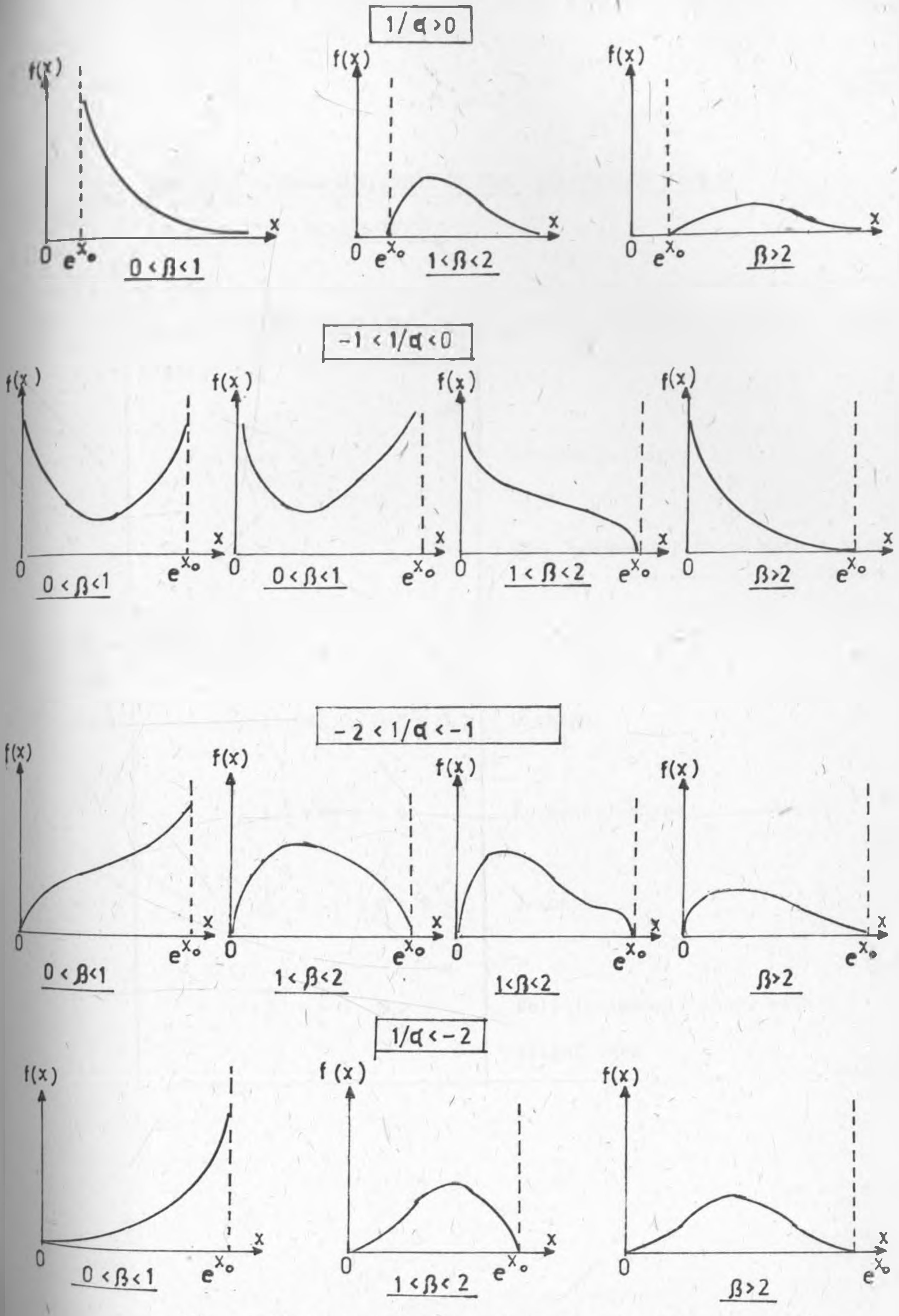

Fig. 5 Density function of the log Pearson distribution

(Bobee 1975)

TABLE 4: CHARACTERISTICS OF THE LOG-PEARSON TYPE 3 DISTRIBUTION (BOBEE 1975)

Parameter value	Form of the distribution
Class A: $\frac{1}{\alpha} > 0$	
$0 < \beta \leq 1$	Reverse J-shape
$\beta > 1$	Bell (unimodal) shape with a skew
Class B: $\frac{1}{\alpha} < 0$	
-1 $\frac{1}{\alpha}$ 0; $0<\beta \leq 1$	U-shape
-1 $\frac{1}{\alpha} < 0$; $\beta > 1$	Reverse J-shape
$\frac{1}{\alpha} < -1$; $0 < \beta \leq 1$	J-shape
$\frac{1}{\alpha} < -1$; $\beta > 1$	Bell (unimodal) shape with a slight skew

Kite (1977) showed that for flood frequency analysis, the only shape of interest is that which is unimodal, continuous from zero to infinity, has an infinitely high order or smooth contact with a lower limit and unbounded at the upper limit. The log-Pearson Type 3 distribution falls within these criteria only when $\beta > 1$ and $\frac{1}{\alpha} > 0$. When the coefficient of skewness of the untransformed data is negative, this corresponds to a negative value of α which is not suitable. Reich (1972) gave examples of the application of the log-Pearson Type 3 distribution to samples with negative skew in which the computed upper bounds were found to be lower than the maximum observed events.

The log-Pearson Type 3 distribution is related to the Pearson Type 3 distribution (in log-domain) in that, if x is log-Pearson Type 3, then $y = \ell n(x)$ is Pearson Type 3 distributed. Hence, the probability density function of x, f(x) is related to the probability density function fo y, f(y) by

$$f(x) = \frac{1}{x} f(y) \qquad \dots . (90)$$

Equation (90) is used to find the log-Pearson Type 3 density function f(x) based on the Pearson Type 3 density of $y = \ell n(x)$ as described earlier. As for the cumulative distribution function of x, F(x), the following relation holds:

$$F(x) = \int_0^x f(x)dx = dx = F(y) = \int_0^{\ell n(x)} f(y)dy \qquad \dots . (91)$$

where F(y) is the cumulative distribution function of y (the cumulative distribution function of the Pearson Type 3 in the log-domain). Thus, from (91) we can see that F(x) may be obtained by integrating the Pearson Type 3 density in the log-domain.

2.4 THE FISHER TIPPET TYPE DISTRIBUTIONS

The Fisher Tippet type distribution of the first, second and third types, in other words the Gumbel, the Fisher Tippet Type 2 and Fisher Tippet Type 3 distributions are well known as models for extreme observations such as annual maximum wind speeds and floods. The distribution functions of these types of probability functions are summarised in Table 5.

TABLE 5: THE FISHER TIPPET TYPE DISTRIBUTIONS

Distribution name	Cumulative distribution	Range of variate
Gumbel	$F(x)=\exp\left[-\exp(x-x_o)/\alpha\right]$	$\alpha<x<\infty$
Fisher Tippet Type 2	$F(x)=\exp\{-\left[1-(x-x_o)/\alpha\right]^{\frac{1}{\beta}}\}$	$\beta<0, x\geq x_o+\alpha/\beta$
Fisher Tippet Type 3	$F(x)=\exp\{-\left[1-(x-x_o)/\alpha\right]^{\frac{1}{\beta}}\}$	$\beta>0, x<x_o+\alpha/\beta$

The Gumbel distribution depends on a location parameter x_0 and a positive valued scale parameter α. Fisher Tippet Type 2 and Type 3 distributions depend on these two parameters and also on a shape parameter β. It can be shown that, the Gumbel distribution correponds to the limit case of $\beta \to 0$ (Salas and Smith, 1980).

There is a practical way of determining, the type of general extreme value distribution for a given sample. Since the theoretical skewness coefficient of a Gumbel variate is 1.14 ($\beta = 0$), for a skewness coefficient greater than 1.14, the distribution is Fisher Tippet Type 2 while for the Fisher Tippet Type 3, the skewness coefficient of the sample is less than 1.14.

Note that none of the general extreme value distributions can be used individually in all cases due to their individual requirements on the sample skewness coefficient. Furthermore, the Fisher Tippet Type 3 distribution is mostly appropriate for the smallest values analyses since it has an upper bound.

There are about five methods available for the etstimation of the parameters of the general extreme values distributions, namely, the method of moments, the method of maximum likelihood, the method of sextiles, the graphical method and the method of probability weighted moments. The detail of these methods can be found in the works by Gumbel

(1958) and WMO (1966). However, the most commonly used methods of parameter estimation are the method of moments and the method of maximum likelihood. The application of these two methods is described below.

The method of moments can be used as follows:-

Using the sample skewness coefficient, $\hat{\gamma}$, the parameter, β, can be estimated by trial and error methods from the following relationship (Sallas and Smith, 1980):

$$\hat{\gamma} = \frac{(-1)^j \left[\Gamma(1+3\beta) - 3\Gamma(1+2\beta)(1+\beta) + 2\Gamma^2(+1+\beta)\right]}{\left[\Gamma(1+2\beta) - \Gamma^2(1+\beta)\right]^{1.5}} \quad \text{.... (92)}$$

in which $j = 1$ for $\hat{\gamma} < 1.14$ and $j = 2$ for $\hat{\gamma} \geq 1.14$; and Γ is the gamma function.

Having obtained an estimate of β, the estimates of ∝ and x_o can be obtained from;

$$\hat{\alpha} = (-1)^j \hat{\beta} \quad \text{.... (93)}$$

and

$$\hat{x}_o = A - \hat{\alpha}/\hat{\beta} \quad \text{.... (94)}$$

where

$$\hat{\beta} = \{\hat{\sigma}^2 \left[\Gamma(1+2\hat{\beta})\right]\}^{\frac{1}{2}} \quad \text{.... (95)}$$

and

$$A = \hat{\mu} + (-1)^j \hat{\beta}\Gamma(1+\beta) \quad \text{.... (96)}$$

and $\hat{\mu}$ and σ^2 are the sample estimates of the mean and variance respectively. As before, j = 1 for $\hat{\gamma} < 1.14$ and j = 2 for $\hat{\gamma} \geq 1.14$. The method of maximum likehood is used as follows:

Using the Jenkinson's approach (Jenkinson 1969), the log-likehood function of the general Fisher Tippet Type distributions takes the form:

$$LL(x;x_o,\alpha,\beta) = n\ell n(\alpha)-(1-\beta)\sum_{i=1}^{n} y_i - \sum_{i=1}^{n} e^{-y_i} \qquad \ldots . (97)$$

where

$$y_i = \frac{1}{\beta}\ell n\left[1 - \frac{x_i - x_o}{\alpha}\right] \qquad \ldots . (98)$$

and n is the sample size

Differentiating with respect to x_o, α, and β yields:

$$-\frac{\partial LL}{\partial x_o} = \frac{Q}{\alpha} \qquad \ldots . (99)$$

$$-\frac{\partial LL}{\partial \alpha} = \frac{P+Q}{\alpha\beta} \qquad \ldots . (100)$$

and

$$-\frac{\partial LL}{\partial \beta} = \left[(R-P+Q)/\beta\right]/\beta \qquad \ldots . (101)$$

where

$$P = n - \sum_{i=1}^{n} e^{-y_i} \qquad \ldots . (102)$$

$$Q = \sum_{i=1}^{n} e^{(\beta-1)y_i} - (1-\beta)\sum_{i=1}^{n} e^{\beta y_i} \quad .... (103)$$

and

$$R = n - \sum_{i=1}^{n} y_i \sum_{i=1}^{n} y_i e^{-y_i} \quad .... (104)$$

The maximum likelihood estimates of $\hat{x}_o$, $\hat{\alpha}$ and $\hat{\beta}$ must therefore satisfy the equations:

$$-\frac{\partial LL}{\partial x_o} = 0, \ -\frac{\partial LL}{\partial \alpha} = 0, \ -\frac{\partial LL}{\alpha\beta} = 0 \quad .... (105)$$

These equations must be solved iteratively as there is no explicit solution. It is required to begin the process with initial values x_o, α, β as for example the moment estimates. Let $\Delta x_{o,k}, \Delta\alpha_k$, be the differences between the maximum likelihood estimates and the current estimates $x_{o,k}, \beta_k$ and α_k at the kth iteration. That is

$$x_{o,k+1} = \hat{x}_{o,k} + \Delta x_{o,k} \quad .... (106)$$

$$\hat{\alpha}_{k+1} = \hat{\alpha}_k + \Delta\alpha_k \quad .... (107)$$

and

$$\beta_{k+1} = \beta_k + \Delta\beta_k \quad .... (108)$$

Expanding the equations of (105) in a Taylor series about the maximum likelihood values and omitting the terms containing higher powers and cross-products of $\Delta x_{o,k}, \Delta\alpha_k$ and $\Delta\beta_k$ yields the matrix equation:

$$
\begin{bmatrix} \Delta x_{o,k} \\ \Delta \alpha_k \\ \Delta \beta_k \end{bmatrix} = \begin{bmatrix} -\partial^2 LL/\partial x_{o,k}^2 & -\partial^2 LL/\partial x_{o,k}\partial\alpha_k & -\partial^2 LL/\partial x_{o,k}\partial\beta_k \\ -\partial^2 LL/\partial\alpha_k\partial x_{o,k} & -\partial^2 LL/\partial\alpha_k^2 & -\partial^2 LL/\partial\alpha_k\partial\beta_k \\ -\partial^2 LL/\partial\beta_k\partial x_{o,k} & -\partial^2 LL/\partial\beta_k\partial\alpha_k & -\partial^2 LL/\partial\beta_k^2 \end{bmatrix}^{-1} \begin{bmatrix} \partial LL/\partial x_{o,k} \\ \partial LL/\partial\alpha_k \\ \partial LL/\partial\beta_k \end{bmatrix}
$$

Jenkinson replaces the elements of the matrix to be inverted by their expected values which results in the large sample maximum likelihood variance - covariance matrix of estimators $\hat{x}_o, \hat{\alpha}$ and $\hat{\beta}$, which is given by:

$$
\begin{bmatrix} \mathrm{Var}\ x_o & \mathrm{Cov}(x_o,\alpha) & \mathrm{Cov}(x_o,\beta) \\ \mathrm{Cov}(x_o,\beta) & \mathrm{Var}\ \alpha & \mathrm{Cov}(x_o,\beta) \\ \mathrm{Cov}(x_o,\beta) & \mathrm{cov}(\alpha,\beta) & \mathrm{Var}\ \beta \end{bmatrix} = \begin{bmatrix} \alpha^2 b & \alpha^2 h & \alpha f \\ \alpha^2 h & \alpha^2 a & \alpha g \\ \alpha f & \alpha g & c \end{bmatrix}
$$

here a,b,c,f,g and h are functions of β and are given for some values of β in Table 6 below:

TABLE 6: VARIATION OF THE COEFFICIENTS IN THE INFORMATION MATRIX WITH VARIATIONS IN β (FROM SALAS & SMITH 1980).

β	a	b	c	f	g	h
- 0.4	1.05	1.29	0.84	0.36	0.09	0.80
- 0.3	0.92	1.29	0.73	0.26	0.03	0.69
- 0.2	0.81	1.28	0.64	0.26	0.04	0.57
- 0.1	0.72	1.27	0.55	0.26	0.10	0.46
0.0	0.65	1.25	0.48	0.26	0.15	0.34
0.0	0.61	1.22	0.39	0.24	0.18	0.21
0.2	0.58	1.20	0.33	0.22	0.4	0.09
0.3	0.58	1.17	0.27	0.19	0.23	-0.03
0.4	0.60	1.14	0.21	0.16	0.24	-0.16
0.5	0.63	1.11	0.15	0.13	0.24	-0.30
0.6	0.68	1.08	0.10	0.09	0.22	-0.43
0.8	0.82	1.02	0.03	0.03	0.15	-0.71
1.0	1.00	1.00	0.00	0.00	0.00	-1.00

Inserting the expressions for the derivatives $\partial LL/\partial x_{o,k} \partial LL/\partial \alpha_k$ and $\partial LL/\partial \beta_k$ in P,Q and R as in (99) to (101) with (109) and using (110) for the inverse of the matrix yields:

$$\begin{bmatrix} \Delta x_{o,k} \\ \Delta \alpha_k \\ \Delta \beta_k \end{bmatrix} = \frac{1}{N} \begin{bmatrix} \alpha_k^2 b & \alpha_k^2 & \alpha_k f \\ \alpha_k b & \alpha_k h & \alpha_k f \\ \alpha_k^2 a & \alpha_k g & c \end{bmatrix} \begin{bmatrix} -Q_k/\alpha_k \\ -(P_k+Q_k)/(\alpha_k \beta_k) \\ -\left[(R_k-P_k+Q_k)/\beta_k\right]/\beta_k \end{bmatrix} \quad \text{.... (111)}$$

This matrix equation can be simplified to:

$$\Delta x_{o,k} = \alpha_k(b_k A_k + h_k B_k + f_k C_k)/N$$

$$\Delta \alpha_k = \alpha_k(h_k A_k + a_k B_k + g_k C_k)/N \quad \text{.... (112)}$$

$$\Delta \beta_k = -(f_k A_k + g_k B_k + g_k B_k + c_k C_k)/N$$

where a_k, b_k, c_k, f_k, g_k and h_k are the estimates of a,b,c,f,g and h for a given β_k in iteration k; and A_k A_k, B_k and C_k are given by:

$$A_k = Q_k \quad .... (113)$$

$$B_k = (P_k+Q_k)/\beta_k \quad .... (114)$$

$$C_k = (R_k-B_k)/\beta_k \quad .... (115)$$

and Q_k, P_k and R_k are the values of Q,P and R in iteration k for a given set of α_k and $x_{o,k}$. The new estimate of parameters are:

$$x_{o,k+1} = x_{o,k}+\Delta x_{o,k} \quad .... (116)$$

$$\alpha_{k+1} = \alpha_k+\Delta\alpha_k \quad .... (117)$$

$$\beta_{k+1} = \beta_k+\Delta\beta_k \quad .... (118)$$

For every iteration k, one requires to set up a routine for estimating a,b,c,f,g and h in Table 6 which can be very time consuming and complicated. However, the following regression equations, valid for $-0.4 \leq \beta \leq 1.0$ and with a Spearman's rank correlation coefficient $R \geq 0.999$ and maximum absolute error MAE ≤ 0.0092 (Salas and Smith, 1980), are usually used.

$$a=0.6528060-0.5598783\beta+1.0876209\beta^2-0.054024\beta^3-0.1270214\beta^4 \quad .... (119)$$

$$b=1.2488727-.0.205244\beta-0.2225715\beta^2+0.0962481\beta^3+0.0813214\beta^4 \quad .... (120)$$

$$c=0.4725506-0.7603083\beta+0.2836171\beta^2-.1886466\beta^3+0.1931182\beta^4 \quad .... (121)$$

$$f=0.2597848-0.1727130\beta-0.1370156\beta^2-0.2437380\beta^3+0.2937678\beta^4 \quad .... (122)$$

$$g=0.1432656+0.4419219\beta-0.4261602\beta^2-0.156171\beta^3 \quad .... (123)$$

$$h=0.3386324-1.2041691\beta-0.1333794\beta^2 \quad .... (124)$$

The numerical procedure may be finished when all of the increments of $\Delta x_{o,k}$, $\Delta\alpha_k$ and α_k become less than a selected percentage of the correct values of x_o, α_k and β_k usually taken as 1×10^{-6}.

For a given set of parameters of the Fisher Tippet type distributions, the probability density function and cumulative distribution function can be computed directly from the corresponding equations since these have explicit exponetial forms.

2.5 THE LOG-FISHER TIPPET DISTRIBUTIONS

Actually, the log-Fisher Tippet distributions are the Fisher Tippet distributions in the log-domain. That is, if x is log-Fisher Tippet distributed, then $y = \ln(x)$ is Fisher Tippet distributed. That is:

$$f(y) = \frac{1}{\alpha}\left[1-\beta(y-y_o)/\alpha\right]^{\frac{1}{\beta}} F(y) \quad .... (125)$$

where

$$F(y) = \exp\{-\left[1-\beta(y-y_o)/\alpha\right]^{\frac{1}{\beta}}\} \quad .... (126)$$

and α, β and y_o are the scale, shape and location parameters, respectively, of the distribution in the log-domain.

Hence,

$$f(x) = \frac{1}{x}(f(y)) \quad .... (127)$$

and

$$F(x) = F(y) \quad .... (128)$$

It is evident from (127) that the parameters of the log-Fisher Tippet distribution are exactly those of the Fisher Tippet distribution in the log-domain. The methods of estimating these parameters have already been discussed in the previous sections.

2.6 THE WALTER BOUGHTON DISTRIBUTION

The Walter Boughton distribution is essentially a three parameter probability distribution which is based on the non-linear relationship between the frequency factor K and the function $\ell n\{\ell n\left[T/(T-1)\right]\}$ of the recurrence interval T. It is important to note that the Gumbel distribution assumes a linear relationship between K and $\ell n\{\ell n\left[T/(T-1)\right]\}$ which is given by the following equation.

$$K = -0.45+0.779\,\ell n\left[\ell n\left(\frac{T}{T-1}\right)\right] \quad .....(129)$$

Boughton (1980), Mutua (1984) and others have shown that a non-linear relationship exists without exception between K and $\ell n\{\ell n[T/(T-1)]\}$ using flood data in Australia and Kenya respectively. In fact, there is sufficient evidence to show that this relationship holds true for all flood data (Boughton 1980).

Generally, this non-linear relationship is of the form

$$(K-a)\left[\ell n \ell n\left(\frac{T}{T-1}\right)-b\right] = c \qquad \text{.... (130)}$$

where a and b are the asymptotes of the frequency factor K and $\ell n \ell n[T/(T-1)]$ respectively, and c is a positive constant. Equation (130) can also be rewritten as

$$K = a+c/\{\ell n \ell n\left(\frac{T}{T-1}\right)-b\} \qquad \text{.... (131)}$$

In flood frequency analyses, it is customary to relate the estimate of a flood magnitude x and its corresponding return period, T, to the population mean μ and variance σ^2 (or the corresponding sample estimates) and the frequency factor K by an equation of the form

$$x = \mu + K\sigma \qquad \text{.... (132)}$$

so that

$$K = (x-\mu)/\sigma \qquad \text{.... (133)}$$

Using (133), (131) can also be rewritten as:

$$x = \mu + a\sigma + c\sigma/\{\ell n \ell n\left(\frac{T}{T-1}\right)-b\} \qquad \text{.... (134)}$$

The return period T which corresponds to a flood of magnitude x is related to the cumulative distribution F(x), by the equation

$$F(x) = (T-1)/T \qquad \ldots. (135)$$

Hence, (134) can further be rewritten as

$$x = \mu+a\sigma+c\sigma/\{\ell n\left[-\ell n(Fx)\right]-b\} \qquad \ldots. (136)$$

from which we get

$$F(x) = \exp\{-\exp\left[\alpha+\beta/(x-x_o)\right]\} \qquad \ldots. (137)$$

where $\alpha = b$, the scale parameter, $\beta = c\sigma$, the shape parameter, and $x_o = \mu+c\sigma$, the location parameter.

The corresponding probability density function f(x) is therefore given by

$$f(x) = \beta.ex\left[\alpha+\beta/(x-x_o)\right]/(x-x_o)^2.F(x) \qquad \ldots. (138)$$

There are two methods, which can be developed for the estimation of the parameters of the Walter Boughton distribution, namely, the method of least squares and the method of maximum likelihood. Equation (138) can also be rewritten as

$$\frac{1}{x_i-x_o} = \frac{\alpha}{\beta} + \frac{1}{\beta}\ell n\{-\ell n\left[F(x_i)\right]\},\ i=1,2,\ldots,n \qquad \ldots. (139)$$

when n is the number of observations and the subscript i denotes the ith observation.

The method of least squares can be therefore developed from (139) as follows:

(i) By the use of an appropriate plotting position formula, $F(x_i)$ can be estimated for all x_i, $i=1,2,\ldots,n$, where n is the number of observations.

(ii) A value of x_o is adopted and a regression analysis is done between $1/(x_i-x_o)$ and $\ell n\{-\ell nF(x_i)\}$. By repeating this process several times, each time with a different estimate x_o, the optimum values of x_o can be obtained as the one which gives the minimum sum of squared error. If I and S are the intercept and slope of the regression analysis at minimum sum of squared error, then we can see from (139) that

$$\beta = 1/S \qquad \ldots. (140).$$

and

$$\alpha = - I/S \qquad \ldots. (141)$$

It should be noted that the method of least sqaures gives the estimates of x_o, α and β which are fairly close to the maximum likelihood estimates, especially when a good plotting position formula is used.

From (137) and (138), it can be shown that the log-likelihood function of the Walter Boughton distribution is given by:

$$LL(x:x_o\alpha,\beta) = - \sum_{i=1}^{n} \exp\left[\alpha+\beta/(x_i-x_o)\right] + \sum_{i=1}^{n}\left[\alpha+\beta/(x_i-x_o)\right] + n\ell n\beta - 2\sum_{i=1}^{n} \ell n(x_i-x_o)$$

$$\ldots. (142)$$

For convenience, we will denote $LL(x;x_o,\alpha,\beta)$ by LL. Then the maximum likelihood estimation can be obtained from the system of simultaneous equations given by

$$\frac{\partial LL}{\partial x_o} = - \sum\{e^{\left[\alpha+\beta/(x_i-x_o)\right]}\, \beta/(x_i-x_o)^2\} + \sum_{i=1}^{n}\left[\beta/(x_i-x_o)^2\right] + \sum_{i=1}^{n} \frac{2}{x_i-x_o} = 0$$

$$\frac{\partial LL}{\partial \alpha} = - \sum_{i=1}^{n} e^{\left[\alpha+\beta/(x_j-x_o)\right]} + n = 0. \qquad \ldots. (143)$$

$$\frac{\partial LL}{\partial \beta} - \sum_{i=1}^{n} \{ e^{\left[\alpha+\beta/(x_i-x_o)\right]}/(x_i-x_o)\} + \sum_{i=1}^{n} \left[1/(x_i-x_o) \right] + n/\beta = 0$$

Following the methodology as given by Carnaham *et al* (1969), the solution of this system of non-linear simultaneous equations can be obtained by an iterative procedure which can be represented in matrix form by

$$\begin{bmatrix} x_{o,k+1} \\ \alpha_{k+1} \\ \beta_{k+1} \end{bmatrix} = \begin{bmatrix} x_{o,k} \\ \alpha_{k} \\ \beta_{k} \end{bmatrix} - \begin{bmatrix} \Delta x_{o,k} \\ \Delta\alpha_{k} \\ \Delta\beta_{k} \end{bmatrix} \quad \text{.... (144)}$$

where $x_{o,k}$, α_k and β_k denote the estimates of x_o, α and β during the k-th iteration and $x_{o,k+1}$, α_{k+1} and β_{k+1} are the updated estimates after the k-th iteration. The increments $\Delta x_{o,k}$, $\Delta\alpha_k$ and $\Delta\beta_k$ during the k-th iteration can be obtained from the solution of the matrix:

$$\begin{bmatrix} \frac{\partial^2 LL}{\partial x_o^2} & \frac{\partial^2 LL}{\partial x_o \partial\alpha} & \frac{\partial^2 LL}{\partial x_o \partial\beta} \\ \frac{\partial^2 LL}{\partial\alpha \partial x_o} & \frac{\partial^2 LL}{\partial\alpha^2} & \frac{\partial^2 LL}{\partial\alpha\partial\beta} \\ \frac{\partial^2 LL}{\partial\beta \partial x_o} & \frac{\partial^2 LL}{\partial\beta\partial\alpha} & \frac{\partial^2 LL}{\partial\beta^2} \end{bmatrix} \begin{bmatrix} \Delta x_{o,k} \\ \Delta\alpha_{k} \\ \Delta\beta_{k} \end{bmatrix} = - \begin{bmatrix} \frac{\partial LL}{\partial x_o} \\ \frac{\partial LL}{\partial\alpha} \\ \frac{\partial LL}{\partial\beta} \end{bmatrix} \quad \text{.... (145)}$$

where,

$$\frac{\partial^2 LL}{\partial x_o^2} = -\sum_{i=1}^{n} \{\exp\left[\alpha + \frac{\beta}{x_i - x_o}\right] \frac{\beta}{(x_i - x_o)^3} \left[\frac{\beta}{(x_i - x_o)} + 2\right] - \frac{2}{(x_i - x_o)^2} \left[\frac{\beta}{(x_i - x_o)} + 1\right]\}$$

.... (146)

$$\frac{\partial^2 LL}{\partial x_o \partial \alpha} = -\sum_{i=1}^{n} \exp\left[\alpha + \frac{\beta}{(x_i - x_o)}\right] \frac{\beta}{(x_i - x_o)^2} \qquad .... (147)$$

$$\frac{\partial^2 LL}{\partial x_o \partial \beta} = -\sum_{i=1}^{n} \{ \exp\left[\alpha + \frac{\beta}{(x_i - x_o)}\right] \left(\frac{1}{x_i - x_o}\right)^2 \left[\frac{\beta}{(x_i - x_o)} + 1\right] - \left(\frac{1}{x_i - x_o}\right) \}$$

.... (148)

$$\frac{\partial^2 LL}{\partial \alpha^2} = -\sum_{i=1}^{n} \exp\left[\alpha + \frac{\beta}{(x_i - x_o)}\right] \qquad .... (149)$$

$$\frac{\partial^2 LL}{\partial \alpha \partial \beta} = -\sum_{i=1}^{n} \exp\left[\alpha + \frac{\beta}{(x_i - x_o)}\right] / (x_i - x_o) \qquad .... (150)$$

$$\frac{\partial^2 LL}{\partial \beta^2} = -\sum_{i=1}^{n} \exp\left[\alpha + \frac{\beta}{(x_i - x_o)}\right] \left(\frac{1}{x_i - x_o}\right)^2 - \frac{n}{\beta 2} \qquad .... (151)$$

It is possible to set up the stopping criterion such that the iterations stop when $\Delta x_{o,k}$, $\Delta\alpha_k$ and $\Delta\beta_k$ are all less than the required error limit. Further, the method of maximum likelihood requires initial estimates of x_o, α and β. The estimates as obtained by the method of least squares can be used for this purpose.

Just like as it is with the Fisher Tippet distribution, the probability density and cumulative distribution functions can be computed directly from the corresponding equations for a given set of parameters.

2.7 THE LOG -WALTER BOUGHTON DISTRIBUTION

If a variate x is log Walter Boughton distributed, then $y = \ell n(x)$ is Walter Boughton distributed. That is

$$f(y) = \beta.\exp\left[\alpha + \beta(y-y_o)\right]/(y-y_o)^2.F(y) \quad .... (152)$$

where

$$F(y)= \exp\{-\exp\left[\alpha + \beta/(y-y_o)\right]\}$$

and α, β and y_o are the scale, shape and location parameters respectively, of the distribution in the log-domain. Hence,

$$f(x)=\frac{1}{x}f(y) \quad .... (153)$$

and

$$F(x)=F(y) \quad .... (154)$$

It is clear from (153) that the parameters of the log-Walter Boughton distribution are essentially those of the Walter Boughton distribution in the log-domain. The methods of estimating such parameters of the Walter Boughton distribution have already been discussed in the previous sections. In the same way, the probability density functions and the cumulative distribution functions can be computed directly from (153) and (154) respectively, for a given set of parameters.

2.8 THE WAKEBY DISTRIBUTION

The distribution which has aroused the most interest for analytical reasons is the five parameter Wakeby distribution. This distribution is most easily defined as an inverse function of the form:

$$x=-a[1-F(x)]^{b}+c[1-F(x)]^{-d}+e \qquad \text{.... (155)}$$

where F(x) is the cumulative probability corresponding to the observation x; a, b, c and d are positive distribution parameters and e is also a distribution parameter which can be negative or positive.

The Wakeby distribution is similar to a five parameter member of the Tukey family of Labdas (Joiner and Rosenblatt, 1971). Given values of a and b that are typical of the flood records, the $-a[1-F(x)]^{b}$ term generally has no effect on x if $F(x)>0.25$. Thus, the Wakeby distribution can be thought of in two parts: the right hand tail $c[1-F(x)]^{-d}+e$ and the left hand tail $-a[1-F(x)]^{b}$ which is in effect an adjustment to the graph of the right hand tail.

The Wakeby distribution was introduced by Houghton (1978a) as a grand parent of the frequency distributions because of its ability to generate flows which mimic most conventional hydrologic distributions if the parameters are chosen correctly. Further, Houghton (1978b) noted the similarity of the Wakeby distribution to the older Fuller (1914) flood model which is expressed as

$$x=a+b[T(x)] \qquad \text{.... (156)}$$

where

$$T(x)=1/[1-F(x)] \qquad .... (157)$$

and a and b are positive constants and F(x) is as defined above.

The Wakeby distribution was introduced mainly because of its ability to absorb more degrees of freedom than the traditional distributions. There is some reason to believe that none of the traditional distributions have the properties on their left hand tails that reflect nature accurately (Greis 1983). If in reality, the lowest observations follow the left-hand tail of a low-skew log-normal distribution and the highest observations follow the right hand tail of a high-skew lognormal distribution, then, no conventional three-parameter distribution can model it accurately. They lack enough kurtosis for any given skew. Fitting a three-parameter curve to a four-parameter nature can distort the whole fit, including the higher quantiles. This is mainly due to the so called 'separation effect' presented by Matalas et al (1975).

When the traditional distributions are used to generate synthetic flood data in Monte Carlo experiments, the standard deviation of skew of such data is usually found to be lower than that which exists in nature. This is the phenomenon refered to as the separation effect.

The accuracy of the procedures involved in the estimation of the parameters of a distribution is very sensitive to the instability of higher moments and their functions such as the coefficient of skew. The higher moments often add more noise than signal to the estimation procedures for conventional distributions. Although the Wakeby distribution has five parameters, some of the parameter estimation methods require only the first sample moment.

There are about four methods which can be used to estimate the parameters of the Wakeby distribution. These methods are: the method of least squares, the method of incomplete means, the method of probability weighted moments and the differential correction technique method. The usage of these methods is discussed in the following subsections.

2.8.1 THE LEAST SQUARES METHOD FOR THE ESTIMATION THE WAKEBY PARAMETERS.

This parameter estimation procedure which was developed by Houghton (1978b) works in two phases, taking advantage of the separation properties of the left and right tails of the Wakeby distribution. Phase one operates on the right hand tail while phase two operates on the left hand tail. The procedure is as follows:

(i) Choose some $F(x_c)$ which is a cutoff point, where $F(x_c)$ denotes the cumulative probability of the variate of magnitude x_c. This cumulative probability is usually estimated by an appropriate plotting position formula. The curve corresponding to the cumulative probability $F(x)$ such that $F(x)>F(x_c)$ is analysed in phase one and that corresponding to $F(x)<F(x_c)$ is analysed in phase two. For phase one,

$$x_k=-a\left[1-F(x_k)\right]^{b}+c\left[1-F(x_k)\right]^{-d}+e \qquad \text{.....(158)}$$

or alternatively,

$$x_k-e+a\left[1-F(x_k)\right]^{b}=c\left[1-F(x_k)\right]^{-d} \qquad \text{.... (159)}$$

therefore,

$$\log\{x_k-e+a\left[1-F(x_k)\right]^{b}=\log(c)-d\log\left[1-F(x_k)\right] \qquad \text{...(160)}$$

for all $x_k>x_c$

(ii) set a = 0, b = 1 and assume an initial value for e. then one can use linear regression analysis on the basis of (160) between $\log\{x_k-e+a[1-F(x_k)]^{b}\}$ and $\log[1-F(x_k)]$ for all the observations x_k such that $F(x_k)>F(x_c)$. The regression analysis then yields estimates of c and d from the intercept and slope respectively.

(iii) a search is then made over various values of e to find the one which minimises the sum of squares of vertical distance from each observation points to the regression line.

Phase one gives improved estimates of c,d and e.

(iv) in phase two, one assumes the values of c,d and e estimated in phase one and evaluates a and b by regression analysis between $\log\{-x_j+e+c[1-F(x_j)]^{-d}\}$

and log $[1-F(x_j)]$ according to the equation.

$$\log\{-x_j+e+c\,[1-F(x_j)]^{-d}\}=\log(a)+b\,\log[1-F(x_j)]. \quad \text{.... (161)}$$

for all x_j such that $F(x_j)<F(x_c)$. Phase two therefore gives estimates of a and b from the intercept and slope respectively.

(v) given new values of a and b, phase one is repeated, then phase two, and so on. In practice, repetitions are unnecessary (Houghton, 1978b). In those cases where repetitions are necessary, one repetition provides most of the change and further repetitions tend to oscillate about the required estimates.

The cut-off cumulative probability $F(x_c)$ is usually not known apriori and can only be determined by a trial and error method. This trial and error method can be started by assuming that the left hand tail of the Wakeby distribution contains the lowest two observations, with the rest of the observations in the right hand tail of the distribution. Thus, if the observations are sorted in an ascending order, the above assumption would correspond to $F(x_c)=F(x_2)$. By continuing this process from x_2 to x_3 and so on to the third quartile x_m, different values of (x_c), can be obtained together with their corresponding parameter estimates which are obtained as explained above. Usually, x_c falls within the lower half of the observations but extending the search to the lowest 75% of the observations ensures that the wrong final estimates of $F(x_c)$ is not obtained.

The criterion by which the optimum value of $F(x_c)$ is chosen (together with its corresponding parameter estimates) is a weighted sum of the squares of the correlation coefficients calculated in phase one and phase two according to the relationship given below;

$$\rho_w = \frac{n_\ell}{n}\rho_\ell^2 + \frac{n-n_\ell}{n}\rho_u^2 \qquad \text{.... (162)}$$

where ρ_w is the weighted sum of squares of the correlation coefficients, n_ℓ is the number of observations in the lower tail of the Wakeby distribution, ρ_ℓ is the correlation coefficient in phase two, ρ_u is the correlation coefficient in phase one and n is the total number of observations.

The optimum cut-off cumulative probability $F(x_c)$ is chosen as the one that has the maximum ρ_w. The optimum estimates of a,b,c,d and e are then chosen to be those which correspond to the optimum value of $F(x_c)$. For a high value of $F(x_c)$, phase one would often result in some calculations involving the logarithms of negative numbers. In such case, the particular cut-of point and all others above it are excluded from further analysis.

2.8.2 ESTIMATION OF THE WAKEBY DISTRIBUTIONS PARAMETERS BY INCOMPLETE MEANS.

The incomplete means estimation uses a combination of means calculated over only part of the total range. This new procedure yields fairly stable estimates with little bias since it uses no moments higher than the first (Boughton 1978b). Consider a sample of n ranked observations $x_1, x_2, \ldots x_n$. Calculate the mean $\bar{x}$; it will fall between two adjacent observations in the sample. It effectively divides the sample into two disjoint sets. Calculate the mean of the upper set and call it $\bar{x}_1$. Similarly calculate the mean of all the observations above $\bar{x}_1$, and lable the incomplete mean $\bar{x}_2$, and so on. The incomplete means and the corresponding n_i, the corresponding rank order from which the computation of the mean start, can be related to the parameters in the Wakeby distribution in a straightforward way. This is due to the inverse definition of the Wakeby; the relation of the incomplete means to the parameters of the lognormal distribution, for example, would not be as simple.

To derive the relationship between the parameters and the incomplete means, recall that

$$\bar{x} = \int_{-\infty}^{\infty} x f(x) dx \qquad \ldots. (163)$$

where f(x) is the probability density function of the variate x. Let $y=F(x)$, $x=F^{-1}(y)$ and $dy=f(x)dx$

where F(.) denotes the cumulative distribution function. Then,

$$\bar{x} = \int_o^1 F^{-1}(y)\,dy \qquad \ldots. \; (164)$$

Define $\bar{x}_{(a,b)}$ to be the mean of the interval (x_a, x_b) in the x-space or of the interval (a,b) in F-space. Then

$$\bar{x}_{(a,b)} = \frac{1}{b-a}\int_a^b F^{-1}(y)\,dy \qquad \ldots. \; (165)$$

In the incomplete means method, end points are determined by functions of n_i. For example, for the Wakeby distribution,

$$\bar{x}_i = \frac{1}{\frac{n_i}{1-\frac{n_i}{n}}}\int_{\frac{n_i}{n}}^{1} -a(1-F)^b + c(1-F)^{-d} + e \; dF \qquad \ldots. \; (166)$$

$$i=O,1,2,3,4$$

which reduces to

$$\bar{x}_i = -\frac{a(1-\frac{n_i}{n})^b}{(1+b)} + \frac{(1-\frac{n_i}{n})^{-d}}{(1-d)} + e \qquad \ldots. \; (167)$$

$$i=O,1,2,3,4$$

Therefore, by using $\bar{x}, \bar{x}_1, \bar{x}_2, \bar{x}_3, \bar{x}_4$; and the corresponding expressions from (167) it is possible to estimate the parameters a,b,c,d and e of the Wakeby distribution .

2.8.3. ESTIMATION OF THE WAKEY DISTRIBUTION PARAMETERS BY PROBABILITY WEIGHTED MOMENTS.

The applicability of the probability weighted moments in the estimation of the parameters of the Wakeby distribution and other distributions was given by Greenwood *et al.*, (1979)

A distribution function $F = F(x)=P(X<x)$ denotes the probability that the independent variate X is less than or equal to a value, can be characterised by probability weighted moments which are defined as

$$M_{\ell,j,k}=E[X^{\ell}F^{j}(1-F)^{k}]=\int_{o}^{1} x^{\ell}F(1-F)^{k}dF \qquad \ldots\ldots (168$$

where ℓ,j and k are real numbers. If j=k=O and ℓ is a nonnegative integer, then $M_{\ell,o,o}$ represents the conventional moments about the origin of order ℓ. If $M_{\ell,o,o}$ exists and X is a continuous function of F, then $M_{\ell,j,k}$ exists for all nonnegative real numbers j and k.

If j and k are nonnegative integers,

$$M_{\ell,j,k} = \sum_{j=o}^{k} \binom{k}{j}(-1)^{j}M_{\ell,j,o} \qquad \ldots\ldots (169)$$

and

$$M_{\ell,j,o} = \sum_{k=o}^{j} \binom{k}{j}(-1)^{k}M_{\ell,o,k} \qquad \ldots\ldots (170)$$

where if $M_{\ell,o,k}$ exists and X is a continuous function of F,$M_{\ell,j,o}$ exists. In general, if $M_{\ell,j,k}$ exists it may be difficult to derive its analytical form, particularly if the inverse x=x(F) of the distribution F=F(x) cannot be analytically defined.

In the special case where ℓ, j and k are nonnegative integers, $M_{\ell,j,k}$ is proportional to $E[X^{\ell}_{j+1,\ k+j+1}]$, the ℓ-th moment about the origin of the (j+1)-th order statistic for a sample of size k+j+1. More specifically,

$$E\left[X^{\ell}_{j+1,k+j+1}\right] = M_{\ell,j,k}/B\left[j+1,\ k+1\right] \qquad \text{.... (171)}$$

where B[•,•] denotes the beta function. If j=0, then $(k+1)M_{\ell,o,k}$ represents the ℓ-th moment about the origin, of the first order statistic for a sample of size (k+1); and if k=0, $(j+1)M_{\ell,j,o}$ represents the ℓ-th moment about the origin of the (j+1)-th order statistic for a sample of size (j+1).

Among the distributions for which only the inverse form x=x(F) is explicitly defined are the generalized lambda (Tukey, 1960) and the Wakeby distribution. There are many distributions which may be explicitly defined as both F=F(x) and x=x(F), among them being the kappa, recently introduced by Milke (1973) in analysing precipitation data, and the more familiar Weibull, Gumbel and logistic distributions.

Let $M_{(k)} = M_{1,o,k}$. Then, for the Wakeby distribution,

$$M_{(k)} = -\frac{a}{1+k+b} + \frac{c}{1+k-d} + \frac{e}{1+k} \qquad \text{.... (172)}$$

The algorithm for estimating the Wakeby parameters from (172) requires estimates of $M_{(k)}$, for several nonnegative values of k. To estimate $M_{(k)}=E[X(1-F)^{k}]$ from a sample of size n, where k is nonnegative, let x_i, i=1,...n, denote the sample values in ranked order, smallest to largest. An unbiased estimator of $M_{(k)}$, where k is constrained to be a nonnegative integer is given by (Landwehr *et al.*, 1979):

$$\hat{M}_{(k)} = \frac{1}{n}\sum_{i=1}^{n-k} x_i \binom{n-i}{k} / \binom{n-1}{k} \qquad .... \ (173)$$

Another estimator constructed in accordance with the concept of plotting positions is given by

$$\hat{M}_{(k)} = \frac{1}{n}\sum_{i=1}^{n} x_i [1-p(x_i)]^k \qquad .... \ (174)$$

where $p(x_i)$ is the plotting position for the observation x_i. Landwehr *et al.* (1979b) and Landwehr *et al.* (1979c) used

$$p(x_i) = (i-0.35)/n \qquad .... \ (175)$$

Note that in (174) it is not necessary to specify that k is an integer; k is assumed only to be nonnegative. After estimating M_k for five values of k, the simultaneous equations, of (172) can then be solved to yield the parameter estimates $\hat{a}, \hat{b}, \hat{c}, \hat{d}$, and $\hat{e}$.

2.8.4. ESTIMATION OF THE WAKEBY DISTRIBUTION PARAMETERS BY THE DIFFERENTIAL CORRECTED TECHNIQUE.

In many applied problems, the form of a function $f(x, c_o, c_1, ..., c_m)$ representing a phenomenon is known, whereas the parameters $c_o, c_1, \ldots. c_m$ are to be determined. For example, according to theoretical considerations, a particular phenomenon might be represented by an exponential, logarithmic, trigonometric or other functions. If a set of observations provides emperical values y_i of the function

$$f(x_i, c_o, c_1, \ldots, c_m),\ i=1,2,\ldots.n,\ (n>m) \ldots (176)$$

then the parameters c_k, $k=0,1,\ldots,m$, can be determined from a set of estimates c_k^o, $k=0, 1,\ldots.m$, by differential correction techniques based on least squares. Such techniques are used extensively in mathematics.

Assume that for a particular phenomenon, there exists an explicit function

$$y = f(x, c_o, c_1, \ldots, c_m) \qquad \ldots. (177)$$

relating the variables x and y where $c_o, \ldots, c_m$ are unknown independent parameters that enter non-linearly into the functional expression. Assume that observation of this phenomenon has provided a set of data pairs (x_i, y_i), $i=1,2,\ldots,n$, at distinct and discrete values of x_i of the independent variable x. The recorded values y_i may contain errors of observation, measurement, recording, transmission, conversion and so forth.

If the values of the parameters c_k were known, it would be possible to evaluate $f(x, c_o, c_1, \ldots, c_m)$

for each x_i to obtain a set of 'true" residuals.

$$r_i=f(x_i,c_o,c_1,....,c_m)-y_i, i=1,2,....n \quad ....(178)$$

A "true" residual then would represent the difference between the actual function value at x_i and the emperical or recorded value y_i, for i=1,2,...,n. These "true" residuals cannot be calculated because the actual values of the parameters c_k, k=0,1,....,m are unknown.

However, if estimates c_k^o, k=0,1,....,m of the parameters can be obtained by some means, then "computed" residuals

$$R_i=f(x_i,c_o^o,c_1^o,....,c_m^o)-y_i \quad .... (179)$$

can be calculated. The problem then is to obtain improved values of the parameters c_k using the data pairs (x_i,y_i), the estimates c_k^o and the computed residuals R_i. This can be accomplished by a differential correction technique based on least squares, provided the estimates c_k^o are sufficiently close to the actual values of the parameters c_k to lead to convergence of the method.

This differential correction technique can be derived by first expanding the function about $(c_k^o,c_1^o,....,c_m^o)$ using a linear Taylor-series expansion of the form (McCalla, 1967):

$$f(x,c_o,c_1,....,c_m)=f(x,c_o^o,c_1^o,....,c_m^o)+$$

$$\frac{\partial f}{\partial c_o}(c_o-c_o^o)+\frac{\partial f}{\partial c_1}(c_1-c_1^o)+.....\frac{\partial f}{\partial c_m}(c_1-c_1^o) \quad ....(180)$$

so that a relation between r_i and R_i can be obtained. This can be obtained by evaluating (180) at each x_i and subtracting y_i from both sides of the equation. Using the definitions

$$\delta c_k = c_k - c_k^o$$

and

$$\frac{\partial f_i}{\partial c_k} = \left.\frac{\partial f_k}{\partial c_k}\right|_{x=x_i, c_k = c_k^o}$$

for k = 0,1,....,m and i = 1,2,....,n, we can write the result in the form

$$f(x_i,c_o,c_1,....,c_m)-y_i=f(x_i,c_o^o,c_1^o,...,c_m^o)-y_i+$$

$$\frac{\partial f_i}{\partial c_o}\partial c_o+.... + \frac{\partial f_i}{\partial c_m}\partial c_m. \quad .... (181)$$

The desired relation between r_i and R_i can be found by substituting (178) and (179) into (181); the result is the relation

$$r_i \simeq R_i + \frac{\partial f_i}{\partial c_o} c_o+....\frac{\partial f_i}{\partial}\partial c_m \quad .... (182)$$

Now let us see how this relation can be used to compute, from c_k^o, a set of parameters c_k

that minimise the sum of squares of the "true" residuals r_i; that is, such that the quantity Q, defined by the relation

$$Q = \sum_{i=1}^{n} r_i^2 = \sum_{i=1}^{n} \left[R_i + \frac{\partial f_i}{\partial c_o}\delta c_o + \ldots \frac{\partial f_i}{c_m}\delta c_m \right] \quad \ldots . \ (183)$$

is minimised.

Equation (183) indicates that Q is a "function" of the terms δc_k. That is, if any δc_k is varied, then the change in Q will result. We therefore consider Q as a function of the terms δc_k and write $Q(c_o, c_1, \ldots, \delta c_m)$. This function Q has a minimum value when all its partial derivatives with respect to the δc_k are simultaneously zero; that is, when

$$\frac{\partial Q}{\partial(\delta c_k)} = 2 \sum_{i=1}^{n} R_{i} + \frac{\partial f_i}{\partial c_o}\delta c_o + \ldots + \frac{\partial f_i}{\partial c_m}\delta c_m \frac{\partial \Delta}{\partial(\delta c_k)} = 0 \quad \ldots . \ (184)$$

for $k=0,1,\ldots,m$,

where

$$\Delta = \frac{\partial f_i}{\partial c_o}\delta c_o + \ldots + \frac{\partial f_i}{c_m} c_m$$

denotes the total differential o f the function f. Substituting the relation

$$\frac{\partial \Delta}{\partial(\delta c_k)} = \frac{\partial f_i}{\partial c_k}$$

into (184), and multiplying each term in the brackets by this quantity, we can write (184) in the form

$$\frac{\partial Q}{\partial(\delta c_k)} = \sum_{i=1}^{n}\left[\frac{\partial f_i}{\partial c_k}R_i + \frac{\partial f_i}{\partial c_o}\delta c_o + \dots + \frac{\partial f_i}{\partial c_k}\frac{\partial f_i}{\partial c_m}\delta c_m\right] = 0 \qquad \dots (185)$$

for k = 0, 1,....,m.

Evaluating this equation for each value of k, and writing the result in matrix form, we obtain the normal set of equations:

$$\begin{bmatrix} \sum_{i=1}^{n}\left(\frac{\partial f_i}{\partial c_o}\right)^2 & \sum_{i=1}\frac{\partial f_i}{\partial c_o}\frac{\partial f_i}{\partial c_1} & \dots & \sum_{i=1}^{n}\frac{\partial f_i}{\partial c_m} \\ \sum_{i=1}^{n}\frac{\partial f_i}{\partial c_1}\frac{\partial f_i}{c_o} & \sum_{i=1}^{n}\left(\frac{\partial f_i}{\partial c_1}\right)^2 & \dots & \sum_{i=1}^{n}\frac{\partial f_i}{\partial c_1}\frac{\partial f_i}{\partial c_m} \\ \vdots & \vdots & & \vdots \\ \sum_{i=1}^{n}\frac{\partial f_i}{\partial c_m}\frac{\partial f_i}{c_o} & \sum_{i=1}^{n}\frac{\partial f_i}{\partial c_m}\frac{\partial f_i}{\partial c_1} & \dots & \sum_{i=1}^{n}\left(\frac{\partial f_i}{\partial c_m}\right)^2 \end{bmatrix} \begin{bmatrix} \delta c_o \\ \delta c_1 \\ \vdots \\ \delta c_m \end{bmatrix} = \begin{bmatrix} -\sum_{i=1}^{n}\frac{\partial f_i}{\partial c_o}R_i \\ -\sum_{i=1}^{n}\frac{\partial f_i}{\partial c_1} \\ \vdots \\ -\sum_{i=1}^{n}\frac{\partial f_i}{\partial c_m}R_i \end{bmatrix}$$

....(186)

The solution $(\delta c_o, \delta c_1, \ldots, \delta c_m)$ of this set of equations is a first order approximation of the change in $(c_c^o, c_1^o, \ldots . c_m^o)$ required to obtain the parameters $(c_o, c_1, \ldots, c_m)$. If any $|\delta c_k| > \varepsilon$, where ε is a required error limit, we replace c_k^o by $c_k^o + \delta c_k$ $(k=0,1,\ldots,m)$ and repeat the entire differential correction technique using the known new estimates.

In order to relate this type of analysis to the Wakeby distribution, let

$$S_{aa} = \sum_{i=1}^{n} (1-F_i)$$

$$S_{ab} = a \sum_{i=1}^{n} \left[(1-F_i)^b\right]^2 \log(1-F_i)$$

$$S_{ac} = - \sum_{i=1}^{n} (1-F_i)^{b-d}$$

$$S_{ad} = c \sum_{i=1}^{n} (1-F_i)^{b-d} \log_e(1-F_i)$$

$$S_{ae} = - \sum_{i=1}^{n} (1-F_i)^b$$

$$S_{bb} = a^2 \sum_{i=1}^{n} (1-F_i)^{2b} \left[\log_e(1-F_i)\right]^2$$

$$S_{bc} = - \sum_{i=1}^{n} (1-F_i)^{b-d} \log(1-F_i)$$

$$S_{bd} = ac\sum_{i=1}^{n}(1-F_i)^{b-d}\left[\log_e(1-F_i)\right]^2$$

$$S_{be} = -\sum_{i=1}^{n}(1-F_i)^{b}$$

$$S_{cc} = \sum_{i=1}^{n}(1-F_i)^{-2d}$$

$$S_{cd} = -c\sum_{i=1}^{n}(1-F_i)^{-2d}\log_e(1-F_i)$$

$$S_{ce} = \sum_{i=1}^{n}(1-F_i)^{-d}$$

$$S_{dd} = c\sum_{i=1}^{n}(1-F_i)^{-2d}\left[\log_e(1-F_i)\right]^2$$

and

$$S_{de} = -c\sum_{i=1}^{n}(1-F_i)^{-d}\log(1-F_i)$$

where n is the number of independent peak flow observations and F_i, the cumulative probability is estimated by an appropriate plotting position formula. Then (186) for the Wakeby parameters becomes

$$\begin{bmatrix} S_{aa} & S_{ab} & S_{ac} & S_{ad} & S_{ae} \\ S_{ab} & S_{bb} & S_{bc} & S_{bd} & S_{be} \\ S_{ac} & S_{bc} & S_{cc} & S_{cd} & C_{ce} \\ S_{ad} & S_{bd} & S_{cd} & S_{dd} & S_{de} \\ S_{ae} & S_{be} & S_{ce} & S_{de} & n \end{bmatrix} \begin{bmatrix} \delta a \\ \delta b \\ \delta c \\ \delta d \\ \delta e \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} (1-F_i)^b R_i \\ a\sum_{i=1}^{n} (1-F_i)^b \log_e (1-F_i) R_i \\ -\sum_{i=1}^{n} (1-F_i)^{-d} R_i \\ c\sum_{i=1}^{n} (1-F_i)^{-d} \log (1-F_i) R_i \\ -\sum_{i=1}^{n} R_i \end{bmatrix}$$

where δ denotes the change in the indicated parameter and

$$R_i = a(1-F_i)^b + c(1-F_i)^{-d} + e - x_i, \quad i=1,2,\ldots,n$$

where a,b,c,d and e are the Wakeby parameters and x_i indicates the i-th peak flow observation (not necessarily ordered).

Having said so much on the estimation of the Wakeby distribution parameters, it now remains to show how the probability density and cumulative distribution functions can be obtained for a given set of parameters and variate x.

Since (155) is a one-to-one and onto function of x, the probability density function, f(x), can be obtained from

$$f(x) = \frac{dF(x)}{dx} = 1./\left[dx/dF(x)\right] = 1./\{ab\left[1-F(x)\right]^{b-1} +$$

$$cd\left[1-F(x)\right]^{-d-1}\} \quad .... (187)$$

This shows that, unlike in other probability distributions, f(x) is defined explicitly in terms of F(x) rather than in terms of x. Thus, in order to obtain f(x), one needs to solve (155) for F(x) by an appropriate method first. The following is one method of solving for F(x) in (155) for a given set of parameters.

For a given set of parameters a,b,c,d and e, the procedure can be summarised as follows:-

(i) for values of $F(x^*)$ between 0 and 0.9999, generate a series of corresponding values of x^* by use of (155).

(ii) for every observation x, two consecutive values of x^* can be found which envelop the observation x.

(iii) Langrangian interpolation can then be used to estimate the required value of F(x).

(iv) repeat steps (ii) and (iii) for all observations of x to obtain F(x) for all data. Consequently, f(x) can be obtained from (187).

2.9 LOG-WAKEBY DISTRIBUTION

If a variate x is log-Wakeby distributed, then,

$y=\ell n(x)$ is Wakeby distributed. That is

$$f(y)=1/\{ab\left[1-F(y)\right]^{b-1}+cd\left[1-F(y)\right]^{-d-1}\} \quad \text{.... (188)}$$

where

$$y= -a\left[1-F(y)\right]^{b}+c\left[1-F(y)\right]^{-d}+e \quad \text{.... (189)}$$

and a,b,c,d and e are the parameters of the Wakeby distribution in the logarithmic domain. Then

$$f(x)= \frac{1}{x}f(y) \quad \text{.... (190)}$$

and

$$F(x)=F(y) \quad \text{.... (191)}$$

As before, (190) and (191) suggest that it suffices to estimate the parameters of the Wakeby distribution in the logarithmic domain in order to define the log-Wakeby distribution. Methods of estimating the Wakeby parameters have already been discussed. Similarly, the proability density function, f(x), and the cumulative distribution function, F(x), for the log-Wakeby distribution can be evaluated by a method similar to the one described for the Wakeby distribution with the consideration of (190) and (191).

We have so far described the properties and methods of application of the various probability distributions which can be used in flood frequency analysis. However, it is important to mention that, often, it is possible to fit more than one of the probability functions to the same peak flow data. In such cases, tests of goodness of fit may be used to identify the most adequate model(s) . Some of the methods of goodness of fit which can be used in such an identification procedure are discussed in the next sections.

2.10. COMPARISON OF THE FLOOD FREQUENCY DISTRIBUTIONS

The previous sections have described the use of various continuous probability distributions for estimating events at return periods larger than those of the recorded events. The question naturally arises as to which of these distributions can be used in general. Hence, the goodness of fit tests of these probability distribution functions are necessary steps in drawing from the data the most reliable information.

Goodness of fit tests require a measuring parameter for fitting the discrete and continuous distribution. Two such parameters (statistics) are discussed here, the chi-square and the Smirnov-Kolmogorov statistics. These two statistics will also be used to supplement the main goodness of fit procedure to be discussed later.

2.10.1 THE CHI-SQUARE GOODNESS OF FIT TEST

Let a discrete variable have k mutually exclusive random events, or a continuous variable be divided into k class intervals of mutually exclusive events. Let f_i, i=1,2,....,k represent the respective frequencies of these random events or of class intervals for a sample of size N, and let the discrete or continuous distribution functions have the probability of class intervals as p_i, i=1,2,...k. For the observed sample size N, the sample absolute frequencies are $n_i=Nf_i$, in which f_i are the relative frequencies, and the expected values of the absolute frequencies are NP_i, i=1,2,....k.

It has been shown (Kendall and Stuart, 1961) that the chi-square statistic, χ^2, which is given by

$$\chi^2 = \sum_{i=1}^{k} \frac{(n_i - NP_i)^2}{NP_i} \quad \text{.... (192)}$$

has a chi-square distribution with the number of degrees of freedom, NDF=k-1, for a sufficiently large N.

The chi-square test prescribes the critical value χ_o^2, for a given probability level α, so that for $\chi^2 < \chi_o^2$, the null hypothesis of a good fit is accepted, otherwise, for $\chi^2 > \chi_o^2$, it is rejected. Figure 6 shows the distribution of χ_o^2 for various values of α.

In general, the number of class intervals should be $k \geq 5$ and the expected absolute frequencies should be $NP_i > 5$ for the distribution with NDF=k-1 to well approximate the distribution of the statistic

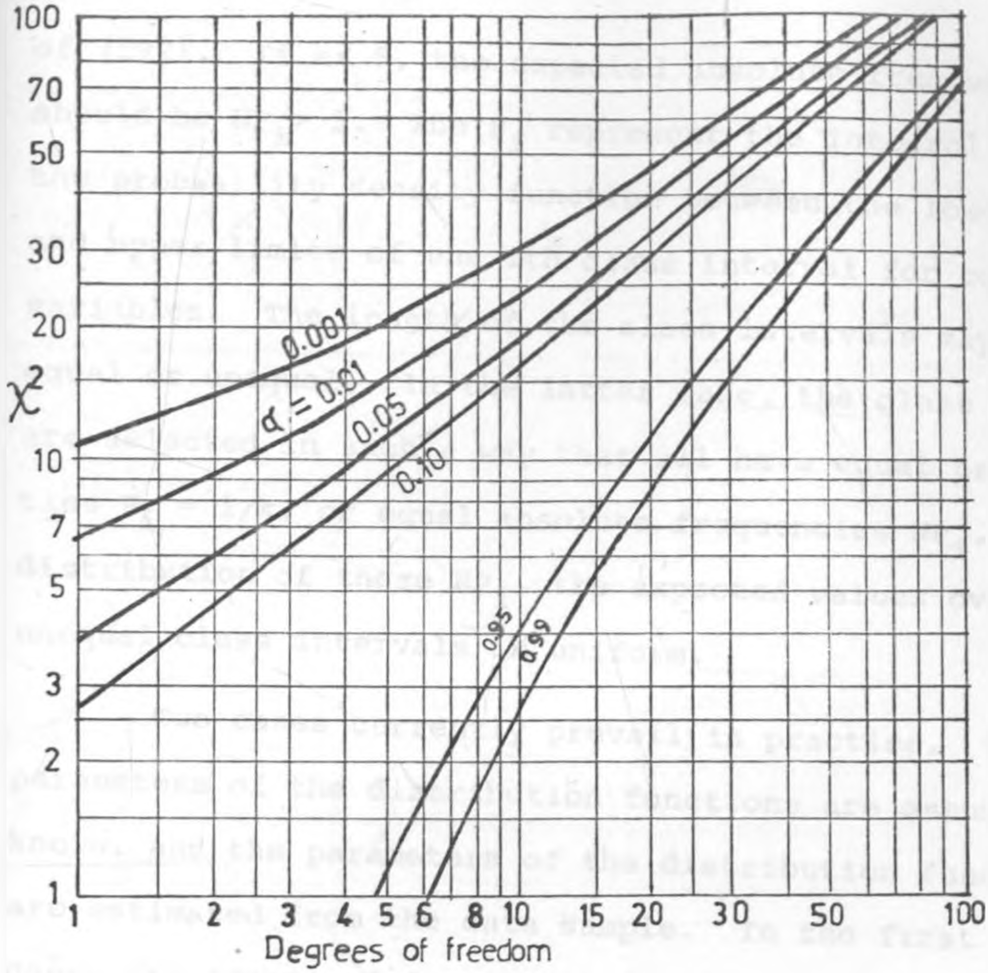

Fig. 6 Probability points of the χ^2 distribution

(Carnahan, et al 1969)

of (192). If $k < 5$, the expected absolute frequencies should be $NP_i > 5$. The P_i represent the integral of the probability density function between the lower and upper limits of the ith class interval for continuous variables. The length of the class intervals may be equal or unequal. In the latter case, the class lengths are selected in such a way that all have equal probabilities $P_i = 1/k$, or equal absolute frequencies NP_i. The distribution of these NP_i, the expected values over the unequal class intervals is uniform.

Two cases currently prevail in practice. The parameters of the distribution functions are assumed known, and the parameters of the distribution functions are estimated from the data sample. In the first case, the number of degrees of freedom NDF=k-1, that is, a degree of freedom is lost because of the constraint that

$$\sum_{i=1}^{k} P_i = 1 \qquad \text{.... (193)}$$

In the second case, the number, h, of parameters of the distribution function, represents additional h constraints, so that NDF=k-h-1.

2.10.2 THE SMIRNOFF KOLMOGOROV TEST

In order to avoid the loss of information due to the grouping suffered by the χ^2 - tests, other methods of goodness of fit have been developed such as the Neyman-Barton (Barton,1953) and the Cramer-Von Mises

W^2-test (Cramer, 1946). The most important of these alternatives to the χ^2-test is the Smirnoff-Kolmogorov statistic, (Kite, 1977). The Smirnoff-Kolmogorov tests is based on the deviations of the sample (emperical) distribution function P(x) from the completely specified continuous hypothetical distribution function F(x); of the continuous variable x, such that

$$\Delta = \text{maximum} \left| P(x) - F(x) \right| \quad \text{.... (194)}$$

Table 7 gives the critical values of the Smirnoff-Kolmogorov, Δ_o, for various values of the sample size N and also for several values of the confidence level α.

TABLE 7: CRITICAL VALUE Δ_0 OF THE SMIRNOFF-KOLMOGOROV STATISTICS Δ, FOR VARIOUS VALUES OF N AND α (YEVJEVICH 1972a).

N	α			
	0.20	0.10	0.05	0.01
5	0.45	0.51	056	0.67
10.	0.32	0.37	0.41	0.49
15	0.27	0.30	0.34	0.40
20	0.23	0.26	0.29	0.36
25	0.21	0.24	0.27	0.32
30	0.19	022	0.24	0.29
35	0.18	0.20	0.23	0.27
40	0.17	0.19	0.21	0.25
45	0.16	0.18	0.20	0.24
50	0.15	0.17	0.19	0.23
N>50	$1.07/\sqrt{N}$	$1.22/\sqrt{N}$	$1.36/\sqrt{N}$	$1.63/\sqrt{N}$

Comparing the Δ - test to the χ^2 - test, there is no condition in a Δ - test that the absolute class frequencies should be greater than five, and no sorting of data in class intervals is required.

Because the parameters of a distribution function F(x), fitted to the emperical distribution P(x), must be estimated from the sample data, the Δ - test is not an exact test, but an approximation. Because Δ is obtained from the absolute maximum difference between F(x) and P(x) under the condition that the parameters of F(x) are estimated from the sample data, Δ is likely to be smaller than the true Δ_t if the distribution parameters are known. The bias $\Delta - \Delta_t$ is not known, however, (Yevjevich,(1972a). Regardless of this bias, a Δ - test is useful and quick, if the critical value Δ_o is selected somewhat smaller than Δ_o of Table (7) (Yevjevich,(1972a).

A typical streamflow record of peak annual events usually contains one or more flood events which may occcur at the higher return period side of the data and which may deviate significantly from the general trend of the other events for whichever probability model is adopted (see Figure 7). Despite the fact that the assumed flood frequency model might be incorrect, these large floods often contain very large measurment errors which may account for this type of deviation. On the other hand, if a truly 100-year return period flood, say, is observed within, say, a 20-year record of annual peak flows, then such an event would be assigned, quite inevitably, a return period of about 20 years. When plotted on any probability paper, this

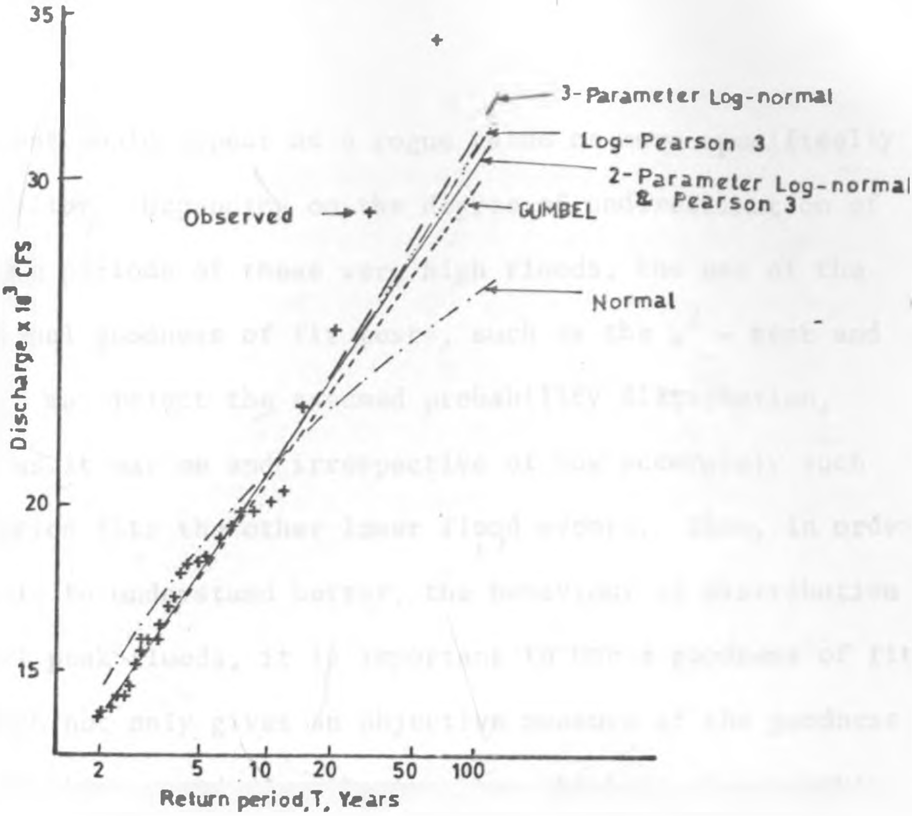

Fig. 7 Comparison of some frequency curves from various distributions. St. Mary's River at Stillwater, U.S.A. 1916 - 1975. (Kite 1977)

flood event would appear as a rogue value or more specifically as an outlier. Depending on the degree of underestimation of the return periods of these very high floods, the use of the conventional goodness of fit tests, such as the χ^2 - test and Δ - test, may reject the assumed probability distribution, correct as it may be and irrespective of how accurately such distribution fits the other lower flood events. Thus, in order to be able to understand better, the behaviour of distribution of annual peak floods, it is important to use a goodness of fit test which not only gives an objective measure of the goodness of fit of the assumed distribution, but which is also capable of detecting the outlying floods. These outlying peak flows, hereafter referred to as outliers occur as a result of measurement error or as a result that their return periods were either underestimated or overestimated. The following section gives a brief review of the theory, of the detection and accommondation of outliers and how the theory can be útilised to construct an objective goodness of fit test statistic.

2.10.3 OUTLIERS AND THE BEST-FITTING MODEL

The intuitive definition of an outlier is "an observation which deviates so much from other observations as to arouse suspicion that is was generated by a different mechanism other than that which generates the majority of the observations."

The concept of an outlier has fascinated experimentalists since the earliest attempts to interpret data. Even before the formal development of statistical methods, arguments ranged over whether or not, and on what basis we should discard observations from a set of observations on the grounds that they are unrepresentative, spurrious, mavericks or rogues. Attitudes have varied from one extreme to another: from the view that we should never sully the purity of the data by daring to adjudge its propriety, to an ultimate pragmatism expressing if in doubt, throw it out.

In observing a set of observations in some practical situation, one or more of the observations jars or stands out in contrast to the other observations usually as extremes or outliers. Such outliers do not fit with the tidy pattern present in our minds at the outset of our enquiry, of what constitutes a reasonable set of data. We have subjective doubts about the propriety of the outlying observations both in relation to the specific data set we have obtained and in relation to our initial views of an appropriate probability model to describe the generation of our data. Our attitudes about the data will in this respect differ quite widely with different possible basic probability models. If we assume a normal distribution for a given set of data, we may react quite strongly to certain observations which would arouse no specific concern if the assumed distribution were longer tailed such as the lognormal or cauchy. The purpose of most of the statisitcal methods for examining outliers is in broad terms to provide means of assessing whether or not our subjective declaration of the presence of outliers in a particular data set has important objective implications for the further analysis of the data. Thus, it is important to analyse the set of data using all the feasible probability models before one makes a final decision that the data contains outliers. Some of the tests which have been used to test for the presence of outliers are discussed below.

2.10.4 TESTS FOR OUTLIERS IN STATISTICAL DATA

Augmenting the classification of Barnett and Lewis, (1979), we can distinguish seven types of outlier - test statistics. These basic types of outliers - test statistics are given below.

2.10.4.1 EXCESS SPREAD STATISTICS

These are the ratios of differences between an outlier and its nearest or next nearest neighbour to the range, or some other measure of spread of the data sample. Examples are:-

(i) the Dixon type statistic (Dixon, 1951) which can be written as $(x_n - x_{n-1})/(x_n - x_2)$ where the subscripts denote the rank order of the data of length n, arranged in ascending order.

(ii) the Irwin type statistic (Irwin, 1925) which can be written as $(x_n - x_{n-1})/\sigma$, where the subscripts have the same meaning as in Dixons test statistic and σ is the population standard deviation.

Irwins statistic assumes that σ is known and is particularly relevant for a normal distribution. Clearly, we could replace σ with an estimate which can be based on a restricted sample which excludes observations we wish to test for; such as the outlier x_n or other extremes (Barnett and Lewis, 1979).

2.10.4.2 RANGE-SPREAD STATISTIC

In this case, the numerator of the Irwins type statistic is replaced with the sample range to get for example, $(x_n-x_1)/S$ (David *et al*, 1954, Pearson and Stephen, 1964). Again, S, the sample standard deviation can be replaced by a restricted sample analogue, independent estimate or a known value of a measure of spread of the population. Using the range has the disadvantage that it is not clear without further investigation whether significant result represent discordancy of an upper outlier, lower outlier or both.

2.10.4.3 DEVIATION-SPREAD STATISTICS

These statistics give a measure of the distance of an outlier from some measure of central tendency in the data. An example for a lower outlier test is $(\bar{x}-x_1)/S$ where $\bar{x}$ is the sample mean (Grubbs, 1950). Both $\bar{x}$ and S can be based on a restricted sample, or replaced with independent estimates, or population values of some convenient measure. A modification uses maximised deviation, for example maximum $|x_i-\bar{x}|/S$ (Halperin *et al.*, 1955), x_i, as before, denotes the ith observation when the data is arranged in ascending order.

2.10.4.4. SUM OF SQUARES STATISTICS

These statistics are expressed as ratios of squares for the restricted and the total sample. For example, the statistic of the form

$$\sum_{i=1}^{n-2}(x_i-\bar{x}_{n,n-1})^2 / \sum_{i=1}^{n}(x_i-\bar{x})^2 \qquad \text{.... (195)}$$

where

$$\bar{x}_{n,n-1} = \sum_{i=1}^{n-2} x_i/(n-2) \qquad \text{.... (196)}$$

and x_i denotes the ith observation when the data is arranged in increasing order and n is the sample size. This statistic was proposed by Grubbs, (1950) for testing two upper outliers x_{n-1} and x_n.

2.10.4.5 HIGHER ORDER MOMENT STATISTICS

Furguson, (1961) showed that statistics such as measures of skewness and kurtosis, although specifically designed for assessing outliers can nevertheless be useful in this context. However, this application was based on the normal alternatives only and no test distribution exists for the other skewed distributions (Barnett and Lewis, 1979).

Another omnibus statistic of relevance for testing for outliers is the W-statistic proposed by Shapiro and Wilk, (1965, 1972) and Shapiro and Wilk, (1968). For normal samples, this statistic consists of the ratio of the square of a particular type of linear combination of all the ordered sample values to the sum of squares of the individual deviations about the sample mean S^2, that is:

$$\left[\sum_{i=1}^{(n/2)} a_{n,n-i+1}(x_{n-i+1}-x_i)\right]^2 / S^2 \qquad \text{.... (197)}$$

where (n/2) denotes the integer part of n/2 and n is the sample size. The coefficients $a_{n,j}$ can be obtained from a table given by Shapiro and Wilk, (1965) and S^2 denotes the sum of squares of the individual deviations about the sample mean, x_j denotes the magnitude of the jth observation when the observations are arranged in ascending order. The Shapiro and Wilk, (1972) statistic for the exponential distribution takes the simple form:

$$n(\bar{x}-x_1)^2 / \left[(n-1)\sum_{j=1}^{n}(x_j-\bar{x})^2\right] \qquad \text{.... (198)}$$

where n is the sample size, x_1 the smallest observation, x_j the jth observation when the data is arranged in ascending order and $\bar{x}$ is the sample mean.

Historically, the motivation for a statistical treatment of outliers came first from the problems of combining astronomical observations, and repeated measurements or determinations must always be one of the main contexts in which discordancy problems arise. In very many cases, errors of measurement may plausibly be assumed to follow a normal distribution, whether through the operation of the central limit theorem on contributory error components, or purely as an empirical fact. It is not surprising therefore that the vast body of published methodology on outliers from the eighteenth century to the present day rest on the working hypothesis of a normal distribution. Indeed, it is only in the last 20 or 30 years that outliers in the exponential (and gamma) models have been specifically considered (Hawkins, 1980).

Barnett and Lewis, (1979) have analysed exhaustively the efficiency, properties and limitations of the above mentioned statistics. One of the most serious disadvantages of these test statistics is the existence of the masking effect. This effect arises when a sample contains more than one outlier. These outliers so increase the spread of the sample that the removal of one outlier makes little improvement in the appearance of the sample, and in particular all values of $(x_i-\bar{x})/S$ are near zero because of the very large value of S (all variables have the same meaning as explained previously). The practical consequences of the masking effect is that any attempt to remove these outliers one at a time proves fruitless.

Furthermore, Tietjen and Moore, (1972) showed that each of these tests is geared to a particular kind of outliers and the application of these procedures to a given set of data at a given level of significance often leads to conflicating conclusions with respect to the null hypothesis. Tiejen and Moore, (1972) also emphasised the importance of the determination of the number of outliers previous to testing - a fact which, not surprisingly leads to 'if you already know how the outliers are distributed in the data, then why test for them?"

Kitagawa, (1979) solved these problems of outlier detection within the context of estimation theory, using the Akaike Information Criterion. With this method, we no longer need a two step procedure of determination of the type and number of outliers and hypothesis testing.

2.10.4.6 THE AKAIKE INFORMATION CRITERION FOR DETECTION OF OUTLIERS.

Akaike (1973) proposed an objective procedure for the identification of an optimal model from a a class of competing models. Kitagawa, (1979) extended the work of Akaike, (1973,1977) and showed that a criterion which he called the Akaike Information criterion, can not only be used to identify the optimal model, but can also be used to detect the existence of any outliers in a given model.

The Akaike Information Criterion (AIC) is given by

AIC= - 2log(maximum likelihood)+2(number of independently adjusted parameters) .... (199)

The AIC has been used for modelling in various fields of statistics (Akaike, 1970, 1976, Otsu *et al.*, 1976), engineering (Otomo *et al.*, 1972) and numerical analysis (Tanabe, 1974).

The Akaike Information Criterion has the desired advantages over all the previously mentioned tests of goodness of fit or outlier detection procedures in that it not only gives an objective and comparative measure of the goodness of fit when tested over various alternatives, but with it, it is possible to detect either low or high outliers without any danger of the masking effect. Thus, with the use of the Akaike Information Criterion, it is now possible to identify, more positively, the optimal flood frequency model from the class of competing frequency models previously discussed. This AIC approach has also been used herein for the detection of outliers in flood data and also for the identification of the best fitting model(s) at each of the river gauging stations used. In this way, it will be possible to choose the overall optimal flood frequency model(s) of general applicability.

Since the Akaike Information Criterion approach is a key method in this project, the details of its development and application are presented in the following sections.

2.10.5 DEVELOPMENT OF THE AKAIKE INFORMATION CRITERION

The maximum likelihood principle has been utilized in two different branches of statistical theories. The first is the estimation theory where the method of maximum likelihood has been used extensively and the second in the test theory where the log-likelihood ratio test plays a very important role in hypothesis testing (Mood *et al.*, 1974). Akaike, (1973) defined a loss function $W(\theta,\hat{\theta})$ and risk function $R(\theta,\hat{\theta})$ where θ is a vector of the true (population) parameter values and $\hat{\theta}$ is the vector of the estimated (sample) parameter values, and suggested that these two parameters can be combined into one problem of statistical decision. Thus, instead of considering a single estimate of the vector θ, the estimates corresponding to various possible restrictions of the distribution can be considered and instead of treating the problem as a multiple decision or test between hypothesis, the problem of general estimation procedure based on the decision theoretic consideration is treated. The whole idea can be simply realized by comparing $R(\theta,\hat{\theta})$ or $W(\theta,\hat{\theta})$ when possible for various values of $\hat{\theta}$ and taking the one with the minimum $R(\theta,\hat{\theta})$ or $W(\theta,\hat{\theta})$ for our final choice. Akaike, (1971) discusses at length, the advantages and disadvantages of this approach.

The problem of statistical model identification is often formulated as a problem of the selection of a density function $f(x,\theta^k)$, $k=1,2,\ldots,L$ based on the observations X, where θ^k is the projection of θ, a vector on L-dimensional real subspace, where L is the number of population parameters, to a lower dimensional real subspace of dimension k. The value of k, which is the number of parameters of the emperical distribution is often called the order of the model. Wald, (1943) shows that this problem can be treated as a subject of composite hypothesis testing and the use of the log-likelihood ratio criterion is well established for this purpose.

Consider the situation where the results x_i, $i=1,2,\ldots,N$ of N independent observations of the variate X have been obtained. Let $\hat{\theta}^k$ denote the maximum likelihood estimates in the space of θ^k, that is, $\hat{\theta}^k$, is the value of the vector θ^k which gives the maximum of the likelihood function

$$\prod_{i=1}^{N} f(x_i, \theta^k) \qquad \text{.... (200)}$$

Then, Akaike, (1973) strongly suggested the use of

$$\omega^k = -\frac{2}{N} \sum_{i=1}^{N} \log\left[\frac{f(x_i, \theta^k)}{f(x_i, \theta)}\right] \qquad \text{.... (201)}$$

as an estimate of the loss function $W(\theta, \hat{\theta}^k)$.

The statistic

$$\eta^k = N.\omega^k \qquad \text{.... (202)}$$

is the familiar log-likelihood ratio test statistic which is asymptotically distributed as a non-central chi-square variable with (L-k) degrees of freedom when the true parameter vector θ is in the space of θ^k.

Assuming the tendency towards a Gaussian distribution of $\sqrt{N}(\hat{\theta}-\theta)$ and the consistency of $\hat{\theta}^k$ and $\hat{\theta}$ as the estimates of θ^k and θ respectively, it can be shown (Akaike, 1973) that, an asymptotic equality in distribution for the log-likelihood ratio statistic η^k exists and is given by

$$\eta^k = N||\hat{\theta}-\theta^k||_c^2 - N||\hat{\theta}^k-\theta^k||_c^2 \qquad \text{.... (203)}$$

where $||\alpha||$ denotes the norm in the space of α and is defined by

$$||\alpha||^2_C = \sum_{\ell=1}^{L} \sum_{m=1}^{L} \alpha_\ell \alpha_m C(\ell,m)(\theta) \quad .... (204)$$

where $C(\ell,m)(\theta)$ is the (ℓ,m) element of the Fisher's information matrix given by

$$C(\ell,m)(\theta) = -\int \{\frac{\partial^2 \log[f(x,\theta)]}{\partial\theta_\ell \partial\theta_m}\} f(x,\theta)dx \quad .... (205)$$

for all x

By using the identity

$$||\hat{\theta}-\theta^k||^2_c = ||(\hat{\theta}-\theta)-(\theta^k-\theta)||^2_c \quad .... (206)$$

and the definition of the norm on the first term of the right hand side of (203), we get

$$\eta^k = N||\theta^k-\theta||^2_c + N||\hat{\theta}-\theta||^2_c - ||\hat{\theta}^k-\theta^k||^2_c - 2N(\hat{\theta}-\theta,\theta^k-\theta)_c$$

.... (207)

Extending (201), Akaike, (1973) defined the loss function $W(\theta,\hat{\theta}^k)$ as

$$W(\theta,\hat{\theta}^k) = \sum_{\ell=1}^{L} \sum_{m=1}^{L} (\hat{\theta}^k_\ell - \theta_\ell)(\hat{\theta}^k_m - \theta_m) C(\ell,m)(\theta) \quad .... (208)$$

which can also be rewritten as:

$$W(\theta,\hat{\theta}^k) = ||\theta^k-\theta||^2_c + ||\hat{\theta}^k-\theta^k||^2_c \quad .... (209)$$

Using (207), $W(\theta,\hat{\theta}^k)$ can be further rewritten as

$$W(\theta,\hat{\theta}^k)=\eta^k-\{N\ ||\hat{\theta}-\theta||_c^2-N||\hat{\theta}^k-\theta^k||_c^2\} +$$

$$2N(\hat{\theta}-\theta,\theta^k-\theta)_c + ||\hat{\theta}^k-\theta^k||_c^2 \quad .... (210)$$

Geometrically, $\hat{\theta}-\theta^k$ is approximately the projection of $\hat{\theta}-\theta$ into the space of θ^k. From this result, Akaike (1973) showed that $N||\hat{\theta}-\theta||_c^2-N||\hat{\theta}^k-\theta^k||_c^2$ and $N||\hat{\theta}^k-\theta^k||_c^2$ are asymptotically independently distributed as chi-square variables with (L-k) and k degrees of freedom respectively. Akaike, (1973) further showed that the standard deviation of the asymptotic distribution of $N(\hat{\theta}-\theta,\ \theta^k-\theta)_c$ is equal to $\sqrt{N}\,||\theta^k-\theta||_c^2$. Thus, if $N||\theta^k-\theta||_c^2$ is of comparable magnitude with L-1 or k and these are large integers, then, the contribution of $N(\hat{\theta}-\theta,\theta^k-\theta)_c$ in (210) remains relatively insignificant. Further, if $N||\theta^k-0||_c^2$ is significantly larger than L, the contribution of $N(\hat{\theta}-\theta,\theta^k-\theta)_c$ is also relatively insignificant. Also, if $N||\theta^k-\theta||_c^2$ is significantly smaller than L and k, again, the contribution of $N(\hat{\theta}-\theta,\theta^k-0)_c$ remains insignificant compared to the other variables of the chi-square type. These observations suggest that (210) can be approximately written as

$$W(\theta,\hat{\theta}^k)=\eta^k-\{N||\hat{\theta}-\theta||_c^2-N||\hat{\theta}^k-\theta^k||_c^2\}+||\hat{\theta}^k-\theta^k||_c^2 \quad .... (211)$$

For a given model, $f(x_i,\hat{\theta}^k)$ is a constant. Hence, taking the statistical expectation operation E on both sides of (211), we get

$$E\left[W(\theta,\hat{\theta}^{k})\right] = \eta^{k}+2k-L \quad \ldots. \ (212)$$

Using (201) and (202), (212), can further be rewritten as

$$E\left[(W(\theta,\hat{\theta}^{k})\right] = -2\sum_{i=1}^{N} \log\left[f(x_i,\hat{\theta}^{k})\right]+2k-\{L+2\sum_{i=1}^{N}\log \left[f(x_i,\theta)\right]\} \quad \ldots. \ (213)$$

The two last terms of (213), in braces, are both constants for a given set of observations x_i, since they represent population values. Hence, these two terms cannot have any decisive power on the model discrimination test based on $E\left[W(\theta,\hat{\theta}^{k})\right]$. On this basis therefore, Akaike, (1973) derived a criterion which is herein called the Akaike Information Criterion for model identification and other statistical inference situations given by

$$AIC = - 2 \sum_{i=1}^{N} \log\left[f(x_i,\hat{\theta}^{k})\right]+ 2k \quad \ldots. \ (214)$$

The interpretation of (214) is given by (199).

The AIC has a clear interpretation in model fitting. The first term of (214) indicates the badness of fit and the second the increased unreliability due to the increased number of parameters. The best approximating model is the one which achieves the most satisfactory (minimum AIC) compromise. The process of finding the model with the minimum AIC is equivalent (approximately) to finding the model with the maximum entropy(Kitagawa, 1979). This approach can be used in the identification of an optimal flood frequency as higlighted in the following subsection.

2.10.6 APPLICATION OF THE AIC IN FLOOD FREQUENCY MODELLING

The Akaike Information Criterion can be utilised in two main ways. Firstly, it can be used purely as a model identification process and secondly, both as a model identification process and as an outlier detection process simultaneously. Obviously, the latter use of the Akaike Information Criterion is more profitable in most statistical analysis in which the basic data is prone to contamination by outliers.

A common type of outlier problem arises in situations where a set of data can be divided into distinct subsamples. The subsamples may correspond with different levels of some set of factors. The subdivision of the sample into the sub-samples may take place after we have collected a random sample from some overall population in which case, subsample, sizes m_i, are random quantities. Alternatively, and more likely, we may choose random samples of prescribed size, m_i, at different factor levels or under different circumstances: these samples in combination serve as subsamples in the overall data set. This is the case in many designed experiments and one of interest, examinable by analysis of variance techniques on the customary assumptions of normality, addidivity and homoscedasticity, is in the comparison of the means or variances of the populations from which the subsamples arise. Barnett and Lewis (1979) have discussed this problem in the context of the other problems in details. In flood studies however, we are mainly concerned with the existence of subsamples of the former type.

Flood flows are the very events subject to the maximum measurement error. In fact, maximum flows are seldom, if ever, measured because of the difficulties involved, firstly in predicting the time at which maximum flow might occur, secondly the difficulty in getting to the gauging site at that time and, finally the difficulties of actually carrying out the gauging at the high stage and high velocities. The automatic recording stations are usually not very helpful during these high flows because they are either clogged up by the characteristically muddy waters or are simply destroyed by the debris or washed away by the fast flowing waters. As a result, high flows are normally estimated by extrapolation of the rating curve, estimation of the mean velocity from an isolated surface velocity measurement, use of slope area method or other similar estimation procedures. The resulting estimates of discharge of these high flows contain a high error component which Blench, (1959) estimated to be at least 25%. This is partly the reason why the highest few flood observations show up as discordant values in relation to the smaller flood values, hence depicting the existence of at least two subsamples.

Another factor which can give rise to subsamples in flood data is the plotting position. It was mentioned earlier that the effect of the plotting position is most serious at both extremes of the flood data. Hence, most often, some of the smallest flood values and some of the highest flood values are unintentionally overestimated or underestimated in terms of their return period, depending on the length of the record.

Consequently, a flood data sample may have three distinct subsamples. Since the lowest and the highest - valued subsamples are generated by mechanisms which either originate from measurement error, underestimation or overestimation, we can refer to the observations in these subsamples, as low outliers and high outliers respectively.

In order to make this existence of subsamples compatible with the requirements of the Akaike Information Criterion analysis, we can assume that the subsamples arise from populations with identical means, with the alternative prospect that in just one or two of the subsamples, the mean has slipped up or down from the predominant level. Or perhaps the variance of one or two of the subsamples is larger than the common variance of the majority of the subsamples.

Due to the uncertainities involved in the computation of the variances of both the low and high outlier subsamples, (since they contain very few observations) we can only use the mean slippage alternative. This alterntive can be simply stated as follows: all the observations have a common probability distribution only that the low and high subsamples have slipped means μ_1 and μ_2 respectively, while the inlying observations have the true mean μ. This can be further interpreted mathematically as follows: Let the sample of n observations be denoted in order of increasing magnitude by $x_1 \leq x_2 \leq \ldots\ldots \leq x_n$. Assume that x_i is a realization of the flood variable specified by a probability function $f(x_i)$. Then, we can write that

$$f(x_i) = \begin{cases} g(x_i, \mu_1), \; i=1,\ldots,n \\ g_{i-n_1, n-n_1-n_2}(x_i, \mu), \; i=n+1,\ldots n-n_2 \\ g(x_i, \mu_2), \; i=n-n_2+1,\ldots n \end{cases} \quad \ldots. (215)$$

where f(.) is the underlying probability distribution which in this project can be any of the previously discussed probability density functions, and $g_{\ell,m}(.,.)$ is the probability density function of the ordered observations $x_\ell \leq x_{\ell+1} \leq \ldots \leq x_m$, for any ℓ less than m and m less than n, from the population with density g(.,.). The integers n_1, and n_2 denote the sizes of the lower and upper subsamples respectively. It is important to note that (Mood et al., 1974).

$$g_{i-n_1, n-n_1-n_2}(x_i, \mu) = C\left[(i-n_1), (n-i-n_2+1)\right]$$

$$\left[G(x_i, \mu)\right]^{i-n_1-1} . \{1-G(x_i, \;)\}^{n-i-n_2} g(x_i, \mu) \quad \ldots. (216)$$

where

$$C\left[(i-n_1), (n-i-n_2)\right] = \frac{(n-n_1-n_2)!}{(i-n_1-1)!(n-i-n_2)!} \quad \ldots (217)$$

and

$$G(x_i, \mu) = \int_{-\infty}^{x_i} g(\xi, \mu) d\xi \quad \ldots. (218)$$

The model of (215) means that n_1 observations $x_1, \ldots x_{n_1}$, $n-n_1-n_2$ observations $x_{n_1+1}, \ldots, x_{n-n_2}$ and n_2 observations $x_{n-n_2+1}, \ldots, x_n$ each are realizations of the distribution $g(.,.)$ with common parameters . except the means which are μ_1, μ and μ_2 respectively. In this project we will consider the observations in the lower and upper subsamples as outliers.

The likelihood function $L(x; n_1, n_2, \mu_1, \mu, \mu)$ of (215) is given by

$$L(x; n_1 n_2 \mu_1, \mu_2) = \prod_{i=1}^{n_1} g(x_i, \mu_1) . \prod_{i=n_1+1}^{n-n_2} g_{i-n_1, n_1-n_2}(x_i, \mu) .$$

$$\prod_{i=n-n_2+1}^{n} g(x_i, \mu_2) \qquad \ldots . (219)$$

Thus, the log-likelihood function $\ell(x; n_1, n_2, \mu_1, \mu_2, \mu)$ of (219) is given by

$$\ell(x; n_1, n_2, \mu_1, \mu_2 \mu) = \sum_{i=1}^{n_1} \log\left[g(x_i, \mu_1)\right] + \log(n-n_1-n_2)!$$

$$- \sum_{i=n_1+1}^{n-n_2} \{\log\left[(i-n_1)! + \log(n-i-n)!\right] + (i-n_1-1) \log G(x_i, \mu) +$$

$$(n-i-n_2) \log\left[1-G(x_i, \mu)\right] + \log\left[g(x_i, \mu)\right]\} + \sum_{i=n-n_2+1}^{n}$$

$$\log\left[g(x_i, \mu_2)\right] \qquad \ldots . (220)$$

Hence, it is now clear that the minimum Akaike Information Criterion of the best approximating numbers of the outliers in the low side and high side are those values of n_1 and n_2 which minimize AIC(i,j) given by

$$AIC(i,j) = \begin{cases} -2\ell(x;i,j,.,.\mu) + 2k_1, & i = j = 0 \\ -2\ell(x;i,j,.,\mu_2,\mu) + 2k_2, & i = 0,\ j \neq 0 \\ -2\ (x;i,j,\mu_1,.,\mu) + 2k_3, & i \neq 0,\ j = 0 \\ 2\ (x;i,j,\mu_1,\mu_2,\mu) + 2k_4, & i \neq 0,\ j \neq 0 \end{cases} \quad \ldots.\ (221)$$

with the maximum likelihood estimates μ_1, μ_2 and μ and all other common parameters. The integers k_1, k_2, k_3 and k_4 in (221) are the number of independently estimated parameters when no outliers are assumed, when there are no high outliers, when there are no low outliers and when there are outliers both in the low and upper subsamples respectively.

Hence, it is clear that by changing the underlying density function f(.,.), the best fitting distribution can be obtained as the one with an overall minimum Akaike Information Criterion (MAIC). The required number of low and high outliers would then be those values of i and j in (221) which correspond to the MAIC.

Theoretically, i and j in (221) can attain any integer values such that $i+j \leq n$. However, experieince

(Kite, 1977; Morel-Seytoux, 1981 and Greis, 1983) show that when discordant flood observations do exist, there are hardly ever more than three of them on either side of the extremes. Thus, in this project the minimum Akaike Information Criterion MAIC is found from all the AIC (i,j) in (221) such that $i \leq 3$ and $j \leq 3$.

The use of the mean slippage alternative which was utilized to arrive at (221) requires that in each of the nine chosen flood probability distributions, the estimate $\hat{\mu}$ of the inlying subsample remain unchanged throughout the changes of i and j. The total sample mean gives a good estimate of $\hat{\mu}$. The estimates $\hat{\mu}_1$ and $\hat{\mu}_2$ are obtained from the subsample means of the lower and upper subsamples respectively. This formulation is quite appropriate for a distribution which contains the μ, as one of its parameters. However, the probability distributions in consideration do not always possess this property. Nevertheless, at least one of the distribution parameters is a function of the mean when the other higher moment sample statistics are held constant as is required by the mean slippage alternative. This is especially true with the moment parameter estimates. For the sake of convenience, these moments parameter relationship are used to define the variations of a required parameters with variations of the mean. Thus, through $\hat{\mu}_1$, and $\hat{\mu}_2$

the concerned parameter or parameters can be changed accordingly in the lower and upper subsamples.

In the three parameter lognormal distribution, the overall estimates of the location parameter $\hat{x}_o$ and the shape parameter $\hat{\sigma}_y$ in (28) are assumed common in the three subsamples. The scale parameter $\hat{\mu}_y$ of the inlying subsample is taken as the one that is computed from the total sample, while $\hat{\mu}_y^{(1)}$ and $\mu_y^{(2)}$ the scale parameters of the lower and upper subsample are obtained as follows

$$\hat{\mu}_y^{(1)} = \frac{1}{n_1} \sum_{i=1}^{n_1} \ell n(x_i - x_o) \qquad \ldots. \ (222)$$

and

$$\hat{\mu}_y^{(2)} = \frac{1}{n_2} \sum_{i=1-n_2+1}^{n} \ell n(x_i - x_o) \qquad \ldots. \ (223)$$

where n_1 and n_2 are the lengths of the lower and upper subsamples and n is the length of the total sample of the variate x.

In the Pearson and log Pearson Type 3 distributions, the scale parameter estimate $\hat{\alpha}$ and the shape parameter estimate $\hat{\beta}$ of the subsamples are taken as those estimated from the total sample. In both the Pearson Type 3 and the log Pearson Type 3 the location parameter $\hat{x}_o$ (in Pearson) and $\hat{y}_o$ (in log Pearson) of the inlying subsample are taken as those computed from the whole sample. However, the location parameters $x_o^{(1)}$ and $x_o^{(2)}$ of the lower and upper subsamples for the Pearson Type 3 are obtained from,

$$\hat{x}_o^{(1)} = \frac{1}{n_1} \sum_{i=1}^{n_1} x_i + \hat{\sigma}\sqrt{\hat{\beta}} \quad \text{.... (224)}$$

and

$$x_o^{(2)} = \frac{1}{n_2} \sum_{i=n-n_2+1}^{n} x_i + \hat{\sigma}\sqrt{\hat{\beta}} \quad \text{.... (225)}$$

where n_1, n_2 and n are as defined before and $\hat{\sigma}_2$ is the sample estimate of the variance of the untransformed variate x. The estimate of the location parameter $\hat{y}_o^{(1)}$ of the lower subsample and the location $\hat{y}_o^{(2)}$ of the upper subsample for the log Pearson Type 3 are obtained from

$$\hat{y}_o^{(1)} = \frac{1}{n_1} \sum_{i=1}^{n_1} \ln(x_i) + \hat{\sigma}_y\sqrt{\hat{\beta}} \quad \text{.... (226)}$$

and

$$\hat{y}_o^{(2)} = \frac{1}{n_2} \sum_{i=n-2_2+1}^{n} \ln(x_i) + \hat{\sigma}_y\sqrt{\hat{\beta}} \quad \text{.... (227)}$$

where $\hat{\sigma}_y^2$ is the estimate of the variance of the transformed variate $y = \ln(x)$ and $\hat{\beta}$ is an estimate of the shape parameter for the total sample in the log domain.

In the AIC analysis with the Fisher Tippet and log-Fisher Tippet distribution, the scale parameter estimate $\hat{\alpha}$ and the shape parameter estimate from the total data sample are assumed to be common in all the subsamples. The location parameter then, which is a function of the mean when $\hat{\alpha}$ and $\hat{\beta}$ are invariable is then the one which is changed within the subsamples. The location parameter x_o of the Fisher Tippet distribution and the location parameter $\hat{y}_o$ of the log-Fisher Tippet distribution of the inlying subsample is again taken as the one computed from the total sample. However, the location parameters $\hat{x}_o^{(1)}$ and $x_o^{(2)}$ for the lower and upper subsamples in the Fisher Tippet distribution are obtained as follows:

$$\hat{x}_o^{(1)} = \frac{1}{n_1} \sum_{i=1}^{n_1} x_i + \frac{\hat{\alpha}}{\hat{\beta}} \left[1-\Gamma(1+\hat{\beta})\right] \quad \ldots\ldots (228)$$

and

$$\hat{x}_o^{(2)} = \frac{1}{n_2} \sum_{i=n-n_2+1}^{n} x_i + \frac{\hat{\alpha}}{\hat{\beta}} \left[1-\Gamma(1+\hat{\beta})\right] \quad \ldots\ldots 229)$$

where n_1, n_2 and n are as defined before and Γ is the gamma function of (55). The corresponding parameter estimates $\hat{y}_o^{(1)}$ and $\hat{y}_o^{(2)}$ for the log-Fisher Tippet distribution are obtained as follows.

$$\hat{y}_o^{(1)} = \frac{1}{n_1} \sum_{i=1}^{n_1} \ell n(x_i) + \frac{\hat{\alpha}}{\hat{\beta}} \left[1-\Gamma(1+\hat{\beta})\right] \quad \ldots\ldots (230)$$

and

$$\hat{y}_o^{(2)} = \frac{1}{n_2} \sum_{i=n-n_2+1}^{n} \ell n(x_i) + \frac{\hat{\alpha}}{\hat{\beta}} \left[1-\Gamma(1+\hat{\beta})\right] \quad \ldots\ldots (231)$$

In (230) and (231), the parameters $\hat{\alpha}$ and $\hat{\beta}$ are for the log-transformed data unlike the $\hat{\alpha}$ and $\hat{\beta}$ in (228) which are computed from the untransformed x-variate.

The scale parameter α and the shape parameter β of the Walter Boughton and log-Walter Boughton distributions are independent of the mean μ of the variate. Hence, under the mean slippage alternative, the estimates $\hat{\alpha}$ and $\hat{\beta}$ as computed from the total sample are assumed common to all the subsamples. However, the location parameter x_o (in the Walter Boughton distribution) or y_o (in the log-Walter Boughton distribution) is a function of the mean of the variate under consideration. Thus, the location parameter has to be changed accordingly. The estimate of x_o or y_o for the inlying subsample is taken as that which is computed from the total sample. In the Walter-Boughton distribution analysis, the location parameter $\hat{x}_o^{(1)}$ of the upper subsample is obtained from

$$\hat{x}_o^{(1)} = \hat{x}_o + \frac{1}{n_1} \sum_{i=1}^{n_1} x_i - \frac{1}{n_1} \sum_{i=1}^{n_1} x_i \qquad \ldots. (232)$$

and

$$\hat{x}_o^{(2)} = \hat{x}_o + \frac{1}{n_2} \sum_{i=n-n_2+1}^{n} (x_i) - \frac{1}{n_2} \sum_{i=n-n_2+1}^{n} x_i \qquad \ldots(233)$$

where $\hat{x}_o$ is the estimate of the location parameter for the total sample of the untransformed variate x; n_1, n_2 and n are as defined before. For the log-Walter Boughton distribution, the corresponding location parameter estimates $\hat{y}_o^{(1)}$ and $\hat{y}_o^{(2)}$ of the lower and upper sub-samples are given

$$\hat{y}_o^{(1)} = \hat{y}_o + \frac{1}{n_1} \sum_{i=1}^{n_1} \ell n(x_i) - \frac{1}{n_1} \sum_{i=1}^{n_1} \ell n(x_i) \quad \ldots (234)$$

and

$$\hat{y}_o^{(2)} = \hat{y}_o + \frac{1}{n_2} \sum_{i=n-n_2+1}^{n} \ell n(x_i) \; \frac{1}{n_2} \sum_{i=n-n_2+1}^{n} \ell n(x_i)$$

where $\hat{y}_o$ is the location parameter of the Walter-Boughton distribution in the log-domain for the total sample.

Finally, in the Wakeby and log-Wakeby distributions, the estimates of the parameters a,b,c and d as computed from the total sample are assumed common to all the subsamples. This way, the parameter e becomes a function of the expected values of the variate in question. Thus, in this case, the estimate of the parameter e is changed (within the samples) accordingly. The estimate $\hat{e}$ (in the real or log-domain) of the in-lying subsample is assumed to be that estimated from the whole sample while $\hat{e}_1$ of the lower subsample and $\hat{e}_2$ of the upper subsample are obtained from

$$\hat{e}_1 = \mu_1 - \hat{a}/(1+\hat{b}) + \hat{c}/(1+\hat{b}) \quad \ldots (236)$$

and

$$\hat{e}_2 = \mu_2 - \hat{a}/(1+\hat{b}) + \hat{c}/(1+\hat{b}) \quad .... \ (237).$$

where μ_1 is the lower subsample mean of the untransformed data in the case of the Wakeby and of the log transformed data in the case of the log-Wakeby and μ_2 is the upper sample mean of either the untransformed data in the case of the Wakeby or of the log-transformed data in the case of the log-Wakeby. Further, the estimates $\hat{a}, \hat{b}, \hat{c}$ and $\hat{d}$ in (236) and (237) for the Wakeby (real domain) are different from those of the log-Wakeby (log domain).

Table 8 below gives the appropriate values of k_1, k_2, k_3 and k_4 of (221) for the various distributions under consideration.

Using the methods which have been reviewd above, the Akaike Information Criterion was used to identify the most optimum flood frequency model, as well as the outliers relevant to the nine probability models described earlier. The results are presented and discussed in the next chapter. However, for completeness sake, the distribution parameter estimates for each of the nine probability models and for each of the chosen catchments are presented and discussed first in the next chapter.

TABLE 8 VALUES OF k_1, k_2, k_3 AND k_4 OF (221)

DISTRIBUTION	k_1	k_2	k_3	k_4
Lognormal Type 3	3	4	4	5
Pearson Type 3	3	4	4	5
Log-Pearson Type 3	3	4	4	5
Fisher Tippet Type 2 or 3	3	4	4	5
Log Fisher Tippet Type 2 or 3	3	4	4	5
Walter Boughton	3	4	4	5
Log-Walter Boughton	3	4	4	5
Wakeby	5	6	6	7
Log-Wakeby	5	6	6	7

CHAPTER 3

3.0 RESULTS AND DISCUSSION

In this section, we will present the major results of the study under two subsections. In the first subsection, the parameters obtained for the various distributions are presented while in the final subsection, the results of the goodness-of-fit tests are presented.

3.1 PARAMETERS OF THE DISTRIBUTIONS

It is often accepted that the maximum likelihood method is the most efficient in the estimation of the parameters of a probability distribution. Furthermore, the Akaike Information theory requires that the maximum likelihood estimates should be used in the computation of the various alternatives of the Akaike Information Criteria given in (221).

Thus, except for the Wakeby and the log-Wakey distributions for which there were no maximum likelihood routines for the estimation of their parameters, the maximum likelihood estimates were used, whenever possible, in all the other probability distributions. However, the maximum likelihood routines for most of the skewed distributions, the flood probability distributions included, involve trial and error or successive approximation procedures which unfortunately do not always guarentee convergence to the required solutions. An error limit of 0.00001 was used in all the maximum likelihood routines. Further, whenever 200 iterations were reached without the required convergence error limit, then the routine in questions was assumed to have no convergence for the catchment in consideration. In such case, the estimates by the method of moments were computed and adopted except for the flood probability distributions which do not posses methods of moments, such as the Walter Boughton, Log-Walter Boughton, Wakeby and the Log-Wakeby distributions. The alternative to the method of maximum likelihod in the cases of the Walter Boughton and the Log-Walter Boughton distributions was the method of least squares. The parameters of the Wakeby and Log-Wakeby distributions were computed solely by the method of probability weighted moments.

In Table 9(a) through 9(i), the estimates of the parameters of the various distributions under investigation are presented. Together with these parameter estimates, it is also remarked against each river gauging station (RGS) which estimation routine accomplished to give the shown estimates. The method of maximum likelihood is denoted by MLE;

TABLE 9 (a): PARAMETERS OF THE LOG-NORMAL 3 DISTRIBUTION

RGS:NO.	PARAMETERS LOCATION x_o	SCALE μ_y	SHAPE σ_y	SAMPLE STATISTICS MEAN	STD. DEV.	SKEW	METHOD USED
1CE1	-623.20	8.44	0.67	5073.1	3688.0	1.05	MLE
1CB2	20.48	4.52	1.29	225.5	325.8	3.21	"
1BA1	27.66	5.50	1.03	418.2	380.9	1.48	"
1BC1	—	—	—	613.3	163.6	-0.37	"
1BG4	43.83	3.95	1.01	127.8	89.6	1.87	"
1BD1	-44.99	5.99	0.80	492.9	416.4	1.13	"
1BE1	-647.60	7.10	0.14	577.6	171.8	0.60	"
1BE2	5.19	2.34	0.94	21.4	17.3	2.21	"
1BE5	56.35	5.06	0.53	236.7	97.7	1.14	"
1BB1	-55.15	6.52	0.54	729.2	472.1	1.97	"
1GG1	11.16	6.83	1.28	2269.5	5355.4	5.57	"
1JG1	-1213.33	8.92	0.53	7377.4	4716.2	1.09	"
1KA5	-1.57	2.68	0.65	16.5	13.2	1.74	"
1LA3	—	—	—	1027.5	303.4	-0.17	"
1CB5	—	—	—	636.2	289.0	-0.17	"
1CB3	6.26	4.54	0.79	136.9	140.9	3.84	"
1CD1	-24.22	5.03	0.74	176.7	166.9	2.27	"
1CA2	—	—	—	894.5	472.3	-0.16	"
1CB1	73.95	5.93	1.15	763.8	787.8	1.72	"
1DA2	87.09	8.73	0.51	7207.6	4289.8	2.37	"
1FG1	-576.71	8.20	0.33	3274.8	1287.8	0.75	"
1JCL3	4.76	2.23	1.02	21.7	31.3	4.96	
1KC3	-389.62	9.38	0.70	14574.0	11049.1	1.64	"
1FG2	1235.16	7.76	0.73	4263.5	2298.2	1.49	"
1HA4	152.85	6.14	0.88	847.2	914.8	5.07	"

TABLE 9(a): PARAMETER OF THE LOG-NORMAL 3 DISTRIBUTION

	PARAMETERS			SAMPLE STATISTICS			
RGS. NO.	LOCATION x_o	SCALE μ_y	SHAPE σ_y	MEAN	STD. DEV.	SKEW	MEAN METHOD
1FF2	80.65	4.87	1.12	298.7	207.0	1.53	MLE
1FE2	83.91	7.97	0.70	3744.9	2676.9	1.50	"
1FE1	—	—	—	2378.9	784.5	0.00	"
2B2	24.72	4.52	0.35	122.7	35.9	1.16	"
2EB3	3.41	4.71	0.90	171.5	195.1	3.72	"
2EC2	0.32	3.01	0.94	30.2	24.7	1.01	"
2FA2	- 18.07	4.28	0.66	70.2	55.6	0.92	"
2FC5	6.21	4.59	2.04	343.6	407.5	1.32	"
2EC3	1.25	3.49	1.15	66.0	100.7	3.66	"
2GD2	- 0.05	4.59	2.44	397.2	444.7	1.14	"
3BA10	7.26	3.90	1.69	202.0	451.2	4.09	"
3BA17	2.33	2.37	1.05	19.9	19.0	2.44	"
3BC12	68.26	6.97	1.28	2079.1	1963.4	0.85	"
3F2	3366.80	9.11	0.99	17834.3	16225.9	2.50	"
3DA2	200.03	9.02	1.07	13331.1	10989.4	0.75	"
3BA18	- 0.01	3.94	0.85	71.6	60.2	1.79	"
3BB10	- 14.46	4.22	0.53	63.7	40.6	0.71	"
4AA2	85.62	5.78	0.98	604.3	544.4	1.58	"
4CA2	0.00	7.35	0.78	2102.9	1970.9	3.69	"
4BE1	- 178.63	8.07	0.71	3857.0	2897.2	1.64	"
4AA4	55.84	5.34	1.23	478.4	570.9	3.00	"
4BC2	0.00	8.55	0.52	5885.8	3139.1	1.17	"
4AB5	75.57	6.27	0.81	797.8	630.4	2.11	"

TABLE 9(a): PARAMETERS OF THE LOG-NORMAL 3 DISTRIBUTION

	PARAMETERS			SAMPLE STATISTICS			
RGS. NO.	LOCATION x_o	SCALE μ_y	SHAPE σ_y	MEAN	STD. DEV.	SKEW	METHOD USED
4G1	7745.93	9.95	1.19	52515.1	73639.6	2.99	MLE
4CB4	-4098.96	8.63	0.14	1560.9	816.8	0.11	"
4BB1	0.00	6.32	0.55	643.4	360.2	1.08	"
4AA1	0.00	5.90	0.78	483.3	357.6	0.90	"
5AA1	- 1.79	5.72	0.73	396.2	352.0	3.27	"
5AA5	- 14.44	5.26	0.78	238.9	173.7	0.63	"
5AB2	21.03	5.09	1.37	342.1	302.1	0.75	"
5BC4	34.78	6.68	0.81	1112.6	840.9	1.10	"
5BC8	19.22	5.92	1.68	1165.8	1725.0	2.44	"
5BE4	- 11.43	5.55	0.85	354.7	341.0	1.89	"
5BE20	358.05	6.51	1.04	1457.1	1089.8	1.32	"
5D5	680.39	7.08	1.14	2717.1	2042.0	1.49	"

TABLE 9(b): PARAMETERS OF THE PEARSON TYPE III DISTRIBUTION

RGS. NO.	PARAMETERS LOCATION x_o	SCALE α	SHAPE β	SAMPLE STATISTICS MEAN	STD.DEV.	SKEW	METHOD USED
1CE1	271.8	3161.9	1.52	5073	3688.0	1.05	MLE
1CB2	-365876.8	0.3	1302143.0	225.5	325.8	3.21	"
1BA1	-344221.8	0.4	847919.0	418.2	380.9	1.48	"
1BC1	———	———	———	613.3	163.6	-0.37	MOM
1BG4	32.0	83.7	1.15	127.8	89.6	1.87	"
1BD1	- 242.1	235.9	3.12	492.9	416.4	1.13	"
1BE1	- 280.4	33.1	25.95	577.6	171.8	0.60	MLE
1BE2	5.7	19.1	0.82	21.4	17.3	2.21	MOM
1BE5	92.0	66.5	2.17	236.7	97.7	1.14	MLE
1BB1	87.3	300.7	2.13	729.2	472.1	1.97	MLE
1GG1	-2851881.4	9.6	295930.0	2269.5	5355.4	5.57	MLE
1JG1	103.3	2921.1	2.49	7377.4	4716.2	1.09	"
1KA5	0.4	9.0	1.78	16.5	13.2	1.74	"
1LA3	———	———	———	1027.5	303.4	-0.17	MOM
1CB5	- 88492.2	0.9	99393.0	636.2	289.0	-0.17	MLE
1CB3	63.5	270.7	0.0	136.9	140.9	3.84	MOM
1CD1	-196474.8	0.1	1418741.0	176.7	166.9	2.27	MLE
1CA2	-136423.3	1.6	88331.0	894.5	472.3	-0.16	MLE
1CB1	- 154.6	675.7	1.36	763.8	789.8	1.72	MOM
1DA2	1279.9	2439.2	2.43	7207.6	4289.8	2.37	MLE
1FG1	451.0	584.1	4.83	3274.8	1287.8	0.75	MLE
1JC13	-16664.6	0.1	297762.0	21.7	31.3	4.96	MLE
1KC3	1095.5	9057.6	1.49	14574.0	11049.1	1.64	MOM
1FG2	1649.9	2103.1	1.24	4263.5	2298.2	1.79	MLE

TABLE 9(b): PARAMETERS OF THE PEARSON TYPE III DISTRIBUTION

RGS. NO.	PARAMETERS LOCATION x_o	SCALE α	SHAPE β	SAMPLE STATISTICS MEAN	STD.DEV.	SKEW	METHOD USED
1HA4	486.3	2310.6	0.0	847.2	914.8	5.07	MOM
1FF2	-111872.4	0.4	307832.0	298.7	207.0	1.53	MLE
1FE2	689.6	2956.2	1.03	3744.9	2676.9	1.50	MLE
1FE1	-791162.7	0.8	1023122.0	2378.9	734.5	0.00	MOM
2B2	45.6	15.5	4.99	122.7	35.9	1.16	MLE
2EB3	66.5	362.4	0.0	171.5	195.1	3.72	MOM
2EC2	-13436.2	0.0	305850.0	30.2	24.7	1.01	MLE
2FA2	- 51.0	25.5	4.76	70.2	55.6	0.92	MOM
2FC5	- 273.8	268.9	2.30	343.6	407.5	1.32	MOM
2EC3	11.0	184.4	0.0	66.0	100.7	3.66	MOM
2GD2	-200196.4	1.0	209283.0	397.2	444.7	1.14	MLE
3BA10	- 18.8	922.2	0.24	202.0	451.2	4.09	MOM
3BA17	4.2	23.2	1.0	19.9	19.0	2.44	MOM
3BC12	- 2523.7	837.5	5.50	2079.1	1963.4	0.85	MOM
3F2	-6840634.5	36.9	186035.0	17834.3	16225.9	2.56	MLE
3DA2	-4211280.2	27.4	153953.0	13331.1	10989.4	0.75	MLE
3BA18	4.6	54.0	1.24	71.6	60.2	1.79	MOM
3BB10	- 50.6	14.4	7.94	63.7	40.6	0.71	MOM
4AA2	- 85.7	429.6	1.61	604.3	544.4	1.58	MOM
4CA2	1033.3	3631.8	0.0	2102.9	1970.9	3.69	MOM
4BE1	316.9	2371.0	1.79	3859.0	2897.2	1.64	MOM
4AA4	97.7	856.3	0.0	478.4	570.9	3.00	MOM
4BC2	1343.4	2187.4	2.08	5885.8	3139.1	1.17	MLE
4AB5	201.5	666.3	1.0	797.8	630.4	2.11	MOM
4G1	3197.5	109956.6	0.45	52515.1	73639.6	2.99	MOM

TABLE 9(b): PARAMETERS OF THE PEARSON TYPE III DISTRIBUTION

	PARAMETERS			SAMPLE STATISTICS			
RGS.NO.	LOCATION x_o	SCALE α	SHAPE β	MEAN	STD.DEV.	SKEW	METHOD USED
4CB4	-497.0	348.4	5.91	1560.9	816	0.11	MLE
4BB1	136	262.5	1.93	643.4	360.2	1.08	MLE
4AA1	-309.8	161.2	4.92	483.3	357.6	0.90	MOM
5AA1	181.0	575.8	0.00	396.2	352.0	3.27	MOM
5AA5	-224504.1	0.1	1751306.00	238.9	173.7	0.63	MLE
5AB2	-462.1	113.5	7.09	342.1	302.1	0.75	MOM
5BC4	-414.1	463.2	3.30	1112.6	840.9	1.10	MOM
5BC8	-246.9	2106.4	0.67	1165.8	1725.0	2.44	MOM
5BE4	-5.7	322.7	1.12	354.7	341.0	1.89	MOM
5BE20	-190.8	720.7	2.29	1457.1	1089.8	1.32	MOM
5D5	-28.7	1518.6	1.81	2717.1	2042.0	1.49	MOM

TABLE 9(c): PARAMETERS OF THE LOG-PEARSON TYPE 3 DISTRIBUTION

	P A R A M E T E R S			SAMPLE STATISTICS			
RGS. NO.	LOCATION y_o	SCALE α	SHAPE β	MEAN	STD.DEV.	SKEW	METHOD USED
1CE1	- 341.8	0.0	179336.0	8.2	0.8	-0.57	MLE
1CB2	2.6	0.4	5.15	4.8	1.0	0.65	MLE
1BA1	-5.5	0.1	161.0	5.7	0.9	0.08	MLE
1BC1	—	—	—	6.4	0.3	-0.91	MOM
1BG4	3.8	0.4	2.26	4.7	0.6	0.95	MLE
1BD1	—	—	—	5.8	1.0	-0.78	MOM
1BE1	-114.8	0.0	153244.0	6.3	0.3	-0.67	MLE
1BE2	1.6	0.3	3.84	2.8	0.6	0.87	MLE
1BE5	3.5	0.1	23.02	5.4	0.4	0.28	MLE
1BB1	—	—	—	6.4	0.6	-0.18	MOM
1GG1	-70.3	0.0	3840.0	6.9	1.3	0.05	MLE
1JG1	-333.4	0.0	256219.0	8.7	0.7	-0.63	MLE
1KA5	—	—	—	2.5	0.8	-0.51	MOM
1LA3	-105.6	0.0	131927.0	6.9	0.3	-0.48	MLE
1CB5	—	—	—	6.3	0.6	-1.39	MOM
1CB3	-0.9	0.1	58.00	4.6	0.7	0.39	MLE
1CD1	—	—	—	4.8	1.0	-0.86	MOM
1CA2	—	—	—	6.6	0.9	-2.16	MOM
1CB1	3.7	0.3	7.11	6.2	0.9	0.43	MLE
1DA2	-1.3	0.0	392.00	8.7	0.5	0.17	MLE
1FG1	-834.8	0.0	453364.00	8.0	0.4	-0.24	MLE
1JC13	1.3	0.3	4.85	2.8	0.7	1.71	MLE
1KC3	—	—	—	9.3	0.7	-0.17	MOM
1FG2	7.0	0.2	6.28	8.2	0.5	0.54	MLE
1HA4	4.2	0.2	12.98	6.5	0.6	0.76	MLE

TABLE 9(c): PARAMETERS OF THE LOG-PEARSON TYPE 3 DISTRIBUTION

RGS. NO.	PARAMETERS LOCATION y_o	SCALE α	SHAPE β	SAMPLE STATISTICS MEAN	STD.DEV.	SKEW	METHOD USED
1FF2	4.4	0.5	2.16	5.5	0.7	0.29	MLE
1FE2	—	—	—	8.0	0.7	-0.03	MLE
1FE1	—	—	—	7.7	0.4	-0.63	MOM
2B2	2.6	0.0	62.00	4.8	0.3	0.32	MLE
2EB3	-10.6	0.0	317.00	4.8	0.9	0.13	MLE
2EC2	—	—	—	3.0	0.9	-0.24	MOM
2FA2	—	—	—	3.8	1.1	-1.18	MOM
2FC5	-549.3	0.0	113011.00	4.8	1.7	-0.32	MLE
2EC3	-4.7	0:1	59.00	3.8	1.1	0.29	MLE
2GD2	—	—	—	4.6	2.5	-1.23	MOM
3BA10	1.8	0.9	2.81	4.2	1.4	0.70	MLE
3BA17	-0.7	0.2	16.33	2.6	0.8	0.29	MLE
3BC12	—	—	—	7.1	1.2	-0.22	MOM
3F2	7.9	0.3	5.40	9.5	0.7	0.67	MLE
3DA2	—	—	—	9.1	1.0	-0.38	MOM
3BA18	—	—	—	3.9	0.9	-0.23	MOM
3BB10	—	—	—	3.9	0.7	-0.55	MOM
4AA2	4.2	0.3	6.25	6.1	0.8	0.67	MLE
4CA2	—	—	—	7.3	0.8	-0.00	MOM
4BE1	-255.5	0.0	118507.00	8.0	0.8	-0.30	MLE
4AA4	3.9	0.5	3.70	5.7	0.9	0.60	MLE
4BC2	-861.8	0.0	2777431.00	8.5	0.5	0.00	MOM
4AB5	1.0	0.1	63.00	6.4	0.7	0.21	MLE
4G1	8.9	0.5	3.13	10.4	0.9	1.21	MLE
4CB4	-488.0	0.0	573860.00	7.2	0.7	-0.86	MLE

TABLE 9(c); PARAMETERS OF THE LOG-PEARSON TYPE 3 DISTRIBUTION

	PARAMETERS			SAMPLE STATISTICS			
RGS. NO.	LOCATION y_o	SCALE α	SHAPE β	MEAN	STD:DEV.	SKEW	METHOD USED
4BB1	———	———	———	6.3	0.6	-0.00	MOM
4AA1	-164.9	0.0	49493.00	5.9	0.8	-0.01	MLE
5AA1	-175.9	0.0	60798.00	5.7	0.7	-0.06	MLE
5AA5	-306.9	0.0	125752.00	5.2	0.9	-0.50	MLE
5AB2	-507.8	0.0	217657.00	5.3	1.1	-0.25	MLE
5BC4	-231.5	0.0	96888.00	6.7	0.8	-0.08	MLE
5BC8	-3.0	0.2	37.44	6.1	1.5	0.15	MLE
5BE4	———	———	———	5.5	0.9	-0.23	MOM
5BE20	5.9	0.4	2.96	7.1	0.7	0.60	MLE

TABLE 9 (d): PARAMETERS OF THE FISHER TYPE DISTRIBUTIONS

	P A R A M E T E R S			SAMPLE STATISTICS			
RGS. NO.	LOCATION x_o	SCALE α	SHAPE β	MEAN	STD.DEV.	SKEW	METHOD USED
1CE1	3123.30	2339.54	-0.24	5073.1	3688.0	1.05	MLE
1CB2	78.24	183.53	-0.19	225.8	325,8	3.21	MOM
1BA1	244.02	276.88	-0.05	418.2	380.9	1.48	MOM
1BC1	565.38	169.61	0.40	613.3	163.6	-1.37	MLE
1BG4	86.63	60.48	-0.10	127.8	89.6	1.87	MOM
1BD1	305.65	325.45	0.00	492.9	416.4	1.13	MOM
1BE1	508.86	156.32	0.15	577.6	171.8	0.60	MLE
1BE2	13.42	11.03	-0.13	21.4	17.3	2.21	MOM
1BE5	188.81	65.81	-0.14	236.7	97.7	1.14	MLE
1BB1	512.08	283.35	-0.16	729.2	472.1	1.97	MLE
1GG1	-51.03	2575.33	-0.25	2269.5	5355.4	5.57	MOM
1JG1	5011.14	3047.51	-0.18	7377.4	4716.2	1.09	MLE
1KA5	10.05	6.81	-0.29	16.5	13.5	1.74	MLE
1LA3	954.38	316.67	0.50	1027.5	303.4	-0.17	MLE
1CB5	556.89	297.78	0.43	636.2	289.0	-0.17	MLE
1CB3	78.37	50.13	-0.38	136.9	140.9	3.84	MLE
1CD1	96.71	81.18	-0.32	176.7	166.9	2.27	MLE
1CA2	748.19	474.88	0.36	894.5	472.3	-0.16	MLE
1CB1	401.90	547.02	-0.08	763.8	787.8	1.72	MLE
1DA2	5313.94	2472.11	-0.16	7207.6	4289.8	2.37	MOM
1FG1	2701.95	1036.92	0:03	3274.8	1287.8	0.75	MLE
1JC13	8.02	15.55	-0.24	21.7	31.3	4.96	MOM
1KC3	9035.05	6034.27	-0.28	14574.0	11049.1	1.64	MLE
1FG2	3062.64	1220.42	-0.34	4263.5	2298.2	1.49	MLE

TABLE 9(d): PARAMETERS OF THE FISHER TIPPET TYPE DISTRIBUTION

RGS. NO.	PARAMETERS LOCATION x_o	SCALE α	SHAPE β	SAMPLE STATISTICS MEAN	STD:DEV.	SKEW	METHOD USED
1HA4	520.84	287.90	-0.34	847.2	914.8	5.07	MLE
1FF2	203.92	148.93	-0.06	298.7	207.0	1.53	MLE
1FE2	2378.00	1477.20	-0.29	3744.9	2676.9	1.50	MLE
1FE1	2127.46	772.15	-0.33	2378.9	784.5	0.00	MLE
2B2	106.68	26.92	-0.02	122.7	35.9	1.16	MLE
2EB3	84.24	105.07	-0.21	171.5	195.1	3.72	MOM
2EC2	19.20	19.82	0.02	30.2	24.7	1.01	MOM
2FA2	41.41	37.09	-0.19	70.2	55.6	0.92	MLE
2FC5	158.40	306.01	-0.03	343.6	407.5	1.32	MOM
2EC3	20.90	54.51	-0.20	66.0	100.7	3.66	MOM
2GD2	197.08	347.11	0.00	397.2	444.7	1.14	MOM
3BA10	1.50	236.34	-0.22	202.0	451.2	4.09	MOM
3BA17	11.14	11.75	-0.14	19.9	19.0	2.44	MOM
3BC12	1218.45	1632.94	0.05	2079.1	1963.4	0.85	MOM
3F2	10414.63	9848.03	-0.15	17834.3	16225.9	2.56	MOM
3DA2	8579.22	9361.91	0.08	13331.1	10989.4	0.75	MOM
3BA18	40.00	30.96	-0.37	71.6	60.6	1.79	MLE
3BB10	43.76	29.48	-0.09	63.7	40.6	0.71	MLE
4AA2	354.78	388.07	-0.06	604.3	544.4	1.58	MOM
4CA2	1206.76	857.63	-0.36	2102.9	1970.9	3.69	MLE
4BE1	2404.96	1700.22	-0.24	3857.0	2897.2	1.64	MLE
4AA4	219.31	328.86	-0.18	478.4	570.9	3.00	MOM
4BC2	4337.04	2079.47	-0.15	5885.8	3139.1	1.17	MLE

TABLE 9 (d): PARAMETERS OF THE FISHER TIPPET TYPE DISTRIBUTION

	PARAMETERS			SAMPLE STATISTICS			
RGS.NO.	LOCATION x_o	SCALE α	SHAPE β	MEAN	STD.DEV.	SKEW	METHOD USED
4AB5	488.14	300.34	-0.35	797.8	630.4	2.11	MLE
4G1	19093.38	42479.42	-0.18	52515.1	73639.6	2.99	MOM
4CB4	1320.52	825.90	0.39	1560.9	816.8	0.11	MLE
4BB1	459.39	233.27	-0.19	643.4	360.2	1.08	MLE
4AA1	325.74	294.15	0.04	483.3	357.6	0.90	MOM
5AA1	244.15	164.14	-0.26	396.2	352.0	3.27	MLE
5AA5	145.19	117.97	-0.21	238.9	173.7	0.63	MLE
5AB2	211.40	257.10	0.07	342.1	302.1	0.75	MOM
5BC4	735.25	661.74	0.01	1112.6	840.9	1.10	MOM
5BC8	375.58	1064.37	-0.14	1165.8	1725.0	2.44	MOM
5BE4	197.86	229.35	-0.10	354.7	341.0	1.89	MOM
5BE20	951.78	817.93	-0.03	1457.1	1089.8	1.32	MOM
5D5	1783.27	1481.82	-0.05	2717.1	2042.0	1.49	MOM

TABLE 9(e) :PARAMETERS OF THE LOG-FISHER TIPPET TYPE DISTRIBUTIONS

RGS.NO.	PARAMETERS LOCATION y_o	SCALE α	SHAPE β	SAMPLE STATISTICS MEAN	STD.DEV.	SKEW	METHOD USED
1CE1	8.04	0.89	0.50	8.2	0.8	-0.57	MLE
1CB2	4.42	0.85	0.08	4.8	1.0	0.65	MLE
1BA1	5.37	0.88	0.32	5.7	0.9	0.08	MLE
1BC1	6.32	0.33	0.60	6.4	0.3	-0.91	MOM
1BG4	4.40	0.40	-0.10	4.7	0.6	0.95	MLE
1BD1	5.56	1.07	0.51	5.8	1.0	-0.78	MLE
1BE1	6.24	0.34	0.51	6.3	0.3	-0.67	MOM
1BE2	2.55	0.48	-0.02	2.8	0.6	0.67	MLE
1BE5	5.24	0.36	0.22	5.4	0.4	0.28	MLE
1BB1	6.21	0.62	0.31	6.4	0.6	-0.18	MLE
1GG1	6.39	1.25	0.23	6.9	1.3	0.05	MLE
1JG1	8.54	0.74	0.50	8.7	0.7	-0.63	MLE
1KA5	2.30	0.83	0.39	2.5	0.8	-0.51	MLE
1LA3	6.81	0.33	0.50	6.9	0.3	-0.48	MLE
1CB5	6.16	0.63	0.51	6.3	0.6	-1.39	MLE
1CB3	4.34	0.69	0.18	4.6	0.7	0.39	MLE
1CB1	4.55	1.12	0.59	4.8	1.0	-0.86	MOM
1BA2	6.58	0.86	1.05	6.6	0.9	-2.16	MOM
1CB1	5.82	0.81	0.14	6.2	0.9	0.43	MLE
1DA2	8.56	0.50	0.23	8.7	0.5	0.17	MLE
1FG1	7.89	0.41	0.37	8.0	0.4	-0.24	MLE
1JC13	2.47	0.53	0.02	2.8	0.7	1.71	MLE
1KC3	9.10	0.76	0.36	9.3	0.7	-0.17	MLE

TABLE 9(e):PARAMETERS OF THE LOG-FISHER TIPPET TYPE DISTRIBUTIONS

	PARAMETERS			SAMPLE STATISTICS			
RGS. NO.	LOCATION y_o	SCALE α	SHAPE β	MEAN	STD.DEV.	SKEW	METHOD USED
1FG2	8.04	0.41	0.10	8.2	0.5	0.54	MLE
1HA4	6.23	0.57	0.11	6.5	0.6	0.76	MLE
1FF2	5.24	0.59	0.18	5.5	0.7	0.29	MLE
1FE2	7.78	0.69	0.33	8.0	0.7	-0.03	MLE
1GE1	7.63	0.38	0.50	7.7	0.4	-0.63	MLE
2B2	4.67	0.26	0.21	4.8	0.3	0.32	MLE
2EB3	4.43	0.85	0.23	4.8	0.9	0.13	MLE
2EC2	2.80	0.99	0.49	3.0	0.9	-0.24	MLE
2FA2	3.53	1.20	0.51	3.8	1.1	-1.18	MLE
2FC5	4.44	1.76	0.50	4.8	1.7	-0.32	MLE
2EC3	3.14	1.03	0.21	3.6	1.1	0.29	MLE
2GD2	3.94	2.62	0.51	4.6	2.5	-1.23	MLE
3BA10	3.56	1.08	-0.01	4.2	1.4	0.70	MLE
3BA17	2.33	0.76	0.20	2.6	0.8	0.29	MLE
3BC12	6.81	1.22	0.50	7.1	1.2	-0.22	MLE
3F2	9.22	0.59	0:08	9.5	0.7	0.67	MLE
3DA2	8.81	1.09	0.50	9.1	1.0	-0.38	MLE
3BA18	3.68	0.88	0.38	3.9	0.9	-0.23	MLE
3BB10	3.75	0.78	0.50	3.9	0.7	-0.55	MLE
4AA2	5.76	0.62	0.06	6.1	0.8	0.67	MLE
4CA2	7.06	0.77	0.24	7.3	0.8	-0.00	MLE
4BE1	7.76	0.80	0.39	8.0	0.8	-0.30	MLE
4AA4	5.28	0.75	0.02	5.7	0.9	0.60	MLE

TABLE 9(e): PARAMETERS OF THE LOG-FISHER TIPPET TYPE DISTRIBUTIONS

RGS.NO.	PARAMETERS LOCATION y_o	SCALE α	SHAPE β	SAMPLE STATISTICS MEAN	STD:DEV.	SKEW	METHOD USED
4BC2	8.38	0.52	0.32	8.5	0.5	0.00	MLE
4AB5	6.18	0.66	0.23	6.4	0.7	0.21	MLE
4G1	9.99	0.62	-0.06	10.4	0.9	1.21	MLE
4CB4	7.01	0.69	0.51	7.2	0.7	-0.86	MLE
4BB1	6.14	0.55	0.34	6.3	0.6	-0.00	MLE
4AA1	5.69	0.81	0.44	5.9	0.8	-0.01	MLE
5AA1	5.44	0.74	0.26	5.7	0.7	-0.06	MLE
5AA5	4.93	0.94	0.50	5.2	0.9	-0.50	MLE
5AB2	5.06	1.17	0.50	5.3	1.1	-0.25	MLE
5BC4	6.51	0.79	0.40	6.7	0.8	-0.03	MLE
5BC8	5.56	1.45	0.29	6.1	1.5	0.15	MLE
5BE4	5.20	0.95	0.37	5.5	0.9	-0.23	MLE
5BE20	6.73	0.52	-0.03	7.1	0.7	0.60	MLE
5D5	7.38	0.59	0.11	7.7	0.7	0.42	MLE

TABLE 9(f):PARAMETERS OF THE WALTER BOUGHTON DISTRIBUTION

	PARAMETERS			SAMPLE STATISTICS			
RGS. NO.	LOCATION x_o	SCALE α	SHAPE β	MEAN	STD.DEV.	SKEW	METHOD USED
1CE1	-270762.6	-99.7	27340986.4	5073.1	3688.0	1.05	LSQ
1CB2	192.6	- 4.2	1171.5	225.5	325.8	3.21	MLE
1BA1	-1259.3	- 7.9	11751.7	418.2	380.9	1.48	LSQ
1BC1	-61700.2	532.1	33119753.7	613.3	163.6	-0.37	LSQ
1BG4	-73.7	-4.9	762.4	127.8	89.6	1.87	MLE
1BD1	-20000.0	-67.2	1363222.0	492.9	416.4	1.13	LSQ
1BE1	-46549.0	-364.4	1714211.1	577.6	171.8	0.60	LSQ
1BE2	-18159.0	-1430.9	26004156.1	21.4	17.3	2.21	LSQ
1BE5	-18099.9	-246.5	4508787.2	236.7	97.7	1.14	LSQ
1BB1	—	—	—	729.2	472.1	1.97	LSQ
1GG1	-3310.8	-5.5	22420.7	2269.5	5355.4	5.57	LSQ
1JG1	-103769.6	-33.1	3603245.9	7377.4	4716.2	1.09	LSQ
1KA5	-18700.5	-1894.3	35445297.8	16.5	13.2	1.74	LSQ
1LA3	-69390.0	-327.6	23021910.7	1027.5	303.4	-0.17	LSQ
1CB5	-82683.0	-396.0	32940553.3	636.2	289.0	-0.17	LSQ
1CB3	-240.8	-6.5	2095.1	136.9	140.9	3.84	MLE
1CD1	-531.2	-7.8	4921.1	176.7	166.9	2.27	MLE
1CA2	-113035.0	334.6	35054653.8	894.5	472.5	-0.16	LSQ
1CB1	-664.8	-4.1	4127.1	763.8	787.8	1.72	LSQ
1DA2	-19009.6	-10.6	257070.6	7207.6	4289.8	2.37	LSQ
1FG1	-187189.0	-195.2	37069947.0	3274.8	1287.8	0.75	LSQ
1JC13	-18.8	-5.8	176.8	21.7	31.3	4.96	LSQ

TABLE 9(f): PARAMETERS OF THE WALTER BOUGHTON DISTRIBUTION

	PARAMETERS			SAMPLE STATISTICS			
RGS.NO.	LOCATION x_o	SCALE α	SHAPE β	MEAN	STD.DEV.	SKEW	METHOD USED
1KC3	-53806.7	-10.2	645605.3	14574.0	11049.1	1.64	LSQ
1FG2	-18602.9	-15.1	329262.7	4263.5	2298.2	1.49	LSQ
1HA4	-1343.9	-7.2	13416.2	847.2	914.8	5.07	LSQ
1FF2	-2756.0	-21.5	63622.0	298.7	208.7	1.53	LSQ
1FE2	-17082.0	-12.2	238247.6	3744.9	2676.9	1.50	LSQ
1FE1	-145377.0	-257.0	37887803.9	2378.9	784.5	0.00	LSQ
2B2	-19800.0	-727.5	14481815.3	122.7	35.9	1.16	LSQ
2EB3	-301.7	-0.0	2328.1	171.5	195.1	3.72	MLE
2EC2	-19640.0	-1054.9	20737529.0	30.2	24.7	1.01	LSQ
2FA2	-19879.2	-414.3	9450564.4	70.2	55.6	0.92	LSQ
2FC5	-2267.0	-10.7	25877.9	343.6	407.5	1.32	LSQ
2EC3	-68.4	-4.4	409.4	66.0	100.7	3.66	MLE
2GD2	—	—	—	397.2	44.7	1.14	LSQ
3BA10	-239.9	-4.7	1353.5	202.0	451.2	4.09	LSQ
3BA17	-18630.0	2327.4	24744824.1	19.9	19.0	2.44	LSQ
3BC12	-158658.0	-115.0	18390300.0	2079.1	1963.4	0.85	LSQ
3F2	-28903.0	-6.4	250170.4	17834.3	16225.9	2.56	LSQ
3DA2	-743890.0	-94.2	70366389.4	13331.1	10939.4	0.75	LSQ
3BA18	-465.1	-13.9	7057.3	71.6	60.2	1.79	LSQ
3BB10	-19829.1	-629.1	12502844.4	63.7	40.6	0.71	LSQ
4AA2	-502.9	-4.5	3710.0	604.3	544.4	1.58	MLE
4CA2	-6824.7	-9.1	73709.6	2102.9	1970.9	3.69	LSQ
4BE1	-25527.2	-15.4	341945.9	3857.0	2897.2	1.64	LSQ

TABLE 9(f):PARAMETERS OF THE WALTER BOUGHTON DISTRIBUTION

	PARAMETERS			SAMPLE STATISTICS			
RGS.NO.	LOCATION x_o	SCALE α	SHAPE β	MEAN	STD.DEV.	SKEW	METHOD USED
4AA4	-1176.8	-6.7	9328.5	478.4	570.9	3.00	LSQ
4BC2	-85400.0	-40.1	3597125.3	5885.8	3139.1	1.17	LSQ
4AB5	-2395.8	-9.0	26158.1	797.8	630.4	2.11	LSQ
4G1	-36599.8	-4.5	265919.1	52515.1	73639.6	2.99	LSQ
4CB4	-147190.0	-250.2	37123456.1	8160.9	816.8	0.11	LSQ
4BB1	-35171.1	-132.2	4712931.7	643.4	360.2	1.08	LSQ
4AA1	-63672.1	-243.1	1555495.3	483.3	357.6	0.90	LSQ
5AA1	———	———	———	396.2	352.0	3.27	MLE
5AA5	-35143.6	260.5	9197085.0	238.9	173.7	0.63	LSQ
5AB2	-56061.0	-254.0	14292606.2	342.1	302.1	0.75	LSQ
5BC4	-113084.3	-184.8	1032679.5	1112.6	840.9	1.10	LSQ
5BC8	-3233.0	-6.0	21627.2	1165.8	1725.0	2.44	LSQ
5BE4	-639.7	-5.8	4794.7	354.7	341.0	1.89	MLE
5BE20	-3418.1	-8.0	34406.1	1457.1	1089.8	1.32	LSQ
5D5	-5994.5	-7.2	55257.1	2717.1	2042.0	1.49	LSQ

TABLE 9(g): PARAMETERS OF THE LOG-WALTER BOUGHTON DISTRIBUTION

RGS.NO:	PARAMETERS LOCATION y_o	SCALE α	SHAPE β	SAMPLE STATISTICS MEAN	STD.DEV.	SKEW	METHOD USED
1HA4	———	———	———	6.5	0.6	0.76	LSQ
1FF2	———	———	———	5.5	0.7	0.29	LSQ
1FE2	———	———	———	8.0	0.7	-0.03	MLE
1FE1	———	———	———	7.7	0.4	-0.63	MLE
2B2	———	———	———	4.8	0.3	0.32	MLE
2EB3	———	———	———	4.8	0.9	0.13	LSQ
2EC2	———	———	———	3.0	0.9	-0.24	MLE
2FA2	———	———	———	3.8	1.1	-1.18	MLE
2FC5	———	———	———	4.8	1.7	-0.32	MLE
2EC3	———	———	———	3.6	1.1	0.29	MLE
2GD2	———	———	———	4.6	2.5	-1.23	MLE
3BA10	———	———	———	4.2	1.4	0.70	LSQ
3BA17	———	———	———	2.6	0.8	0.29	MLE
3BC12	———	———	———	7.1	1.2	-0.22	MLE
3F2	———	———	———	9.5	0.7	0.67	MLE
3DA2	———	———	———	9.1	1.0	-0.38	MLE
3BA18	———	———	———	3.9	0.9	-0.23	MLE
3BB10	———	———	———	3.9	0.7	-0.55	MLE
4AA2	———	———	———	6.1	0.8	0.67	MLE
4CA2	———	———	———	7.3	0.8	-0.00	MLE
4BE1	———	———	———	8.0	0.8	-0.30	MLE
4AA4	———	———	———	5.7	0.9	0.60	LSQ

TABLE 9(g):PARAMETERS OF THE LOG-WALTER BOUGHTON DISTRIBUTION

	PARAMETERS			SAMPLE STATISTICS			
RGS. NO.	LOCATION x_o	SCALE α	SHAPE β	MEAN	STD.DEV.	SKEW	METHOD USED
1CE1	—	—	—	8.2	0.8	-0.57	MLE
1CB2	—	—	—	4.8	1.0	0.65	MLE
1BA1	—	—	—	5.7	0.9	0.08	MLE
1BC1	—	—	—	6.4	0.3	-0.91	MLE
1BG4	—	—	—	4.7	0.6	0.95	LSQ
1BD1	—	—	—	5.8	1.0	-0.78	MLE
1BE1	—	—	—	6.3	0.3	-0.67	LSQ
1BE2	—	—	—	2.8	0.6	0.87	LSQ
1BE5	—	—	—	5.4	0.4	0.28	MLE
1BB1	—	—	—	6.4	0.6	-0.18	LSQ
1GG1	—	—	—	6.9	1.3	0.05	LSQ
1JG1	—	—	—	8.7	0.7	-0.63	MLE
1KA5	—	—	—	2.5	0.8	-0.51	MLE
1LA3	—	—	—	6.9	0.3	-0.48	MLE
1CB5	—	—	—	6.3	0.6	-1.39	MLE
1CB3	—	—	—	4.6	0.7	0.39	LSQ
1CD1	—	—	—	4.8	1.0	-0.86	MLE
1CA2	—	—	—	6.6	0.9	-2.16	MLE
1CB1	—	—	—	6.2	0.9	0.43	LSQ
1DA2	—	—	—	8.7	0.5	0.17	LSQ
1FG1	—	—	—	8.0	0.4	-0.24	MLE
1JC13	—	—	—	2.8	0.7	1.71	LSQ
1KC3	—	—	—	9.3	0.7	-0.17	MLE
1FG2	—	—	—	8.2	0.5	0.54	MLE

TABLE 9(g) PARAMETERS OF THE LOG-WALTER BOUGHTON DISTRIBUTION

	PARAMETERS			SAMPLE STATISTICS			
RGS. NO.	LOCATION x_0	SCALE α	SHAPE β	MEAN	STD.DEV.	SKEW	METHOD USED
4BC2	—	—	—	8.5	0.5	0.00	LSQ
4AB5	—	—	—	6.4	0.7	0.21	MLE
4G1	—	—	—	10.4	0.9	1.21	LSQ
4CB4	—	—	—	7.2	0.7	-0.86	MLE
4BB1	—	—	—	6.3	0.6	-0.00	MLE
4AA1	—	—	—	5.9	0.8	0.01	MLE
5AA1	—	—	—	5.7	0.7	-0.06	MLE
5AA5	—	—	—	5.2	0.9	-0.50	MLE
5AB2	—	—	—	5.3	1.1	-0.25	MLE
5BC4	—	—	—	6.7	0.8	-0.03	MLE
5BC8	—	—	—	6.1	1.5	0.15	LSQ
5BE4	—	—	—	5.5	0.9	-0.23	MLE
5BE20	—	—	—	7.1	0.7	0.60	LSQ
5D5	—	—	—	7.7	0.7	0.42	MLE

TABLE 9(h):PARAMETERS OF THE WAKEBY DISTRIBUTION

	PARAMETERS					SAMPLE STATISTICS			
RGS.NO.	A	B	C	D	E	MEAN	STD.DEV.	SKEW	METHOD USED
1CE1	25757.90	0.211	0.00	0.000	26349.64	8.24	0.70	-0.575	PWM
1CB2	0.00	0.000	261.22	0.435	-236.95	4.85	1.00	0.649	"
1BA1	0.00	0.000	3483.77	0.096	-3437.23	5.66	0.80	0.080	"
1BC1	114.39	2.991	346.47	0.153	307.90	6.38	0.09	-0.908	"
1BG4	0.00	0.003	299.93	0.207	-250.41	4.68	0.32	0.947	"
1BD1	8405.29	0.058	0.00	0.000	8434.09	5.80	1.04	-0.777	"
1BE1	349.29	7.733	945.32	0.114	-449.00	6.31	0.10	-0.667	"
1BE2	0.00	0.000	42.56	0.252	-35.54	2.84	0.40	0.868	"
1BE5	671.45	0.218	0.00	0.000	787.85	5.39	0.15	0.279	"
1BB1	352.11	7.304	1076.40	0.249	-661.89	6.42	0.37	-0.176	"
1GG1	0.00	0.000	1794.74	0.545	-1675.34	6.86	1.60	0.050	"
1JG1	45358.46	0.138	0.00	0.000	47245.32	8.70	0.47	-0.628	"
1KA5	0.00	0.000	172.30	0.069	-168.55	2.53	0.62	-0.506	"
1LA3	1102.38	1.028	0.00	0.000	1571.12	6.89	0.10	-0.478	"
1CB5	1019.81	1.010	0.00	0.000	1143.57	6.61	0.38	-1.392	"
1CB3	35.10	5.692	146.92	0.375	-93.00	4.62	0.55	0.395	"
1CD1	57.39	4.425	409.68	0.239	-350.93	4.76	1.05	-0.861	"
1CA2	1617.74	1.088	0.00	0.000	1669.35	6.56	0.78	-2.165	"
1CB1	0.00	0.000	2219.53	0.235	-2138.71	6.20	0.86	0.431	"
1DA2	2936.81	4.328	7189.90	0.289	-2357.73	8.75	0.26	0.171	"
1FG1	5864.83	0.440	0.00	0.000	7348.39	8.02	0.16	-0.238	"
1JC13	9.31	4.123	4.82	0.692	7.89	2.75	0.46	1.707	
1KC3	4041.07	12.128	101850.32	0.092	-97323.36	9.33	0.55	-0.165	"
1FG2	0.00	0.000	994643.56	0.002	992824.79	8.24	0.23	0.542	"
1HA4	0.00	0.000	2920.75	0.167	-2658.24	6.49	0.41	0.764	"

TABLE 9(h): PARAMETERS OF THE WAKEBY DISTRIBUTION

RGS. NO.	PARAMETERS A	B	C	D	E	SAMPLE STATISTICS MEAN	STD. DEV.	SKEW	METHOD USED
1FF2	5234.89	0.045	0.00	5308.70	5308.70	5.49	0.43	0.290	PWM
1FE2	2599.43	65.532	42748.41	0.059	-41633.98	8.00	0.48	-0.026	"
1FE1	2894.01	0.892	0.00	0.000	3908.85	7.72	0.10	-0.633	"
2B2	222.80	0.252	0.00	0.000	300.74	4.77	0.08	0.317	"
2EB3	69.07	1.774	154.55	0.417	-68.63	4.76	0.77	0.132	"
2EC2	224.50	0.148	0.00	0.000	225.78	3.03	0.87	-0.235	"
2FA2	356.80	0.238	0.00	0.000	358.42	3.81	1.32	-1.184	"
2FC5	0.00	0.000	2428.64	0.138	-2474.53	4.85	2.82	-0.318	"
2EC3	0.00	0.000	76.52	0.442	-71.02	3.56	1.19	0.291	"
2GD2	0.00	0.000	71376.49	0.007	-71448.39	4.57	6.41	-1.227	"
3BA10	0.00	0.000	126.02	0.617	-127.37	4.20	1.91	0.700	"
3BA17	0.00	0.000	75.38	0.179	-71.97	2.65	0.68	0.291	"
3BC12	26121.43	0.092	0.00	0.000	25991.90	7.09	1.33	-0.216	"
3F2	4428.87	27.923	30975.07	0.287	-25470.79	9.52	0.50	0.670	"
3DA2	62873.59	0.284	0.00	0.000	62299.20	9.06	1.09	-0.383	"
3BA18	69.81	0.764	94.34	0.269	-17.79	3.94	0.74	-0.226	"
3BB10	208.69	0.344	0.00	0.000	219.01	3.93	0.55	-0.554	"
4AA2	0.00	0.000	1652.20	0.222	-1520.02	6.09	0.58	0.668	"
4CA2	0.00	0.000	523224.69	0.003	-522864.51	7.35	0.62	-0.003	"
4BE1	2024.38	2.053	8891.18	0.193	-6496.67	7.99	0.60	-0.297	"
4AA4	0.00	0.000	905.24	0.319	-850.47	5.70	0.87	0.64	"
4BC2	23268.66	0.194	0.00	0.000	25376.02	8.55	0.27	0.001	"
4AB5	1785.95	210.767	2473.08	0.181	-2215.20	6.44	0.48	0.214	"

TABLE 9 (h) PARAMETERS OF THE WAKEBY DISTRIBUTION

RGS.NO	PARAMETERS A	B	C	D	E	SAMPLE STATISTICS MEAN	STD.DEV.	SKEW	METHOD USED
4G1	0.00	0.00	41048.58	0.0.507	-30734.07	10.39	0.75	1.207	PWM
4CB4	2985.75	0.804	0.00	0.000	3215.90	7.18	0.44	-0.865	"
4BB1	2738.40	0.188	0.00	0.000	2947.97	6.32	0.31	-0.002	"
4AA1	3262.77	0.147	0.00	0.000	3326.72	6.90	0.61	-0.009	"
5AA1	268.08	2.449	176.68	0.493	125.75	5.71	0.56	-0.064	"
5AA5	871.70	0.360	0.00	0.000	879.93	5.15	0.81	-0.501	"
5AB2	2161.49	0.205	0.00	0.000	2135.75	5.33	1.27	-0.245	"
5BC4	11227.36	0.091	0.00	0.000	11399.14	6.74	0.61	-0.030	"
5BC8	0.00	0.000	1915.39	0.394	-1997.09	6.07	2.29	0.146	"
5BE4	0.00	0.000	1753.01	0.149	-1706.21	5.48	0.85	-0.226	"
5BE20	0.00	0.000	8598.76	0.111	-8215.13	7.05	0.26	0.601	"
5D5	0.00	0.000	33752.75	0.058	-33101.58	7.67	0.47	0.421	"

TABLE 9(i):PARAMETERS OF THE LOG-WAKEBY DISTRIBUTION

	P A R A M E T E R S					SAMPLE STATISTICS			
RGS.NO.	A.	B	C	D	E	MEAN	STD.DEV.	SKEW	METHOD USED
1CE1	3.14	1.270	0.00	0.000	9.62	8.24	0.70	-0.575	PWM
1CB2	4.02	0.00	0.00	0.000	7.33	4.85	1.00	0.649	"
1BA1	3.49	0.819	0.00	0.000	7.58	5.66	0.80	0.080	"
1BC1	1.40	9.115	0.80	0.189	5.53	6.38	0.09	-0.908	"
1BG4	3.01	0.375	0.00	0.000	6.86	4.68	0.32	0.947	"
1BD1	3.71	1.109	0.00	0.000	7.49	5.80	1.04	-0.777	"
1BE1	1.31	1.767	0.00	0.000	6.79	6.31	0.10	-0.667	"
1BE2	3.37	0.347	0.00	0.000	5.34	2.84	0.40	0.868	"
1BE5	1.70	0.770	0.00	0.000	6.35	5.39	0.15	0.279	"
1BB1	2.00	10.624	2.04	0.176	4.11	6.42	0.37	-0.176	"
1GG1	2.99	4.837	5.20	0.122	1.45	6.86	1.60	0.050	"
1JG1	2.65	1.016	0.00	0.000	10.02	8.70	0.47	-0.628	"
1KA5	2.64	0.986	0.00	0.000	3.86	2.53	0.62	-0.506	"
1LA3	2.17	14.639	1.18	0.203	5.55	6.89	0.10	-0.478	"
1CB5	2.75	9.708	4.70	0.072	1.50	6.31	0.38	-1.392	"
1CB3	2.83	0.744	0.00	0.000	6.25	4.62	0.55	0.395	"
1CD1	3.34	1.453	0.00	0.000	6.12	4.76	1.05	-0.861	"
1CA2	3.70	8.378	3.69	0.090	2.90	6.56	0.78	-2.165	"
1CB1	4.00	0.610	0.00	0.000	8.69	6.20	0.86	0.431	"
1DA2	1.97	12.187	1.58	0.208	6.90	8.75	0.26	0.171	"
1FG1	1.65	13.642	4.71	0.073	3.05	8.02	0.16	-0.238	"

TABLE 9(i):PARAMETERS OF THE LOG-WAKEBY DISTRIBUTION

RGS.NO.	PARAMETERS					SAMPLE STATISTICS			METHOD USED
	A	B	C	D	E	MEAN	STD.DEV	SKEW	
1JC13	1.29	4.683	0.38	0.457	2.28	2.75	0.46	1.707	PWM
1KC3	2.55	11.776	9.46	0.063	-0.56	9.33	0.55	-0.165	"
1FG2	3.02	32.237	3.22	0.140	4.58	8.24	0.23	0.542	"
1HA4	2.22	1.130	0.00	0.000	7.54	6.49	0.41	0.764	"
1FF2	2.85	0.707	0.00	0.000	7.16	5.49	0.43	0.290	"
1FE2	2.73	13.524	3.90	0.137	3.67	8.00	0.48	-0.026	"
1FE1	2.46	15.750	1.24	0.206	6.30	7.72	0.13	-0.633	"
2B2	1.55	31.395	3.71	0.071	0.82	4.77	0.08	0.317	"
2EB3	3.21	0.903	0.00	0.000	6.45	4.76	0.77	0.132	"
2EC2	3.23	1.110	0.00	0.000	4.56	3.03	0.87	-0.235	"
2FA2	3.78	1.890	0.00	0.000	5.12	3.81	1.32	-1.184	"
2FC5	0.05	1.585	0.20	0.481	6.43	4.85	2.82	-0.318	"
2EC3	4.00	0.771	0.00	0.000	5.82	3.56	1.19	0.291	"
2GD2	8.48	2.366	0.00	0.000	7.09	4.57	6.41	-1.227	"
3BA10	2.04	1.383	68.25	0.013	-64.08	4.20	1.91	0.700	"
3BA17	3.14	0.711	0.00	0.000	4.48	2.65	0.68	0.291	"
3BC12	4.17	1.134	0.00	0.000	9.05	7.09	1.33	-0.216	"
3F2	2.16	14.456	3.79	0.150	5.20	9.52	0.50	0.670	"
3DA2	4.01	1.292	0.00	0.000	10.81	9.06	1.09	-0.383	"
3BA18	3.07	1.111	0.00	0.000	5.40	3.94	0.74	-0.226	"
3BB10	2.62	1.309	0.00	0.000	5.06	3.93	0.55	-0.554	"

TABLE 9(i): PARAMETERS OF THE LOG-WAKEBY DISTRIBUTION

RGS.NO.	PARAMETERS A	B	C	D	E	SAMPLE STATISTICS MEAN	STD.DEV.	SKEW	METHOD USED
4AA2	3.63	0.471	0.00	0.000	8.56	6.09	0.58	0.668	PWM
4CA2	2.86	1.085	0.00	0.000	8.72	7.35	0.62	-0.003	"
4BE1	2.18	7.195	9.47	0.057	-1.78	7.99	0.60	-0.297	"
4AA4	4.28	0.496	0.00	0.000	8.56	5.70	0.87	0.604	"
4BC2	2.17	16.718	8.09	0.060	0.06	8.55	0.27	0.001	"
4AB5	1.83	11.919	16.01	0.038	-10.06	6.44	0.48	0.214	"
4G1	1.60	11.061	2.88	0.204	6.90	10.39	0.75	1.027	"
4CB4	2.47	1.751	0.00	0.000	8.07	7.18	0.44	-0.865	"
4BB1	2.24	0.889	0.00	0.000	7.71	6.32	0.31	-0.002	"
4AA1	2.98	0.909	0.000	0.000	7.46	5.90	0.61	-0.009	"
5AA1	1.99	5.680	1.17	0.256	4.43	5.71	0.56	-0.064	"
5AA5	3.28	1.361	0.00	0.000	6.54	5.15	0.81	-0.501	"
5AB2	4.14	1.141	0.00	0.000	5.27	5.33	1.27	-0.245	"
5BC4	3.24	0.855	0.00	0.000	8.48	6.74	0.61	-0.030	"
5BC8	5.89	0.762	0.00	0.000	9.41	6.07	2.29	0.146	"
5BE4	3.38	0.941	0.00	0.000	7.23	5.48	0.85	-0.226	"
5BE20	3.53	0.485	0.00	0.000	9.43	7.05	0.46	0.601	"
5D5	3.17	0.650	0.00	0.000	9.59	7.67	0.47	0.421	"

the method of moments is denoted by MOM, the method of least squares is denoted by LS and the method of probability weighted moments by PWM.

Except for the Walter Boughton and the Pearson Type 3 distributions, all the other flood probability distributions which are under investigation and which have maximum likelihood parameter estimation routines gave maximum likelihood parameter estimates in most of the catchments. It was noticed that the maximum likelihood method cannot work with the Walter Boughton distribuiton when the estimate of β in (138) is at least equal to 20000. This is mainly due to the creation of an arithmetic overflow in the maximum likelihood equations.

The most serious problem with the moment estimates of the Pearson and the log-Pearson Type 3 distributions was that the upper or lower bounds were most frequently within the range of the actual observations. This is certainly not permissible and therefore in such cases, no further analysis could be done. Similarly, whenever the final estimate of the shape parameter β in (54) or (78) as obtained by either the method of moments of the method of maximum likelihood was greater than 57, no goodness of fit tests were carried out. This was mainly because the ICL/2950 Fortran subroutine F4GAMMA, which was used to compute $\Gamma(\beta)$ in (55) could not be used with $\beta > 57$.

The existence of a lower or upper bound in a fitted distribution usually depends on the sample skewness coefficient. Figure (8) shows a histogram of the sample skewness coefficients for both the transformed and untransformed flood data for the catchments used in this study. Individual values range from 5.56 to -0.37 with a mean values of 1.77 for the untransformed data and from 1.66 to -2.18 with a mean value of -0.07 for the log-transformed data. In general, 55% of the selected catchments had negative coefficients of skewness for the log-transformed data and only about 7% of the selected catchments had skewness coefficients which were at least equal to 1.14 for the untransformed data while only about 3% had sample skewness coefficients which were at least equal to 1.14 for the log-transformed data. In general, this implies that if all the chosen distributions' parameters were obtainable, then,

(i) about 7% of the fitted three parameter log-normal distributions would have upper bounds.

(ii) about 62% of the fitted Fisher Tippet distributions would be of type 2 with lower bounds and the rest 38% would be of type 3 with upper bounds.

(iii) about 3% of the fitted log-Fisher Tippet distributions would be of Type 2 with lower bounds and the rest 97% of Type 3 with upper bounds.

(iv) about 7% of the fitted Pearson Type 3 dis-

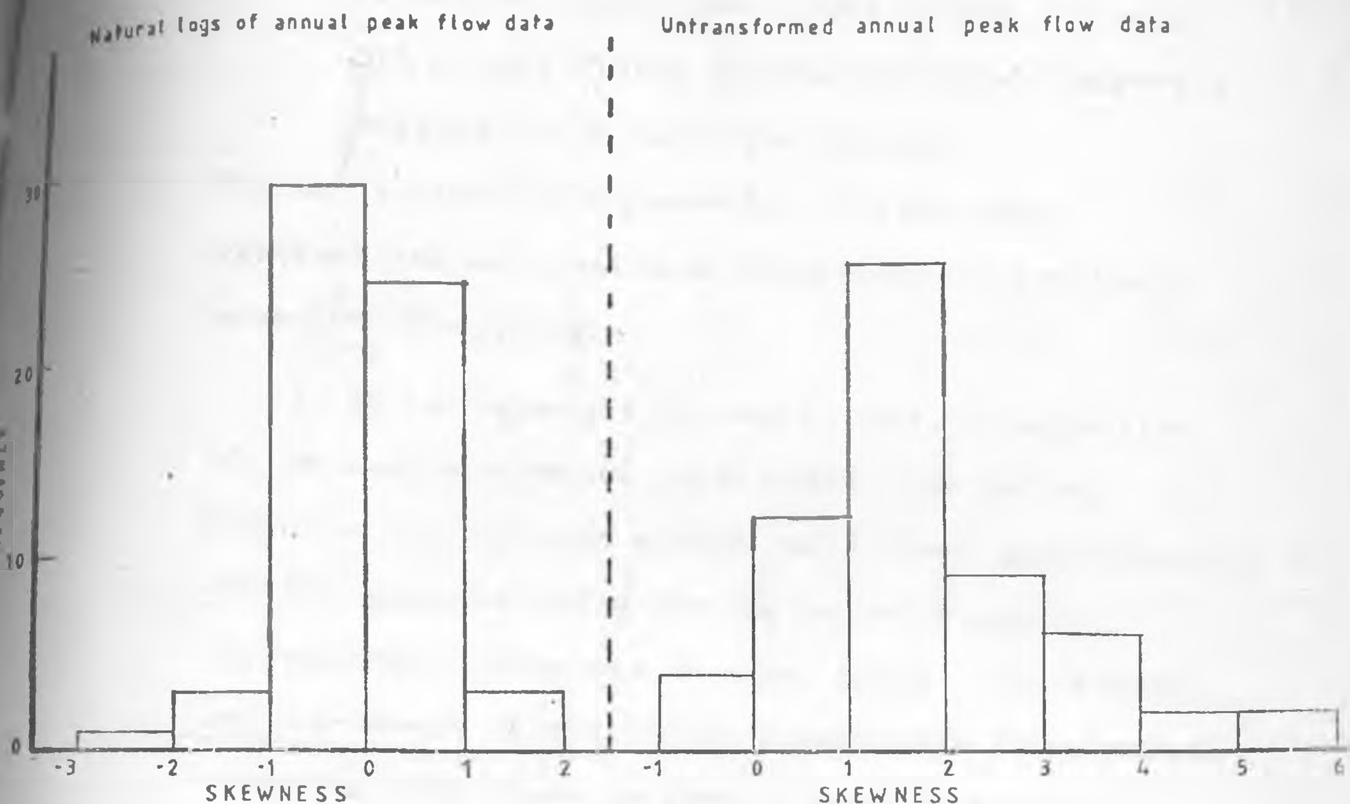

Fig. 8 Histograms of the skewness coefficients of the annual peak flow data

tributions would have upper bounds and about 55% of the fitted log-Pearson Type 3 distributions would have upper bounds.

The conclusions were generally in line with observations although some distribution parameters were not obtainable.

It is important to recall that, irrepsective of the sample skewness coefficient, the Walter Boughton distribution always has a lower bound which is usually negative while the log-Walter Boughton distribution always has an upper bound. The Wakeby and log-Wakeby distributions always have lower bounds which are very close to zero in both domains.

The existence of an upper bound in a probability distribution is unrealistic in flood frequency analysis. For this reason therefore, only those cases of the catchments for which a given probability distribution gives only lower bounds which are smaller than the smallest observed value in the catchment are considered.

The flood probability distributions in the real domain normally require that the lower bound when it exists be positive. However, for the negative lower bound flood probability distributions, corresponding truncated probability distributions were used. This requires that, a finite probability, p_o, given by

$$p_o = \int_{x_o}^{o} f(x)dx \quad .... (238)$$

be used for the value x = O, where x_o is the negative lower bound and f(x) is the negative lower bounded probability density function of the variate x. In this way, all the negative values of the variate x in f(x) are assigned zero probability.

Lastly, it is noteworthy to mention that the methodologies for estimating the parameters of all the other distributions except the Wakeby and log-Wakeby distributions were generally straight-forward and relatively less complicated to program for the computer. It was found that the least squares approach which was developed for the estimation of the parameters of the Wakeby and Log-Wakeby distributions did not give an overall least sum of squared error as is required in all the chosen catchments. During the search for the optimum location parameter e in equation (160) it was noticed that the sum of squared error decreased fairly fast for values of e close to zero and then continued to decrease only but very slowly with further decrements or increments in e until e became so large in magnitude that it virtually overdominated the other components in equation (160), hence making it impossible to continue with the required regression analysis. Due to this problem, it was found necessary to adopt the method of probability weighted moments for the estimation of the prameters of the Wakeby and Log-Wakeby distributions. A plotting position formula of the form,

$$p(x_i) = (i-0.35)/n \qquad \ldots . \ (239).$$

was found to be most appropriate for the data chosen, where $p(x_i)$ is the probability of non-exceedance of the i-th smallest observation x_i and n is the total number of observations.

In the following subsection, we will discuss the results from the goodness of fit tests.

3.2 RESULTS OF GOODNESS OF FIT TESTS

The Smirnov-Kolmogorov distance goodness of fit measure, Δ, was computed directly from equation (194) after the computation of distribution parameters in each catchment. The observed cumulative probability p(x) was estimated by the plotting position formular given by equation (239). The distribution of the computed Δ is presented in Table (10). In this table and the others to follow, the distributions are given the notation as follows:

LN3 - three prameter log-normal
P3 - Pearson type 3
LP3 - log-Pearson type 3
FT - Fisher-Tippet
LFT - log-Fisher-Tippet
WB - Walter-Bougnton
LWB - Log-Walter-Bougnton

TABLE 10: GOODNESS-OF-FIT BY THE SMIRNOV--KOLMOGOROV TEST

	DISTRIBUTION								
RGS	LN3	P3	LP3	FT	LFT	WB	LWB	W	LW
1CE1	0.082A	0.077A	-	0.092A	0.083A	0.119A	-	0.066A*	0.120A
1CB2	0.094A	-	0.091A	0.241R	0.088A*	0.107A	-	0.093A	0.131A
1BA1	0.075A*	-	-	0.167A	0.104A	0.128A	-	0.096A	0.119A
BC1	-	-	-	0.078A	0.098A	0.170A	-	0.052A*	0.122A
BG4	0.059A	0.161A	0.058A*	0.136A	0.067A	0.073A	-	0.059A	0.084A
BD1	0.106A	0.137A	-	0.151A	0.100A	0.146A	-	0.089A*	0.129A
BE1	0.105A	0.104A	-	0.110A	0.990R	0.153A	-	0.100A*	0.229R
BE2	0.053A	0.110A	0.052A*	0.125A	0.054A	0.151A	-	0.076A	0.098A
BE5	0.049A*	0.053A	0.051A	0.053A	0.058A	0.068A	-	0.070A	0.124A
1BB1	0.115A	0.119A	-	0.115A	0.123A	-	-	0.105A	0.092A*
1GG1	0.109A	-	-	0.365R	0.117A	0.115A	-	0.113A	0.095A*
1JG1	0.111A	0.121A	-	0.106A	0.123A	0.127A	-	0.077A*	0.145A
1KA5	0.105A	0.135A	-	0.086A*	0.128A	0.156A	-	0.095A	0.139A
1LA3	-	-	-	0.145A	0.136A	0.210A	-	0.114A*	0.169A
1CB5	-	-	-	0.066A	0.098A	0.124A	-	0.064A*	0.088A
1CB3	0.066A	-	-	0.058A	0.063A	0.055A*	-	0.062A	0.105A
1CB1	0.064A*	-	-	0.067A	0.993R	0.074A	-	0.068A	0.150A
1CA2	-	-	-	0.123A	0.985R	0.096A*	-	0.107A	0.153A
1CB1	0.070A	0.198A	0.065A*	0.171A	0.073A	0.087A	-	0.091A	0.083A
1DA2	0.077A	0.008A	-	0.071A	0.090A	0.062A*	-	0.068A	0.127A

TABLE 10: GOODNESS OF-FIT BY THE SMIRNOV-KOLMOGOROV TEST

	DISTRIBUTION								
RGS	LN3	P3	LP3	FT	LFT	WB	LWB	W	LW
1FG1	0.060A	0.056A	-	0.063A	0.056A*	0.074A	-	0.061A	0.095A
1JC13	0.138A	-	0.129A	0.293R	0.125A	0.134A	-	0.070A*	0.078A
1KC3	0.045A	0.072A	-	0.040A	0.060A	0.040A*	-	0.042A	0.074A
1FG2	0.115A	0.111A	0.114A	0.113A	0.123A	0.142A	-	0.100A	0.093A
1HA4	0.124A*	-	0.125A	0.124A	0.125A	0.146A	-	0.125A	0.183A
1FF2	0.111A	-	0.120A	0.145A	0.128A	0.142A	-	0.092A*	0.102A
1FE2	0.062A	0.074A	-	0.054A*	0.077A	0.064A	-	0.057A	0.133A
1FE1	-	-	-	0.095A	0.087A*	0.158A	-	0.111A	0.147A
2B2	0.135A	0.142A	-	0.132A	0.132A	0.127A	-	0.126A	0.084A*
2EB3	0.044A*	-	-	0.129A	0.053A	0.064A	-	0.052A	0.103A
2EC2	0.095A	-	-	0.134A	0.096A	0.136A	-	0.080A*	0.082A
2FA2	0.108A	0.088A	0.123A	0.089A	0.159A	0.076A*	-	0.120A	0.149A
2FC5	0.176A	0.183A	-	0.204A	0.140A	0.189A	-	0.133A	0.116A*
2EC3	0.056A	-	-	0.230R	0.049A*	0.074A	-	0.054A	0.084A
2GD2	0.169A	-	-	0.155A	0.131A	-	-	0.126A	0.086A*
3BA10	0.065A	0.49R	0.070A	0.362R	0.087A	0.36A	-	0.112A	0.058A*
3BA17	0.074A	-	0.070A*	0.119A	0.078A	0.154A	-	0.080A	0.113A
3BC12	0.116A	0.195A	-	0.198A	0.113A*	0.187A	-	0.127A	0.167A
3F2	0.071A*	-	0.081A	0.124A	0.078A	0.078A	-	0.076A	0.110A
3DA2	0.160A	-	-	0.131A	0.150A	0.162A	-	0.124A*	0.154A
3BA18	0.073A	0.080A	-	0.078A	0.065A	0.097A	-	0.064A*	0.105A
3BB10	0.090A	0.103A	-	0.097A	0.088A	0.101A	-	0.074A*	0.097A
4AA2	0.138A	0.269R	0.122A	0.237R	0.115A	0.096A*	-	0.143A	0.135A

TABLE 10: GOODNESS OF FIT BY THE SMIRNOV-KOLMOGOROV TEST

	DISTRIBUTION								
RGS	LN3	P3	LP3	FT	LFT	WB	LWB	W	LW
4CA2	0.063A	-	-	0.073A	0.071A	0.078A	-	0.058A*	0.116A
4BE1	0.081A	0.061A	-	0.077A	0.067A	0.064A	-	0.053A*	0.082A
4AA4	0.072A*	-	0.080A	0.177A	0.088A	0.133A	-	0.087A	0.076A
4BC2	0.071A	0.071A	-	0.067A*	0.076A	0.083A	-	0.068A	0.101A
4AB5	0.090A	-	-	0.074A	0.083A	0.078A	-	0.123A	0.043A*
4G1	0.094A	0.388R	0.073A	0.262R	0.060A*	0.066A	0.068A	0.081A	0.084A
4CB4	0.118A	0.124A	-	0.119A	0.109A	0.160A	-	0.065A*	0.150A
4BB1	0.071A	0.065A	-	0.070A	0.076A	0.091A	-	0.055A*	0.103A
4AA1	0.102A	0.139A	-	0.147A	0.122A	0.153A	-	0.090A*	0.098A
5AA1	0.106A	-	-	0.093A	0.105A	-	-	0.056A	0.046A*
5AA5	0.105A	-	-	0.112A	0.095A	0.102A	-	0.094A*	0.110A
5AB2	0.150A	0.171A	-	0.174A	0.147A	0.168A	-	0.119A	0.113A*
5BC4	0.148A	0.209A	-	0.221R	0.176A	0.21R	-	0.158A	0.138A*
5BC8	0.077A	0.330R	0.096A	0.227R	0.96A	0.191A	-	0.148A	0.086A*
5BE4	0.078A	0.110A	-	0.118A	0.078A	0.100A	-	0.065A*	0.125A
5BE20	0.104A	0.177A	0.095A*	0.198A	0.099A	0.157A	-	0.121A	0.167A
5D5	0.078A	0.148A	-	0.133A	0.075A	0.098A	-	0.067A*	0.107A

W - Wakeby

LW - Log-Wakeby

For the sake of comparison, the 95% confidence level critical values in Table (10) were utilized. This way, it was then possible to classify the computed values Δ accordingly as to whether the fit they represented was acceptable (A) or not acceptable (R). This calssification is also included in Table (10).

Generally, all the tested flood probability distributions gave acceptable fits to the flood data according to the Smirnov-Kolmogorov test. However, there were some isolated cases when fitted distributions were not acceptable according to the test. The probability distribution whose fit was most frequently rejected was the Pearson Type 3 distribution while the Wakeby and the log-Wakeby had the least number of cases of rejection. The most important point to note is that in most cases, all the flood probability distributions usually gave acceptable fits in any one given catchment so that, conventionally, it is impossible to identify the most optimum flood probability model with this test. However, the minimum of the computed values of Δ for all the distributions in a given catchment can be used as an identifier of the best fitting distribution according to the Smirnov-Kolmorov test, irrespective of its location in the test-region, as is done in Kite (1977). In Table 10, this minimum is shown by a "*" in a given catchment. The consistency with which each of the nine probability distributions gives the minimum Δ is shown in Table 11. Thus, using the same procedure as in Kite (1977), we find that the order of goodness with respect to the Smirnov-Kolmogorov test of these nine distributions, from the best to the worst is:

1. Wakeby
2. Log-Wakeby
3. Log-normal type 3
4. Walter Boughton/Log-Fisher-Tippet
5. Log-Pearson Type 3
6. Fisher-Tippet
7. Pearson Type 3/Log-Walter Boughton

Thus, the Smirnov-Kolmogorov test shows that, generally, the Wakeby distribution is significantly superior than the other eight probability distributions.

Theoretically, the chi-square test statistic x^2 is generally easy to compute for any fitted distribution with known parameters, once the number of class intervals and the class limits are known. Thus, the most crucial step in the computation of χ^2, after the successful completion of the determination of the distribution parameters is the determination of the number of classes to be used and subsequently the class limits.

As the hydrologic samples were of the order of n = 20 to 50, seven classes for k, were used in equation (192) for the chi-square test. For the three parameter distri-

TABLE 11: CONSISTENCY OF BEST FIT BY SMIRNOV-KOLMOGOROV TEST.

DISTRIBUTION	CONSISTENCY (PER-CENT)
Log-normal 3	11.7
Pearson 3	0.o
Log-Pearson 3	8.3
Fisher-Tippet	5.0
Log-Fisner-Tippet	10.0
Walter Boughton	10.0
Log-Walter Boughton	0.0
Wakeby	35.0
Log-Wakeby	20.0

butions, seven class intervals were chosen so that the number of degrees of freedom was 3 and for the Wakeby and Log-Wakeby distribution with five parameters, also, seven class intervals were chosen so that in these cases, the number of degrees of freedom was one. For simplicity, these class intervals were assumed to have equal probability of occurrence equal to $p_i=1/k$ so that the expected number of frequencies were $np_i=n/k$. The observed frequencies n_i were then computed by a simple search process which is equivalent to plotting the cumulative distribution function F(x) against x for the variate x, then dividing the F(x) into k equal classes and then finding accordingly the number of observations which fall in each class. For the sake of comparison, the 95% confidence level critical values of χ_o^2 were utilized. For the three parameter distributions, the $\chi_o^2 = 7.815$ and for the five parameter distributions, $\chi_o^2 = 3.54$. In this way, the computed values of χ^2 given in Table (12) (where the notation is similar to that of Table (10)) are also classified accordingly as to whether the fit they represented are acceptable (A) or not acceptable (R).

Generally, all the tested flood probability distributions gave acceptable fits to the flood data according to the chi-square goodness of fit test although the Pearson Type 3, the Fisher Tippet and the Walter Boughton distributions had higher incidences of rejection of fit. Nevertheless, all the flood

TABLE 12: GOODNESS OF FIT BY THE CHI-SQUARE TEST

	DISTRIBUTION								
RGS	LN3	P3	LP3	FT	LFT	WB	LWB	W	LW
1CE1	1.412A	1.000A*	-	1.824A	1.824A	3.882A	-	2.235A	5.941R
1CB2	2.111A*	-	2.889A	25.444R	3.667A	5.222A	-	2.111A	10.667R
1BA1	2.200A*	-	-	10.133R	4.067A	9.200R	-	4.533R	3.133A
1BC1	-	5.600A	7.600R	5.200A	5.200A	6.400A	-	3.200A*	15.200R
1BG4	4.121A	7.091A	2.848A	15.152R	1.152A	1.152A*	-	5.394R	7.091R
1BD1	2.000A*	8.667R	-	12.667R	4.000A	8.667R	-	4.000R	4.667R
1BE1	12.529R	12.529R	-	12.529R	12.529R	10.882R	-	7.176R*	32.706R
1BE2	3.697A	6.242A	3.273A	11.758R	4.121A	16.424R	-	2.424A*	2.424A
1BE5	1.576A*	1.576A	1.576A	1.576A	1.567A	1.567A	-	4.121R	5.818R
1BB1	16.000R	11.200R	-	9.600R	10.800R	-	-	10.800R	3.600A*
1GG1	9.562R	-	-	33.187R	9.562R	7.375A	-	6.937R	5.625R*
1CJ1	7.944R	6.389A	-	5.611A	5.611A	7.944R	-	2.889A*	9.111R
1KA5	7.333A*	10.000R	-	7.333A	10.000R	18.000R	-	8.000R	12.333R
1LS3	-	-	-	5.333A	6.111A	10.778R	-	4.556R*	13.111R
1CB5	-	-	-	1.043A*	2.870A	7.130A	-	2.261A	3.478A
1CB3	5.611A	-	-	4.444A	4.056A*	5.611A	-	5.611R	5.222R
1CD1	11.320R	-	-	7.680A	13.560R	7.680A*	-	7.680R	14.120R
1CA2	-	-	-	12.609R	5.304A	12.699R	-	4.087R*	13.826R
1CB1	4.552A	12.276R	3.586A	19.034R	4.552A	4.552A	-	4.552A	2.138A*
1DA2	3.667A	4.444A	-	4.056A	6.000A	1.333A*	-	2.111A	10.667R
1FG1	1.600A	1.600A	-	1.600A	1.600A	1.200A*	-	3.200A	6.800R
1JC13	8.880R	-	9.440R	22.880R	6.640A	6.640A	-	2.160A*	9.440R
1KC3	1.687A	1.687A	-	0.813A*	0.813A	0.813A	-	1.250A	1.687A
1FG2	6.333A	6.333A	6.333A	6.333A	5.167A	5.167A	-	3.417A	2.833A*
1HA4	11.880R	-	13.000R	16.920R	13.000R	10.480R*	-	12.160R	26.440R

TABLE 12: GOODNESS OF FIT BY THE CHI-SQUARE TEST

RGS	DISTRIBUTION								
	LN3	P3	LP3	FT	LFT	WB	LWB	W	LW
1FF2	1.652A*	-	4.087A	9.565R	4.087A	12.000R	-	3.478A	5.304R
1FE2	2.667A	4.000A	-	6.667A	1.333A*	2.667A	-	5.333R	4.667R
1FE1	-	-	-	4.000A	2.000A*	4.000A	-	4.000R	8.667R
2B2	12.118R	12.118R	-	14.176R	12.118R	14.176R	-	8.000R*	8.142R
2EB3	2.848A	-	-	7.939R	2.000A*	4.545A	-	2.848A	4.970R
2EC2	10.233R	-	-	16.744R	13.163R	18.698R	-	13.163R	2.093A*
2FA2	9.350R	9.350R	6.200A	11.800R	4.800A*	16.000R	-	7.250R	6.900R
2FC5	14.500R	14.500R	-	14.500R	10.000R	18.000R	-	17.000R	6.500R*
2EC3	5.610A	-	-	25.073R	6.293A	8.341R	-	2.878A*	7.659R
2GD2	22.205R	-	-	21.128R	10.718R	-	-	16.103R	7.846R*
3BA10	1.512A	37.220R	4.244A	46.927R	3.902A	19.268R	-	13.805R	2.878A*
3PA17	4.400A	-	2.800A	6.400A	2.000A*	12.000R	-	3.200A	8.800R
3BC12	5.611A	32.833R	-	32.833R	10.278R	14.944R	-	9.111R	9.500R*
3F2	1.655A	-	1.655A	14.690R	1.172A*	1.655A	-	1.655A	4.552R
3DA2	9.000R	-	-	12.769R	5.769A	6.308A	-	11.692R	3.077A*
3BA18	2.200A*	4.067A	-	2.200a	2.200A	5.933A	-	2.200A	5.467R
3BB10	4.370A*	6.963A	-	4.370A	4.889A	5.926A	-	6.963R	6.444R
4AA2	2.848A	18.121R	2.424A	22.364R	2.424A*	7.090A	-	7.515R	5.394R
4CA2	5.720A	-	-	4.600A	5.720A	4.600A	-	3.200A*	7.680R
4BE1	2.235A	1.412A	-	2.235A	3.882A	2.235A	-	1.412A*	3.471A
4AA4	1.824A	-	7.176A	19.941R	5.529A	17.059R	-	4.294R	1.412A*
4BC2	1.733A	1.267A	-	2.667A	0.800A*	1.733A	-	2.667A	4.533R
4AB5	6.062A	-	-	4.750A	6.062A	3.875A	-	6.937R	2.562A*
4G1	4.00A	31.467R	6.489A	51.289R	4.622A	1.511A*	4.622A	7.111R	4.311R
4CB4	5.189A	6.703A	-	13.892R	2.919A	7.081A	-	2.162A*	4.243R

TABLE 12: GOODNESS OF FIT BY THE CHI-SQUARE TEST

	DISTRIBUTION								
RGS	LN3	P3	LP3	FT	LFT	WB	LWB	W	LW
4BB1	2.000A*	2.400A	-	2.000A	2.000A	5.600A	-	2.800A	7.600R
4AA1	6.389A	16.111R	-	16.111R	11.444R	17.278R	-	5.611R*	6.788R
5AA1	5.189A	-	-	6.703A	5.189A	-	-	1.027A*	1.027A
5AA5	7.130A	-	-	10.783R	5.304A*	7.739A	-	5.913R	10.174R
5AB2	5.333A	10.667R	-	10.667R	4.667A*	5.333A	-	10.000R	8.000R
5BC4	9.333R	12.000R	-	12.667R	13.333R	15.333R	-	12.667R	8.667R*
5BC8	2.870A	9.348R	2.261A*	21.130R	2.261A	19.304R	-	5.913R	4.087R
5BC4	1.742A*	8.516R	-	8.968R	1.742A	2.645A	-	2.645A	8.968R
5BE20	5.364A	13.000R	4.727A*	14.273R	4.727A	8.545R	-	5.364R	7.273R
5D5	4.692A	3.077A	-	7.385A	2.538A	2.538A	-	2.000A*	4.154R

probability distributions gave chi-square values which were usually acceptable in any one given catchment. Thus, conventionally, as was mentioned concerning the Smirnov-Kolmogorov test it is impossible to identity uniquely the most optimum flood probability distribution. However, proceeding similarly as in the case of the Smirnov-Kolmogorov test, the minimum of the computed values of χ^2 for all the distributions in a given catchment can be used as an identifier of the best fitting distribution according to the chi-square test, irrespective of its location in the test region. In Table 12, this minimum is shown by a "*" in a given catchment. Again, the consistency with which each of the nine probability distributions gives the minimum χ^2 is shown in Table 13. Accordingly, based on the chi-square test, the order of goodness of the nine distributions from the best to the worst is:

1. Wakeby
2. Log-Wakeby
3. Log-Fisher-Tippet
4. Log-normal 3
5. Walter Boughton
6. Log-Pearson 3/Fisher Tippet
7. Pearson 3
8. Log-Walter Boughton

TABLE 13: CONSISTENCY OF BEST FIT BY CHI-SQUARE TEST

DISTRIBUTION	CONSISTENCY (PERCENT)
Log-normal 3	16.7
Pearson 3	1.7
Log-Pearson 3	3.3
Fisher-Tippet	3.3
Log-Fisher-Tippet	18.3
Walter Boughton	10.0
Log-Walter Boughton	0.0
Wakeby	25.0
Log-Wakeby	21.7

We can notice that the ranking of these probability distributions in this order of goodness, according to the chi-square test is almost similar to that when the Smirnov-Kolmogorov test is used. Thus, the chi-square test, similarly, shows that the Wakeby distribution is significantly more superior compared with the other eight distributions considered. The results of the AIC test are presented and discussed below.

For a given probability distribution, sixteen AIC values were computed for all the sixteen possible combinations of having at most three outlying peak flow observations in either the low side or the high side of extremes in a given sample of observations in a given catchment. The minimum of these sixteen AIC values identified the most optimum combination of the outlying values (Kitagawa, 1979). This minimum AIC value and the corresponding outlier-combination was recorded and the process repeated for all the distributions and catchments in which realistic distribution parameters exist. In Table 14 are presented these minimum AIC values. Against each minimum AIC entry in the table, the corresponding outlier-combination is also shown. This combination of outliers represents the number of observations in the corresponding extreme sides which cannot be modelled adequately by the distribution under consideration in the given catchment. The first entry in brackets represents the number of low outliers, while the other represents the number of high outliers. Since all the

TABLE 14: GOODNESS OF FIT BY AIC TEST

	DISTRIBUTION								
RGS	LN3	P3	LP3	FT	LFT	WB	LWB	W	LW
1CE1	530.7(0,0)	526.8(0,0)	-	534.4(0,0)	524.6(0,0)	538.6(0,0)	-	824.2(0,0)*	590.1(0,0)
1CB2	325.9(0,0)*	-	327.0,0)	462.9(0,0)	327.5(0,0)	337,9(0,0)	-	328.3(0,0)	344.2(0,0)
1BA1	323.6(0,0)	-	-	361.7(0,0)	324.4(0,0)	3357(0,0)	-	323.0(0,0)*	360.9(0,0)
1BC1	-	-	-	340.9(0,0)	344.3(0,0)	338.9(0,0)	-	338.5(0,0)*	905.2(0,0)
1B64	244.2(0,0)*	304.6(0,0)	244.6(0,0)	305.4(0,0)	245.4(0,0)	247.5(0,0)	-	251.(0,0)	269.5(0,0)
1BD1	244.4(0,0)	254.9(0,0)	-	256.7(0,0)	242.3(0,0)*	254.4(0,0)	-	246.2(0,0)	259.2(0,0)
1BE1	345.9(0,0)*	347.9(0,0)	-	350.1(0,0)	1274.6(1,0)	367.8(1,0)	-	346.2(0,0)	259.2(0,0)
1BE2	134.0(0,0)*	165.8(0,0)	134.5(0,0)	191.0(0,2)	134.1(0,0)	210.5(0,0)	-	138.1(0,0)	143.3(0,0)
1BE5	272.2(0,0)*	272.4(0,0)	272.9(0,0)	273.5(0,0)	272.6(0,0)	278.1(0,0)	-	275.2(0,0)	308.9(0,0)
1BB1	399.0(0,0)*	404.3(0,0)	-	399.0(0,0)	404.0(0,0)	-	-	401.6(0,0)	703.6(0,0)
1GG1	448.4(0,0)*	-	-	660.8)0,0)	452.2(0,0)	453.2(0,0)	-	473.7(0,0)	451.6(0,0)
1JG1	582.4(0,0)	585.3(0,0)	-	582.2(0,0)	586.3(0,0)	588.1(0,0)	-	578.4(0,0)*	626.3(0,0)
1KA5	169.1(0,0)	183(0,0)	-	164.1(0,0)	177.9(1,0)	217.8(0,0)	-	161.1(0,0)	258.6(0,0)
1LA3	-	-	-	212.2(0,0)	213.5(0,0)	232.3(0,0)	-	211.5(0,0)*	322.8(0,0)
1CB5	-	-	-	255.3(0,0)	267.4(0,0)	278.3(0,0)	-	255.2(0,0)*	387.2(0,0)
1CB3	290.2(0,0)	-	-	286.9(0,0)*	290.6(0,0)	287.1(0,0)	-	289.9(0,0)	342.5(0,0)
1CD1	436.9(0,0)*	-	-	437.3(0,0)	182.9(1,0)	437.5(0,0)	-	438.7(0,0)	1121.0(0,0)
1CA2	-	-	-	286.1(0,0)	898.7(0,1)	304.0(0,0)	-	281.1(0,0)*	378.3(1,0)
1CB1	343.3(0,0)*	408.2(0,0)	343.4	403.7(0,0)	344.7(0,0)	350.3(0,0)	-	354.9(0,0)	355.5(1,0)
1DA2	562.9(0,0)	570.5(0,0)	-	559.9(0,0)	566.1(0,0)	559.5(0,0)*	-	562.9(0,0)	779.1(0,0)
1FG1	477.4(0,)	477.0(0,0)	-	477.9(0,0)	476.0(0,0)*	478.1(0,0)	-	478.3(0,0)	846.0(0,0)
1JC13	126.0(0,0)	-	121.9(0,0)	225.9(0,0)	121.0(0,0)	119.2(0,)	-	112.9(0,0)*	187.8(0,0)
1KC3	560.4(0,0)*	566.9(0,0)	-	560.8(0,0)	561.1(0,0)	561.5(0,0)	-	567.0(0,0)	580.5(0,0)
1FG2	356.1(0,0)*	356.8(0,0)	356.4(0,0)	356.9(0,0)	356.8(0,0)	363.3(0,0)	-	358.8(0,0)	633.1(0,0)

TABLE 14: GOODNESS OF FIT BY AIC TEST

	DISTRIBUTION								
RGS	LN3	P3	LP3	FT	LFT	WB	LWB	W	LW
1HA4	609.5(0,0)	-	604.8(0,0)	601.0(0,0)*	608.1(0,2)	611.0(0,0)	-	895.8(0,0)	732.0(0,0)
1FF2	231.0(0,0)*	-	233.6(0,0)	2433(0,0)	233.9(0,0)	241.2(0,0)	-	235.8(0,0)	247.2(0,0)
1FE2	319.2(0,0)*	324.4(0,0)	-	320.0(0,0)	319.3(0,0)	321.8(0,0)	-	325.3(0,0)	338.1(1,0)
1FE1	-	-	-	280.1(0,0)	278.9(0,0)	292.7(0,0)	-	278.8(0,0)*	428.2(0,0)
2B2	230.2(0,0)	233.6(0,0)	-	230.4(0,0)	230.8(0,0).	231.3(0,0)	-	230.2(0,0)*	847.6(0,0)
2EB3	286.4(0,0)	-	-	330.3(0,0)	278.6(0,0)	688.8(0,0)	-	286.0(0,0)*	303.0(0,1)
2EC2	236.6(0,0)	-	-	270.5(0,0)	228.4(0,0)	268.0(0,0)	-	225.7(0,0)*	254.6(0,0)
2FA2	292.0(0,0)	299.7(0,0)	282.8(0,0)	293.6(0,0)	295.5(2,0)	297.7(0,0)	-	282.7(0,0)	631.9(0,0)
2FC5	309.2(0,1)	351.7(0,0)	-	356.2(0,3)	300.1(0,0)	341.8(0,0)	-	326.1(0,0)	295.3(0,0)*
2EC3	269.4(0,0)	-	-	429.8(0,3)	269.1(0,0)	275.5(0,0)	-	267.5(0,0)*	333.4(0,0)
2GD2	459.2(0,0)	-	-	493.8(0,0)	454.3(0,0)	-	-	454.2(0.0)*	468.3(2,0)
3BA10	336.1(0,0)*	118.4(2,0)	339.4(0,0)	663.1(0,0)	345.6(0,0)	397.6(0,0)	-	376.6(0,1)	340.2(0,0)
3BA17	151.5(0,0)	-	151.8(0,0)	188.6(0,0)	151.9(0,0)	212.1(0,0)	-	151.4(0,0)*	187.8(0,0)
3BC12	514.8(0,0)*	578.8(0,0)	-	581.2(1,0)	515.1(0,0)	580.0(1,0)	-	538.2(0,0)	516.0(0,0)
3F2	516.5(0,0)*	-	516.7(0,0)	555.4(0,0)	516.6(0,0)	520.0(0,0)	-	524.4(0,0)	538.7(0,0)
3DA2	471.4(0,0)*	-	-	493.3(0,0)	471.8(0,0)	493.4(0,0)	-	481.3(0,0)	490.8(0,1)
3BA18	216.8(0,0)	216.7(0,0)	-	219.6(0,0)	214.0(0,0)*	222.6(0,0)	-	217.8(0.0)	259.5(0,0)
3BB10	291.6(0,0)	198.5(0,)	-	192.5(0,0)	189.1(0,0)	194.4(0,0)	-	189.0(0,0)*	201.7(0,0)
4AA2	378.1(0,0)	455.5(0,0)	374.4(0,0)	453.5(0,0)	473.5(0,0)	373.8(0,0)	-	383.0(0,0)	390.5(0,0)
4CA2	671.2(0,0)*	-	-	678.6(0,0)	673.6(0,1)	679.7(0,0)	-	693.(0,0)	757.9(0,0)
4BE7	508.5(0,0)	506.1(0,0)	-	509.9(0,0)	506.6(0,0)	509.7(0,0)	-	505.8(0,0)	524.7(0,0)
4AA4	362.5(0,0)*	-	363.1(0,0)	440.1(0,0)	366.0(0,0)	398.5(0,0)	-	377.2(0,0)	373.5(0,0)
4BC2	461.5(0,0)*	461.9(0,0)	-	461.8(0,0)	461.6(0,0)	464.1(0,0)	-	462.3(0,0)	744.1(0,0)
4AB5	375.1(0,0)*	-	-	375.2(0,0)	375.2(0,0)	376.8(0,0)	-	413.8(0,0)	381.7(0,0)
4G1	887.6(0,0)	1260.2(0,0)	879.1(0,0)	114.9(0,0)	877.0(0,0)	879.3(0,0)	890.5(0,0)	874.8(0,0)*	903.3(0,0)
4CB4	490.8(0,0)	489.0(0,0)	-	493.2(1,0)	485.7(0,0)	518.0(0,0)	-	473.9(0,0)*	526.1(0,0)

TABLE 14: GOODNESS OF FIT BY AIC TEST

RGS	DISTRIBUTION								
	LN3	P3	LP3	FT	LFT	LW	LWB	W	LW
4BB1	380.1(0,0)	378.8(0,0)	-	381.2(0,0)	378.9(0,0)	388.2(0,0)	-	378.7(0,0)	407.7(0,0)
4AA1	400.7(0,0)	432.5(0,0)	-	435.3(0,0)	402.3(0,0)	431.8(0,0)	-	400.4(0,0)*	404.1(0,0)
5AA1	386.9(0,0)	-	-	383.3(0,0)	392.2(0,1)	-	-	378.0(0,0)*	381.2(0,0)
5AA5	237.4(0,0)	-	-	239.2(0,0)	231.3(0,0)	240.4(0,0)	-	230.9(0,0)*	242.4(0,0)
5AB2	235.5(0,0)*	251.5(0,0)	-	252.0(0,0)	236.0(0,0)	251.7(0,0)	-	243.5(0,0)	239.7(0,0)
5BC4	281.9(0,0)*	294.2(0,0)	-	296.7(0,0)	282.3(0,0)	295.7(0,0)	-	283.8(0,0)	29[illegible].8(0,1)
5BC8	294.5(0,0)*	404.3(0,0)	296.2(0,0)	362.2(0,0)	295.7(0,0)	334.3(0,0)	-	321.2(0,0)	30[illegible].5(0,0)
5BE4	323.7(0,0)	345.3(0,0)	-	358.2(0,0)	325.6(0,0)	325.0(0,0)	-	322.9(0,0)*	340.0(0,0)
5BE20	292.3(0,0)*	318.4(0,0)	292.6(0,0)	321.7(0,0)	294.3(0,0)	305.9(0,0)	-	300.8(0,0)	316.3(0,0)
5D5	371.0(0,1)*	396.0(0,0)	-	396.1(0,0)	373.3(0,0)	379.5(0,0)	-	377.3(0,0)	398.2(0,0)

distributions have different shapes in their tails (Shen et al, 1980), each fits a given sample differently from the other. Therefore, the minimum AIC values given in Table 14 represent a measure of the objective measure of the degree of goodness of fit of the distributions for the samples of the catchments when the corresponding outliers are taken into consideration. The smaller this measure is, the better is the fit. Thus, for a given catchment, the best fitting model gives the least value of the minimum AIC values, which is shown by "*" in the table. The consistency with which each of the nine probability distributions gives the overall minimum AIC value is shown in Table 15. Accordingly, based on the AIC test, the order of goodness of the nine distributions from the best to the worst is:

1. Wakeby/log-normal 3
2. Log-Fisher-Tippet
3. Fisher-Tippet
4. Log-Wakebt/Walter-Boughton
5. Pearson Type 3/log-Pearson Type 3/log Walter-Boughton.

Regarding the existence of outliers, it was pointed out earlier that, when we assume a particular distribution for a given set of data, we can react quite strongly towards some observations, and classify them as outliers, which may cause no specific concern when we assume another distribution for the data. Thus,

TABLE 15: CONSISTENCEY OF BEST FIT BY AIC TEST

DISTRIBUTION	CONSISTENCY (PERCENT)
Log-normal 3	43.3
Pearson 3	0.0
Log-Pearson 3	0.0
Fisher-Tippet	3.3
Log-Fisher-Tippet	6.7
Walter-Boughton	1.7
Log-Walter-Boughton	0.0
Wakeby	43.3
Log-Wakeby	1.7

it is only when outliers are consistently identified under different basic model assumption, can we brand such observations as true outliers with an improved degree of certainity. In Table 14, we can see that, in any given catchment, outliers are located in not more than two models when the others give no evidence of the existence of such outliers. In most cases, no outliers are found by any distributions. The frequency with which outliers are found within the catchments by the various distributions is as follows:

log-normal 3	-	3.3%
Pearson Type 3	-	1.7%
log-Pearson Type 3	-	0.0%
Fisher-Tippet	-	13.3%
log-Fisher-Tippet	-	16.7%
Walter-Boughton	-	5.0%
log-Walter-Boughton	-	0.0%
Wakeby	-	1.7%
log-Wakeby	-	10.0%

We can also notice that, except in only one case of the three parameter log-normal distribution, all the overall-minimum-AIC cases are in no-outlier cases also.

However, the fact that some models show no outliers at all in any one given catchment proves that, those which are found with the other models are not overall outliers. This implies that all the peak flow samples considered have no true outliers. The computed AIC values, under the assumption of no outliers in the samples show that the distributions can be ranked in order of their goodness of fit from the best to the worst as:

1. Wakeby
2. log-normal 3
3. log-Fisher-Tippet
4. Fisher-Tippet
5. log-Wakeby/Walter Boughton
6. Pearson Type 3/log-Pearson Type 3/log-Walter Boughton.

Thus, identification of the Wakeby distribution by the Smirnov-Kolmogorov and the chi-square tests is further confirmed, without the doubt of coincidence by the more objective and reliable AIC test which again identifies the Wakeby distribution to be more superior compared to the others, although only with a small margin over the three parameter log-normal distribution. Table 16 shows the average consistency of best fit in the catchments based on the results of the Smirnov-Kolmogorov test, the chi-square test and the AIC test. The results of this table show that, again the Wakeby distribution is remarkably more superior than the rest, followed by the three parameter log-normal.

TABLE 16: AVERAGE CONSISTENCY OF BEST FIT BY SMIRNOV-KOLMOGOROV, CHI-SQUARE AND AIC TESTS

DISTRIBUTION	CONSISTENCY (PERCENT)
Log-normal 3	23.4
Pearson 3	0.6
Log-Pearson 3	3.9
Fisher-Tippet	3.9
Log-Fisher-Tippet	12.2
Walter-Boughton	7.2
Log-Walter-Bougnton	0.0
Wakeby	34.4
Log-Wakeby	14.4

CHAPTER 4

4. SUMMARY AND CONCLUSION

There is no general agreement among hydrologists, as to which of the various theoretical probability distributions available, should be used. Presently, no general agreement has been reached for the preferable techniques and no standards have been established for design purposes. Acceptance of a certain distribution is based on the goals and conditions that are to be fulfilled and satisfied by the distribution (Zelenhasic, 1970). Such conditions include the suitability of a distribution for flood analysis as well as the goodness-of-fit.

The most commonly used goodness-of-fit tests in hydrology include the chi-square and the Smirnov-Kolmogorov test statistics. These approaches when used conventionally in the tests of goodness of fit of empirical probability distributions are very subjective in that the results depend entirely on a predetermined level of confidence. Furthermore, these tests can pass more than one probability distributions without any distinction of the relative goodness of fit. This has been one of the major reasons contributing to the existence of the disagreement among hydrologists in the choice of the most optimum flood probability model. However, in this study, the modified usage of these tests is utilised. In all, nine flood probability distributions were fitted to flood data obtained from sixty river gauging stations distributed randomly within the five major river basins of Kenya. This study uses not only the chi-square and the Smirnov-Kolmogorov tests but also the Akaike Information Criterion which is based on the entropy maximisation principle, to test the goodness of fit of the nine probability distributions. The Akaike Information Criterion is used not only for the model identification but also for the identification of any outliers that can exist in peak flow data.

The modified Smirnov - Kolmogorov and the chi-square tests identify the Wakeby distribution as the best for flood frequency analysis compared to the rest. This is further confirmed by the more objective and reliable AIC test. Thus, the three goodness of fit tests identify the Wakeby distribution to be generally the most suitable for flood frequency analysis in Kenya. One of the worst fitting distributions is the log-Pearson Type 3, despite its popularity worldwide.

The results also show that the peak flow data does not have any true outliers. However, some of the poorly fitting distributions occasionally show some outlying observations when the more flexible distributions indicate no such outlying values. The Fisher-Tippet distributions has the highest frequency of showing outlying values.

In conclusion therefore, we can see that the Akaike Information Criterion goodness of fit test, which has not been used in flood frequency model testing before, has been successively used for the identification of an optimum flood frequency model as well as for the identification of outliers that can exist in peak flood data. We also note that, for the purposes of achieving efficient flood designs in the Kenyan drainage basins, the Wakeby distribution should be used whenever possible and not the log-Pearson Type 3 distribution as it has been the tradition.

4.1 SUGGESTIONS FOR FUTURE WORK

The existence of reliable parameter estimates for a distribution is an important aspect in model identification. There are usually many methods available for the estimation of the distribution parameters. However, when the maximum likelihood method of parameter estimation exists for a given probability distribution, then it is considered to be the most efficient method. Except for the Wakeby and the log-Wakeby distribution, all the other probability distributions which were considered in this study have maximum likelihood routines, although mainly implicit in form. The method of probability weighted moments was used for the estimation of the parameters of both the Wakeby and log-Wakeby distributions. It was earlier pointed out that this method occasionally yields unsuitable parameters especially with short records. Thus a study should be carried out to find a more reliable parameter estimation procedure for the Wakeby to improve its efficiency in flood frequency analysis.

In the previous chapters, we described some of the probability distributions that can be used to carry out a frequency analysis on a set of observed or computed data. Using any one of these techniques, the event magnitude corresponding to a given probability of occurrence can be determined This event magnitude applies only to the exact location at which the original observations were made. Frequently, for design purpose, it is necessary to estimate event magnitudes at sites where no observations have been taken. Regional analysis is the term given to the techniques which make such estimations possible. Thus, although the Wakeby distribution was established to be the most optimum for the single site flood frequency analysis in Kenya, a study should be carried out to regionalise its usefulness.

ACKNOWLEDGEMENTS

The author would like to thank the World Meteorological Organisation (WMO) for their scholarship under the auspicies of the Kenya Meteorological Department. To the University of Nairobi for granting me study leave to take up this scholarhsip at Colorado State University, Fort Collins.

The author wishes to express his gratitude to Dr. Jose D, Salas of Colorado State University for his continued guidance and encouragement during the initial stages of this study. To the Director and the staff of Water Resources Department, Ministry of Water Development, Kenya, for assisting me in the collection of the annual flood peak data. To the Chairman and staff of the Meteorology Department, University of Nairobi for their guidance and encouragement during the preparation of this work. To the staff of the Institute of Computer Science, University of Nairobi, for their assistance during the analysis stage of this Study. Special thanks are extended to Dr. Laban Ogallo, my University supervisor, for his helpful comments while compiling this report.

The author also wishes to thank Mr. Munyi for drafting some of the figures and Mrs. G.N. Macharia for her patience while typing the manuscript. Special gratitude is extended to Redempta, my wife and our children for exercising great patience and understanding during the prepartion of this report and especially so when the author was away on study leave.

REFERENCES

Abramowitz, M., and Stegun, I.A., 1965: Handbook of Mathematical Functions: Dover Publication, New York.

Akaike, H., 1970: Statistical Predictor Identification. Annals, Inst. Statis. Math., Vol. 22, pp.20-217

Akaike, H., 1971: Determination of the number of factors by an Extended Maximum likelihood principle. Research Memo. No.44, Inst. Math. Statis., March, 1971.

Akaike, H., 1973: Information Theory and an Extension of the Maximum Likelihood Principle. 2nd International Symposium on Information Theory Eds. B.N. Petrov and F. Csaki, Budapest, Akademia, Kiado pp.267-281.

Akaike, H., 1976: Canonical Correlation Analysis of Time Series and the use of an Information Criterion. Systems Identification. Eds. R.K. Mehra and D.G. Lainiotis, New York; Academic Press, p.27-96.

Akaike, H., 1977: On Entropy Maximization Principle. Proc., Symp. on Applic. of Statis., Ed. P.R. Krishnaiah, Amsterdam, North Holland, p. 27-47.

America Water Works Association, 1966: Spillway Design Practice. AWWA Manual M13, New York.

Barnett, V., and Lewis, T., 1978: Outliers in Statistical Data. John Wiley and Sons, New York.

Barton, D.E., 1953: On Neyman's Smooth Test of Goodness of Fit and its Power with respect to a particular system of Alteratives. Skand Aktuartidskr., Vol. 36 pp.24-36.

Bernard, M., 1944: The Primary Role of Meteorology in Flood Flow Estimation. Trans. ASCE, Vol.109.

Blench, T., 1959: Emperical Methods. Proc. of the Symposium No.1 Spillway Design Floods, National Research Centre, Ottawa, pp. 36-48.

Bobee, B., 1975: The log-Pearson Type 3 Distribution and its Applications in Hydrology. Wat.Resour. Res. Vol. 11, No.5 pp. 681-689.

Boughton, W.C., 1980: A Frequency Distribution for Annual Floods. Wat. Resour. Res. Vol.12 pp.347-354.

Carnahan, B., Luther, H.A. and Wilkes, J.O., 1969: Applied Numerical Methods. John Wiley and Sons Inc., New York.

Chow, V.T., 1951: A General Formula for Hydrologic Frequency Analysis. Trans. Amer. Geophys. Union, Vol.32 pp. 226-230.

Chow, V.T., 1962: Hydrologic Determination of Water-

Way Areas for Design of Drainage Structures in Small Drainage Basins. Univ.Ill. Eng. Exp. Stat. Bulletin K. 462.

Chow, V.T. 1964: Handbook of Applied Hydrology McGraw Hill Co.Ltd., New York.

Clark, R.T., 1973: Mathematical Models in Hydrology. F.A.O. Irrig. and Drainage Paper No.19.

Condie, R. and Nix, G. 1975: Modelling of Low Flow Frequency Distributions and Parameter Estimation Inter. Wat. Resour. Assoc., Proc. Symp. on Water for Arid Lands, Iran.

Cramer, H. 1946: Mathematical Methods of Statistics. Princeton University Press, Princeton.

Dalrymple, T., 1964: Hydrology of Flood Control. Section 25-1 in V.T. Chow (Ed.,) Handbook of Applied Hydrology, McGraw Hill, New York.

David, H.A. Hartely, H.O. and Pearson, E.S., 1954: Distribution of Ratios, in a Single Normal Sample of Range to Standard Deviation. Biometrika, Vol. 44 pp. 482-493.

Dawdy, D.R., and Litchy, R.W., 1968: Methodology of Hydrologic Model Building. Symposium on tne use of Analogue and Digital Computers in Hydrology, Tucson International Association of Scientific Hydrology, Pub. 81 pp.347-355.

Dixon, W.J., 1951: Ratios involving Extreme Values; Annals Math. Statis., Vol. 22 pp. 107-115.

East African Meteorological Department, 1975: Climatological Statistics for East Africa. East African Community, Nairobi.

Fuller, W.E., 1914: Flood Flows. Trans. ASCE, Vol.77, p. 564-617.

Furguson T.S., 1961: On the Rejection of Outliers Proc. of the Fourth Berkely Symposium on Mathematical Statistics and Probability, Berkely pp. 253-287.

Greenwood, J.A. Landwehr, J.M., Matals, N.C. and Wallis, J.R., 1979: Probability Weighted Moments: Definition and Relation to Parameters of several Distributions Expresable in Inverse Form. Water Resour. Res., Vol.15 No.5, pp. 1049-1054.

Greis, N.P., 1983: Review of Geophysics and Space Physics. U.S National Report to International Geodesy and Geophysics, 1979-1982, Vol.21 No.3, pp.699-706.

Grubbs, F.E. 1950: Some Criteria for Testing Outlying Observations. Annals. Math. Statis. Vol.21, pp. 27-58.

Gumbel, E.J., 1958: Statistics of Extremes. Columbia University Press, New York.

Hall, W.A. and Howell, D.T., 1963: Estimating Flood Probabilities within Specific Time Intervals. Journal of Hydrology, Vol.1, No.1, pp.265-271.

Halperin, M., Greenhouse, S., Cornfield, J. and Zolokar, J., 1955: Tables of Percentage Points for Studentized Maximum Absolute Deviation in Normal Samples. Jour.Amer.Statis. Assoc., Vol.50, pp.185-195.

Harter, H.L., 1971: Some Optimisation Problems in Parameter Estimation. Optimisation Methods in Statistics. Eds. J.S. Rustagi, Academic Press, New York. pp. 33-36.

Hastings, C., 1955: Approximation for Digital Computers. Princeton University Press.

Hawkins, D.M., 1980: Identification of Outliers. Chapman and Hall Ltd.

Houghton, J.C., 1978a: Birth of a Parent: The Wakeby Distribution for Modelling Low and High Flows. Wat. Resour. Res., Vol.14 No.6. pp 1150-1109.

Houghton, J.C., 1978b: The Incomplete Means Estimation Procedure Applied to Flood Frequency Analysis. Wat. Resour. Res., Vol. 14 No.6. pp.1111-1115.

Huddart, L. and Woodward, P.G., 1975: The Prediction of Antecedent Catchment Wetness. Proc. of Symposium on Flood Hydrology held in Nairobi in

Oct. 1975. Transport and Road Research Laboratory, Supp. Report No.259. p.141-148.

IBM, 1968: System/360 Scientific Subroutines, 360-A-CM-63X Version III, Program Manual, White Plains, New York.

ICL, 1976: ICL/1900 Series, Scientific Subroutines. 4096 Series, London.

Irwin, J.O., 1925: On a Criterion for the Rejection of Outlying Observations. Biometrika, Vol.17, pp. 238-250.

Jarvis, C.S., 1942: Floods. Chap. XI-G in O.E. Meinzer (ed.), Hydrology. McGraw Hill Book Company Inc., New York.

Jenkinson, A.F., 1969: Estimation of Maximum Floods. World Meteorolgoical Orgnisation, Tech. Note. 98, Chapter 5. p.183-257.

Joiner, B.L. and Rosenblatt, J.R., 1971: Some Properties of the Range in Samples from Tukey's Symmetric Lambda Distributions. Jour., Amer. Static. Assoc., Vol.66, pp. 394-399.

Kalinin, G.P., 1960: Calculation and Forecasts of Streamflow from Scanty Hydrometric Readings. Trans. Inter-Regional Seminar on Hydrologic Networks and Methods, Bangkok, (1959). WMO Flood Control Series, No.15, pp.42-52.

Kendal, M.G. and Stuart, A., 1958: Advanced Theory of Statistics - Distribution Theory. Vol.I, Charles Griffin and Company Ltd., London.

Kendal, M.G. and Stuart, A., 1961: Advanced Theory of Statistics. Vol.II, Charles Griffin and Co.Ltd., London pp.54-56.

Kirpich, Z.P., 1940: Time of Concentration of Small Agricultural Watersheds, Civil Engineering (New York), Vol.10, No.6, pp.362-380.

Kitagawa, G., 1979: On the use of AIC for the detection of Outliers. Technometrics, Vol.21, No.2 p.193-199.

Kite, G.W., 1977: Frequency and Risk Analysis in Hydrology. Water Resources Publication, Fort Collins, Colorado.

Landwehr, J.M.,Matalas, N.C. and Wallis, J.R., 1979a: Probability Weighted Moments Compared to some Traditional Techniques of estimating Gumbel Parameters and Qualities. Wat.Resour.Res., Vol.15,No.5 pp. 1055-1064.

Landwehr J.M.,Matalas,N.C. and Wallis, J.R., 1979b: Estimation of the Parameters and Quantities of the Wakeby Distribution—Known Lower Bounds. Wat. Resour. Res.,Vol.15, No.6 pp. 1361-1372.

Landwehr, J.M.,Matalas, N.C. and Wallis, J.R. 1979c: Estimation of the Parameters and Quantities of the Wakeby Distribution—Unknown Lower Bounds, Wat. Resour. Res., Vol.15, No.6 pp.1373-1379.

Langbein, W.B., 1960: Plotting Positions in Frequency Analysis. U.S.G.S. Water Supply Paper NO.154-A, pp.48-51.

Matalas, N.C., Slack, J.R. and Wallis, J.R., 1975: Regional Skew in Search of a Parent. Wat. Resour. Res., Vol. II, NO.6, pp. 815-826.

McCalla, T.R., 1967: Introduction to Numerical Methods and FORTRAN Programming. John Wiley and Sons, Inc., New York pp.255-260.

Merrill, B., 1932: Discussion of Runoff: Rational Runoff Formulae. Jour. Agric.Res., Vol.34, No.9. pp. 797-823.

Mielke, P.W. (Jr.), 1973: Another Family Distributions of Describing and Analysing Precipitation Data., J. App. Meteorology Vol.12 No.2 pp. 275-280.

Mood, A.M., Graybil, F.A. and Boes, D.C., 1974: Introduction to the Theory of Statistics. 3rd Edition, McGraw Hill Book Company, New York.

Morel-Seytoux, H.J., 1982: Instruction Notes for Engineering Hydrology. Colorado State University, Department of Civil Engineering Fort Collins, COLorado.

Mutua, F.M., 1984: Suitability of the Watler Boughton Distribution to Flood Frequency Analysis for Western Kenya. East. Afric. and Forestry Jour., Vol.47, No.3, pp. 68-75.

O'hara, H. and Smith, F.J., 1969: The Evaluation of Definite Inetegrals by Interval Subdivision, Comp. Jour., Vol.12, pp. 179-182.

Otomo, T., Nakagawa and H. Akaike, 1972: Statistical Approach to Computer Control of Cement Rotary Kiln. Automatica, Vol. 8, pp.35-48.

Otsu, K., Horigome, M. and Kitagawa, G., 1976: On the Prediction and Stochastic Control of Ships Motion. Proceedings of the 2nd IFAC/ICIP. Symposium, Eds. M. Pitkin, J.J. Roche and T.J. Williams. Washington D.C. pp. 69-76.

Pearson, E.S. and Stephens, M.A., 1964: The ratio of Range to Standard Deviation in the Some Normal Samples. Biometrika Vol. 51, pp.484-487.

Rao, D.V., 1980a: Log-Pearson Type 3, A Generalized Evaluation. ASCE, 106 (HY5) pp.853-859

Rao, D.V., 1980b: Log-Pearson Type 3: Method of Mixed Moments. ASCE, 106 (HY6). pp.999-1005.

Rao, D.V., 1981: Three Parameter Probability Distributions. ASCE, 107 (HY3). pp.339-345.

Rasmer, C.E., 1972: Runoff from Small Agricultural Areas. Jour. of Agric. Res., Vol. 34, No.19 pp. 797-823.

Raynal, J.A. and Salas, J.D., (1985): Estimation Procedures for the Type 1 Extreme Value Distribution. Department of Civil Engineering, Colorado State University, Fort Collins, Colorado.

Reich, B.M., 1972: Log-Pearson Type 3 and Gumberl Analysis of Floods. 2nd International Symposium on Hydrology. Fort Collins, Colorado pp. 290-303.

Salas, J.D., and Smith, R.A., 1980. Computer Programs of Probability Distributions Functions in Hydrology and Water Resources. Hydrology and Water Resourntces Program, Department of Civil

Engineering, Colorado State University, Colorado.

Shapiro, S.S. and Wilk, M.B., 1965: Analysis of Variance; Test for Normality (Complete Samples). Biometrika, Vol.52, pp. 591-611.

Shapiro, S.S. and Wilk, M.B., 1968: A comparative Study of Various Tests for Normality. Jour. Amer. Stat. Assoc. Vol. 63, pp. 1343-1372.

Shapiro, S.S. and Wilk, M.B., 1972: An Analysis of Variance; Test for the Exponetial Distribution (Complete Samples). Technometrics, Vol.14 pp.355-370.

Shen, H.W., Bryson, M.C. and Ochoa, I.D., 1980: Effect of Tail Behaviour Assumptions on Flood Predictions. Wat. Resour. Res. Vol.16 No.2 pp.361-364.

Singh, K.P., 1980: Regional and Sample Skew Values in Flood Frequency Analysis of Streams in Illinois. Illinois State Water Survey Contract Report 223. Urbana, Illinois.

Sombroek, W.G. and Braun, H.M.H., 1982: Exploratory Soil Map and Agro-Climatic Zones: Ministry of Agriculture, National Agricultural Laboratory, Kenya, Soil Survey.

Tanabe, K., 1974: Statistical Regularization of a Noisy Ill-Conditioned System of Linear Equations by Akaike Information Criterion. Research Memo. No.60, Inst. Math.Statis.,Tokyo, Japan.

Terstreip, M.L. and Stall, J.B., 1974: The Illinois Urban Drainage Area Simulator, Bulletin 58. Illionois State Water Survey. Urbana, Illnois.

Tietjen, G.L. and Moore, R.M., 1972: Some Grubbs - Type Statistics for the Detection of Several Outliers. Technometrics, Vol.14. pp.583-597.

Turkey, J.W., 1960: The Practical Relationship between the Common Transformations of Percentages of Fractions and of the Amounts. Tech.Rep.Vol.36 Statist. Tech. Res. Group, Princeton Univ., Princeton, N.J.

U.S. Water Resources Council - Hydrology Committee, 1977: Guidelines for Determining Flood Frequency Information. Bulletin 17A, Washington D.C.

U.S. Weather Bureau, 1937-1970: Hydrometeorological Reports and Technical Notes.. U.S. Government Printing Office.

Vicens, G.J., Rodriquez-Iturbe, I. and Schaake, J.C.(Jr.), 1975: A Bayesian Framework for the use of Regional Information in Hydrology. Wat. Resour. Res. Vol.II, NO.3. pp.405-414.

Wald, A., 1943: Tests of Statistical Hypothesis Concerning Several Parameters when the number of Observations is large. Trans. Amer. Math.Soc. Vol.54. pp.426-482.

WMO, 1966: Estimation of Maximum Floods (Report of a Working Group on the Commission for Hydrometeorology)

WMO Tech. Note 98, World Meteorological Offfice, Geneva.(1966)

WMO, 1975: Intercomparison of Conceptional Models in Operational Hydrological Forecasting. Operational Hydrology. Report No.7. World Meteorological Office, Geneva.

Wood, E.F. and Rodriquez-Iturbe, I., 1975: Bayesian Influence and Decision Making for Extreme Hydrologic Events. Wat. Resour. Res., Vol.II, No.4. pp. 533-542.

Yevjevich, V., 1972a: Probability and Statistics in Hydrology. Water Resources Publications, Fort Collins, Colorado.

Yevjevich, V., 1972b: New Vistas for Flood Investigations. Academia Naturnale Dei Lincei, Rome Quadern, D. No. 169, pp.515-546.

Zelenhasic, E., 1970: Theoretical Probability Distributions for Flood Flows, Hydrology Paper No.42, Colorado State University, Fort Collins, Colorado.