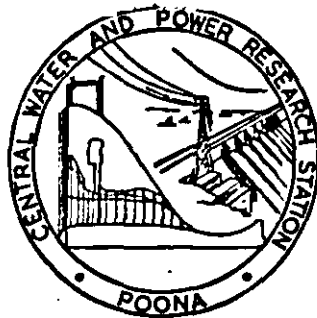


**V P Aggarwal**

with FOREWORD by  
**D V Joglekar**



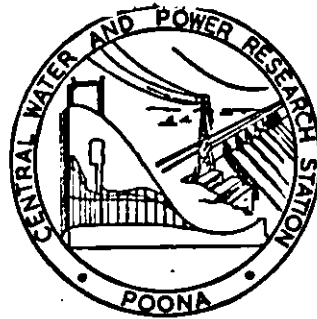
**Technical Memorandum HLO 2**  
**March 1962**

**GOVERNMENT OF INDIA**  
**CENTRAL WATER AND POWER RESEARCH STATION**  
**POONA**

# PEAK FLOW ESTIMATION BY METHOD OF MAXIMUM LIKELIHOOD

by  
**G M Panchang**  
and  
**V P Aggarwal**

Foreword by  
**D V Joglekar**



**Technical Memorandum HLO 2**

**March 1962**

**GOVERNMENT OF INDIA**  
**CENTRAL WATER AND POWER RESEARCH STATION**  
**Poona, India**

## FOREWORD

It is sometime that Gumbel's Extreme Value formula has come to be almost universally accepted for deriving the most reliable estimate of the average peak flow magnitudes of long term return periods of occurrence. Pitfalls nonetheless obtain commonly in evaluating the parameters of the formula from the past observed data of the annual peak flow measures on record. Hypothetical definitions are postulated for the plotting positions representing the various observed flow magnitudes, while theoretically none exist. Gumbel has in fact clearly admitted this position in his latest published valuable book on Extreme Value Distributions(1958).

The maximum likelihood approach which involves "prohibitive" computational labour has been processed in the Technical Memorandum in a schematized form to be readily adopted by all design engineers who are equipped with the simple aid of only a routine calculating machine and ordinary tables of logarithms. The estimates resulting are shown to be precise as well as better consistent of magnitude (ie insusceptible to large alterations of value) to withstand effectively deletions of a few years doubtfully observed flow measures or limited additions of more recent data. The feature comprises a fundamental and a very significant contribution to the subject which has for long proved to be slippery.

Poona

March 17, 1962

D V Joglekar  
Adviser, Central Board  
of Irrigation & Power.

Engineers double-check their work as they go along and then try another method to see if the final answer is the same. That is why it is safe to ride automobiles and electric trains and cross bridges. An Engineer worthy of the name would rather lose a client than design a structure he considers unsafe.

— Harold Coy

The statistician is no longer an alchemist expected to produce gold from any worthless material offered him. He is more like a chemist capable of assaying exactly how much of value it contains, and capable also of extracting this amount, and no more. In these circumstances, it would be foolish to commend a statistician because his results are precise or to reprove because they are not. If he is competent in his craft, THE VALUE OF THE RESULT FOLLOWS SOLELY FROM THE VALUE OF THE MATERIAL GIVEN HIM. IT CONTAINS SO MUCH INFORMATION AND NO MORE. HIS JOB IS ONLY TO PRODUCE WHAT IT CONTAINS.

— R A Fisher

#### PEAK FLOW ESTIMATION BY METHOD OF MAXIMUM LIKELIHOOD

The assessment of the largest magnitudes expected of floods in different rivers constitutes a problem of great significance and importance in engineering constructions and works of River valley development and flood control Projects. The problem is very complex also; for, floods are produced by different combinations of a large number of variable and non-variable conditions. The former include: the rate of rainfall; the temperature conditions; the quantity of water in reservoirs, lakes and ground at the time the flood occurs; the velocity and direction of the storm; and the many other elements which cause one flood to differ from another in the same stream.

The latter conditions include: the prevailing conditions of rainfall; the size, shape and slope of the catchment area; the character of the soil and vegetation on the catchment; the physical characteristics of the stream channel; the storage capacity in reservoirs; and many other physical characteristics of the catchment area and the stream itself. The effect of these on floods is usually invariable.

Again no two floods are exactly alike. Two storms of like intensity, velocity and direction passing over a catchment area may produce different floods. When many conditions tending to large floods occur coincidentally with great rainfall, extraordinary floods are produced. Thus floods which have occurred on some rivers have been greatly in excess of others on the same river. But a characteristic of these great floods is that they come only rarely, and usually in different years for different streams. In order to ascertain, therefore, the flood magnitude(s) that any proposed hydraulic structure(s) across a stream may be expected to meet in the normal course of years, the frequency study of the past observed floods data alone of the stream can at all provide any indication or guidance. By virtue of the natural topographical, climatic and other differences ruling from one region to another, such studies of necessity have also to be confined to individual basins separately.

## 2. EMPIRICAL FORMULAE

While the subject has interested design engineers from very early years, observed data in the past have been meagre. The methods of flow rates observation have been mostly indirect and not sufficiently elaborate in accordance with the modern standards. In order to serve the adhoc needs of constructional requirements therefore, such incomplete

data alone have provided the means for deriving different empirical formulae like

Dickens'	$Q = CA^{3/4}$	$825 < C < 1400$
Ryves'	$Q = CA^{2/3}$	$450 < C < 2700$
Modified Inglis'	$Q = K^{1/3} \times \frac{7000A}{\sqrt{A+4}}$	K, the shape factor = ratio of diameter of standard circular catchment of same area to actual stream length L of catchment.
Ali Nawaz Jung Bahadur's	$Q = CA^{(0.925 - 1/14 \log A)}$	$1700 < C < 2100$
Dredge and Burge's	$Q = 1300 W L^{1/3}$	W = Average width of basin L = Length of basin
Murphy's	$Q = A \left[ \frac{46790}{A + 320} + 15 \right]$	for areas under 10,000 sq miles in north-eastern USA
Ganguillet's	$Q = \frac{1421A}{3.11 + \sqrt{A}}$	for Swiss streams
Cramer's	$Q = \frac{80.6A}{1 + 0.1347 A^{1/3}}$	for Mohawak river in USA
Switzer & Miller's	$Q = 80PW^{1.5}$	for Miami Conservancy district, where P = rainfall factor and W = mean width of drainage area in miles.
etc, ...	etc,	

The principal lacuna in the above formulae and their like is the complete absence in them of a term or terms for the relative frequency or the RETURN PERIOD concept. They all simply aim at evaluating the magnitudes of the very highest floods alone that may ever be expected in the respective streams. But actual observed data of the daily flow rates and of the annual peak rates which have accumulated of recent years for some of the streams show the derived magnitudes from the formulae as not quite realistic always, apart from their looseness. In other words

the formulae do not come up to the modern design engineer's standards of requirements. Economy conditions impose the need for making every design as precise as possible. Thus dams or weirs are designed with fixed spans of expected life. Their heights and or other dimensions are suitably determined to withstand the extreme magnitude of the floods expected during this period and not the highest magnitude for all time. For\*

"However big floods get, there will always be a bigger one coming; so says one theory of extremes, and experience suggests it is true."

It is an obvious fallacy therefore, to formulate any expression for estimating all time maximums.

### 3. RETURN PERIOD CONCEPT

The earliest known attempts towards introducing the frequency or return period concept for different flood magnitudes were made by W E Fuller(1914) and E W Lane(1926) who both expressed the formula for the average highest flood magnitude(Q) expected to occur once every T years as

$$Q = k(\log T + b) \quad \dots (1)$$

where k and b are specific constants for the basins. Different methods of evaluating k and b have been sponsored from time to time. Better suitable forms of expressions have also been proffered for deriving by the application of statistical methods to the observed data [H A Foster (1924) Allen Hazen(1930)]. Slade(1936) proposed the transformation through the partly bounded function

$$F(y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\sqrt{2} \operatorname{clog}_e \frac{d(y+b)}{t^2+1}} e^{-\frac{z^2}{2}} dz \quad \dots (2)$$

---

\* President's Water Resources Policy Commission, 1950. Water Policy for the American People. Report, Washington, D C. Vol I, p 141.

for normalizing the data where  $e$  is the Napierian base and  $b$ ,  $c$ ,  $d$  and  $t$  are constants pertaining to the basin.

In order to explore what additional modifications may or may not be needed for analysing Indian rivers data in accordance with the procedures outlined, the a priori frequency histograms obtaining of the daily observed discharges data of six rivers for some 8 to 20 years as listed below are presented in fig 1.

Tapi at Kathore	..	1941 - 53
Narmada at Gardeshwar	..	1948 - 55
Yamuna at Tajewala	..	1946 - 55
Sone at Dehri	..	1945 - 56
Mahanadi at Sambalpur	..	1944 - 54
Damodar at Rhondia	..	1933 - 53

The distributions in terms of the original variate of the daily discharge are at once seen to be all highly skew. Their logarithmically transformed distributions (also presented in fig 1) show a large part of the skewnesses as removed, though not quite reaching to near normality. On utilising the longer term (1920-54) daily discharges data also (of the Sone at Dehri), both Slade's expression (2) or the alternatively best fitting Pearson type curve (type VI) either (fig 2) are not found to be reasonably adequate. The exceedingly large values\* of  $\chi^2$  obtaining in each case for the different respective fits made, only add to the infallibility of the otherwise correct inference derived from visual inspection of fig 2.

---

\*  $\chi^2 = 338.8$  for 12 degrees of freedom for Slade's expression;  
 $\chi^2 = 277.1$  for 11 degrees of freedom for Pearson-type VI curve;  
 $\chi^2 = 310.4$  for 18 degrees of freedom for log-normal distribution.



#### 4. OBSERVED ANNUAL PEAK FLOW RATES

Other avenues also have, therefore, been explored toward meeting the limited objective of evaluating the magnitude  $Q_T$  of the average highest flood of T years return period of occurrence in specific streams of interest. The annually observed single highest peak flow rates data of the respective streams have alone been mostly utilised for this purpose by suitably processing them. The individual values of such data interse separated by broad twelve months intervals between them subscribe better to the independency criterion usually prescribed for all statistical analyses. Their lengths for most Indian rivers extend only to about 30-40 years. Such data are better easily securable also than complete daily data. They lend better amenable for any computational manipulations by virtue of their limited numbers for the purpose of examining the suitability of various alternative processes of analyses.

Accordingly, fig 3 was prepared to present the type of variations obtaining of such data from year to year. Apart from reflecting the usual year to year irregularities of the flood magnitudes in River Sone (or any other river for that matter) fig 3 does not serve much other useful purpose. Their re-arranged presentation as in fig 4 in ascending order of the year to year peak flood magnitudes will perhaps be found better useful instead, for estimating the graduated future higher magnitudes to expect.

#### 5. FREQUENCY HISTOGRAMS

As the sequence of the years in which they are presented in fig 4 follows only the respective flow magnitudes, it was thought useful in the first instance to enumerate the number of occasions for which the annual peak flood magnitudes were palpably not different. Thus peak floods of the order of

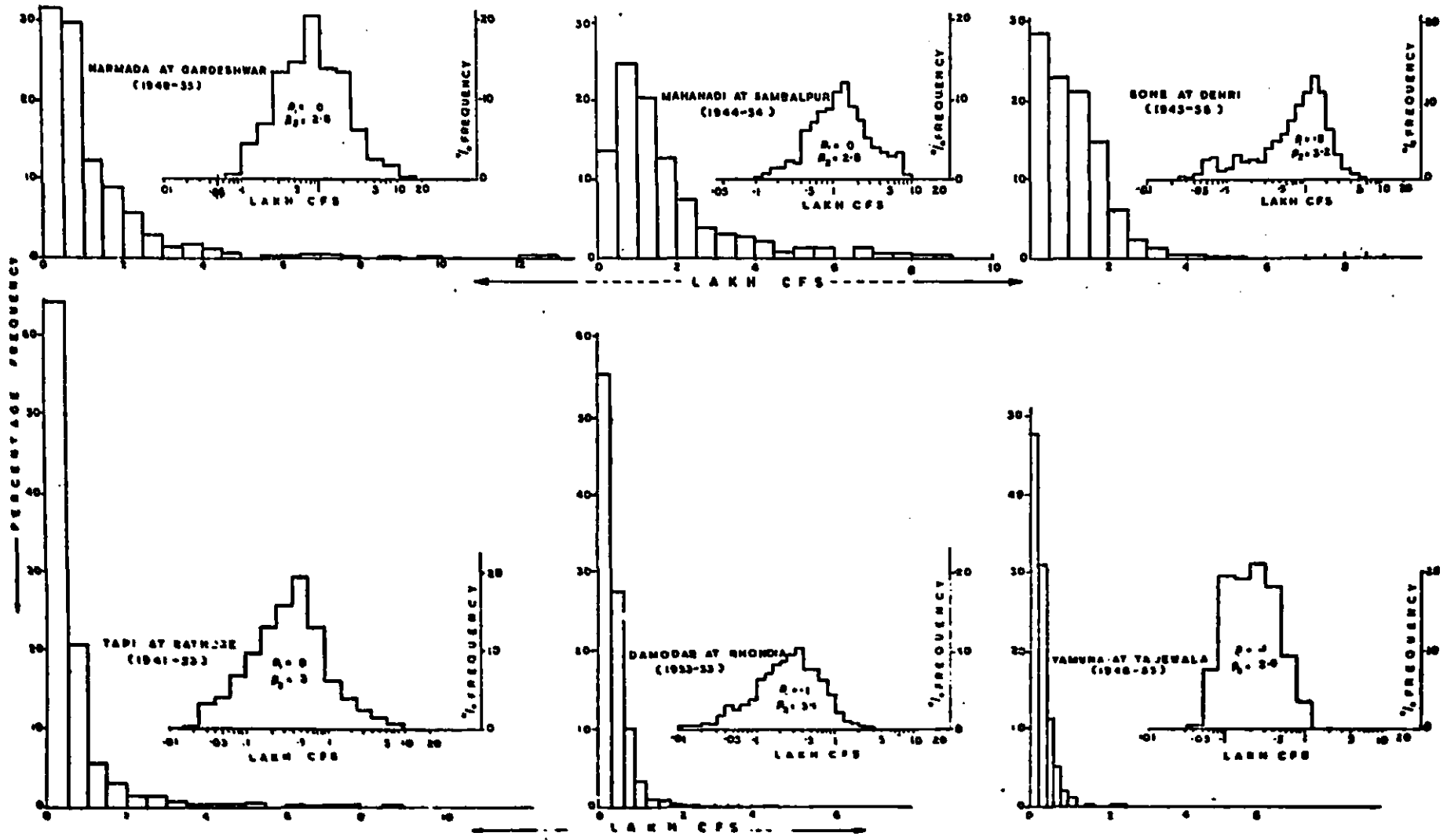
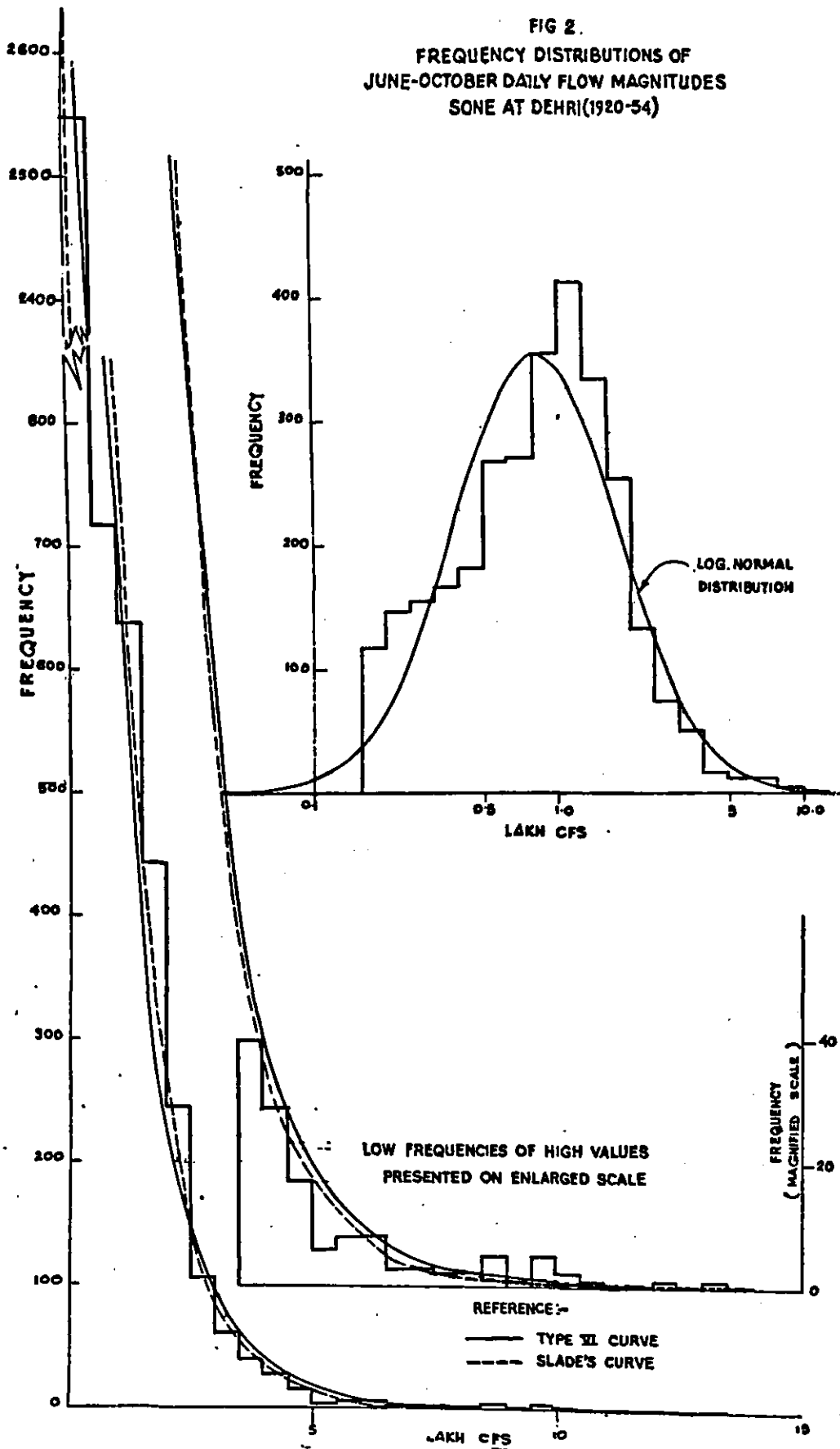


Fig 1 : Frequency histograms for July-Sept daily flows rates and logarithms

FIG 2.  
 FREQUENCY DISTRIBUTIONS OF  
 JUNE-OCTOBER DAILY FLOW MAGNITUDES  
 SONE AT DEHRI(1920-54)



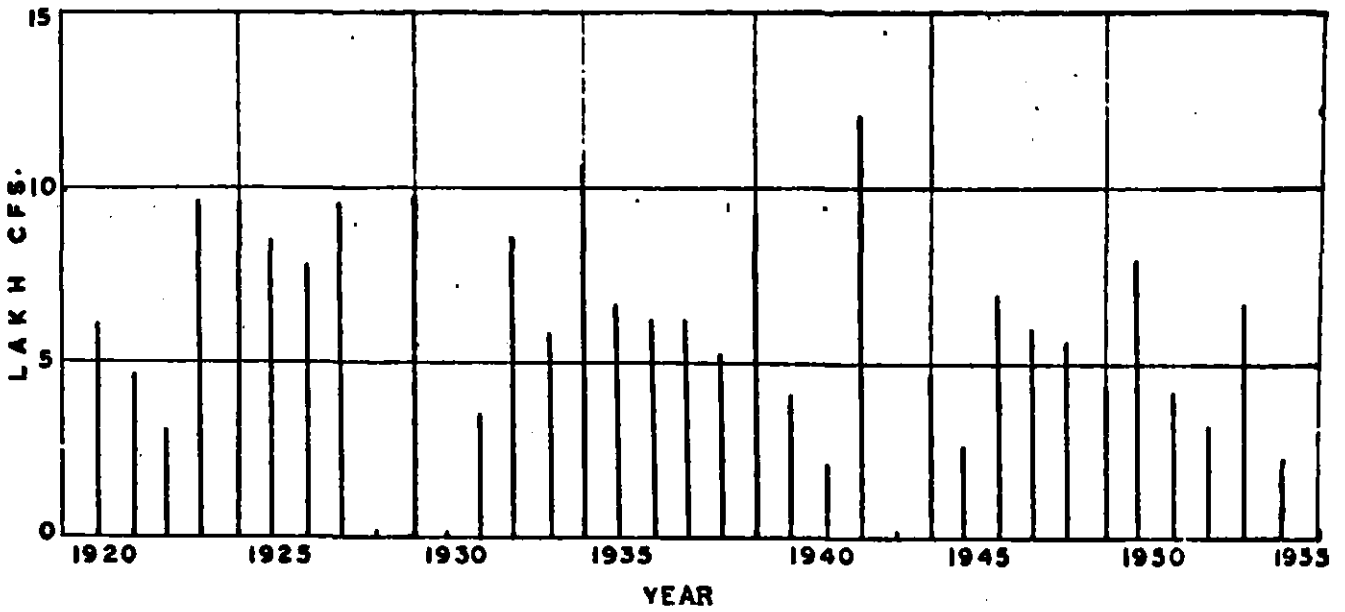


Fig 3 : Sone at Dehri (1920-54); annual highest flow rates

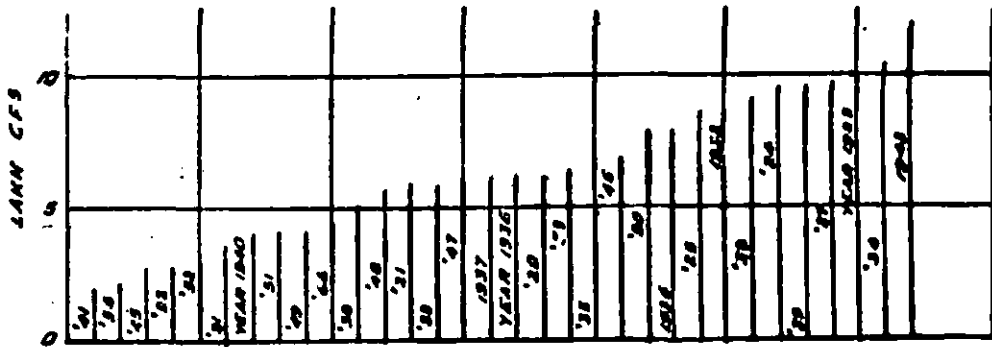


Fig 4 : Sone at Dehri (1920-54); annual highest flow rates (in ascending order)

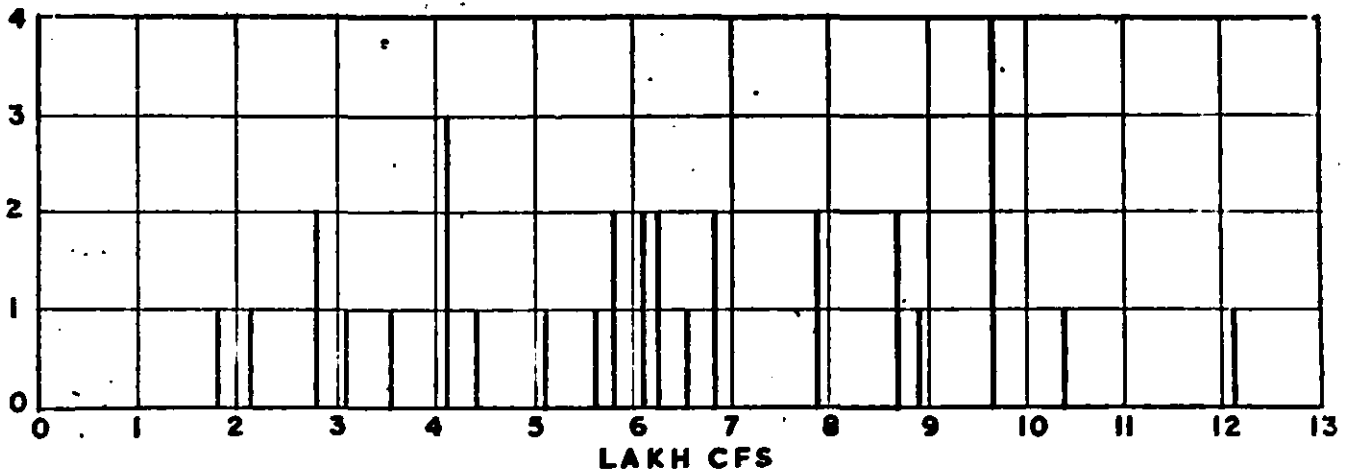


Fig.5: Sone at Dehri (1920-54); frequencies or numbers of years when annual highest flows attained different magnitudinal stages.

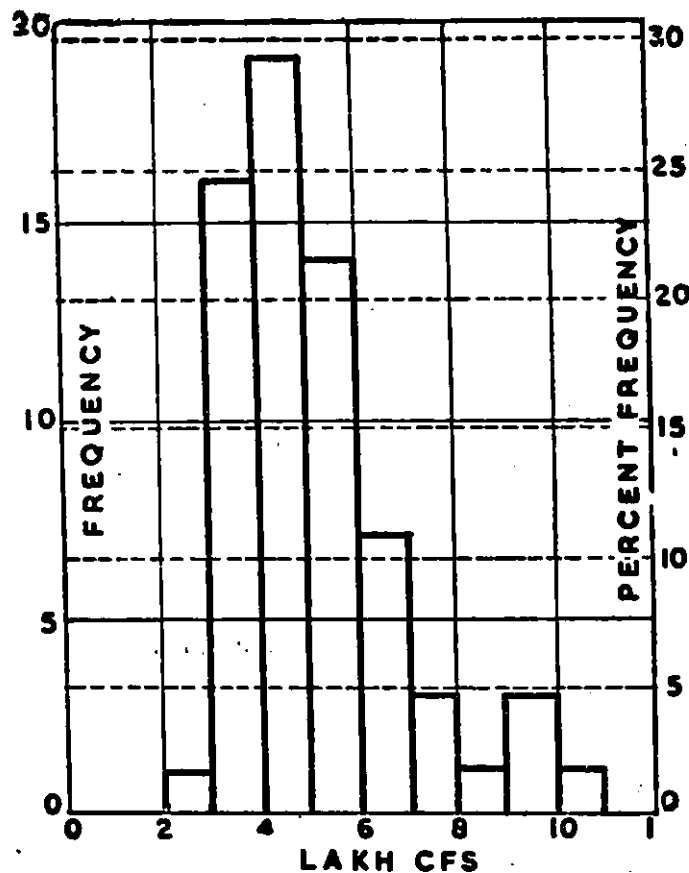


Fig 6 : Krishna at Vijayawada (1894-1958); frequency histogram of annual highest flow rates

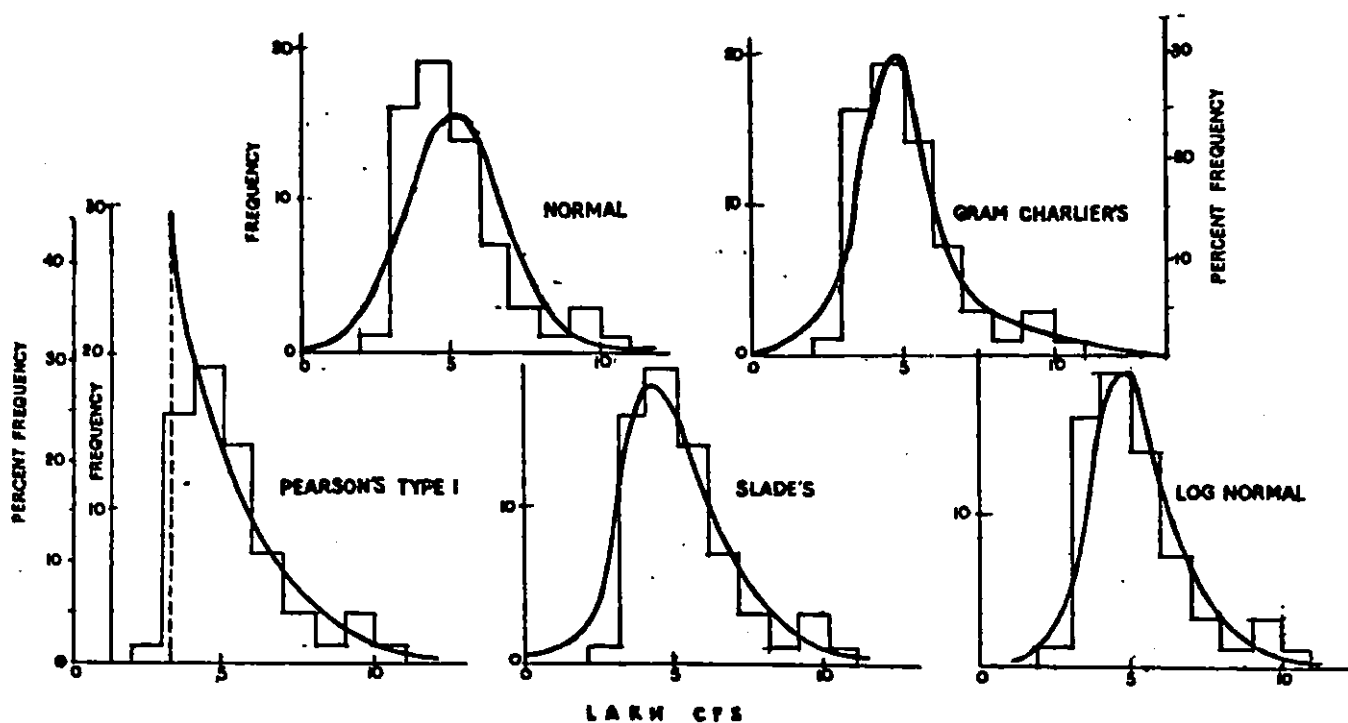


Fig. 7 : Krishna at Vijayawada (1894-1958); frequency histogram of annual highest flow rates with fittings of different smooth distributions

2.75	lakh cfs	were	observed	2	times	in	32	years
4.08	"	"	"	3	"	"	"	"
5.80	"	"	"	2	"	"	"	"
6.03	"	"	"	2	"	"	"	"
6.17	"	"	"	2	"	"	"	"
6.77	"	"	"	2	"	"	"	"
7.81	"	"	"	2	"	"	"	"
8.66	"	"	"	2	"	"	"	"
9.59	"	"	"	4	"	"	"	"

while all other magnitudes occurred only once each. Representing these relative frequencies of occurrence as the ordinates against the respective flood magnitudes as in fig 5 we find the lie of their tops yielding a frequency distribution polygon. It is not difficult to see that the unequal spacings of various ordinates in fig 5 owe primarily to the different concentrations obtaining over different ranges of the flow magnitudes during the period of the years. In order to highlight this aspect in its clearer perspective, a frequency histogram as in fig 6 offers better suitable.

Fig 6 actually shows the frequency histogram for a slightly longer data obtaining of the Krishna at Vijayawada for 65 years(1894-1958). The frequency histogram, as is well known, is the preliminary derivative stage of the ultimate frequency distribution type, which the data, when obtaining in larger numbers may be expected to subscribe or conform to. It is supposed to reflect the principal features of the type and yield a fair index to the values of the parameters contained of its descriptive expression. The broad extents of the curve's coverage to the left and to the right are also nearly indicated by the histogram, as also the relative frequencies expected of all intermediate individual annual peak flood magnitudes.

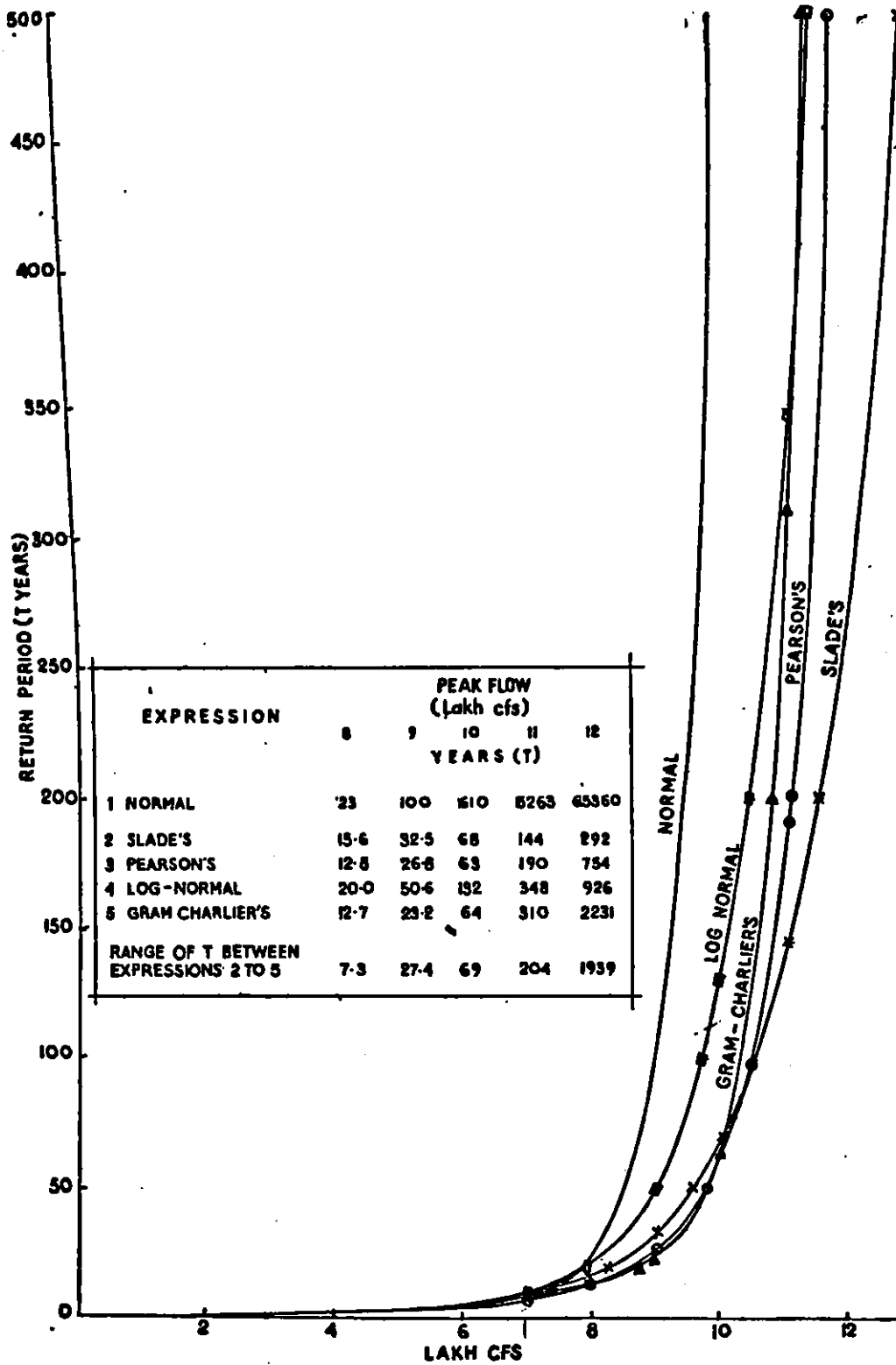


Fig 8 : Krishna at Vijayawada (1894-1958); expected return periods of peak flow magnitudes as derived by fitting different distribution types to data

## 6. MULTIPLICITY OF EMPIRICAL DISTRIBUTIONS

In an attempt to even out the irregularities of the limited histogram form of variations and to avail its inner average form for broad general use elsewhere, five curves evaluated in accordance with the standard procedures are shown superimposed in fig 7. As eye inspection alone may not provide sufficient or reliable evidence of the goodness of fit (appropriateness of the curve types to conform to the histogram pattern), aid has been sought of the non subjective statistical  $\chi^2$  tests. While the elegant symmetrical Gaussian or normal type is immediately found inappropriate for describing the skew pattern of the histogram, the multiplicity of the remaining four types all fitting the histogram nearly adequately, lands the investigation in an awkward dilemma. For, despite the four curves each providing adequate hugs to the overall lie of the histogram, they are not exactly coalescent. As an immediate corollary of the non-coalescence, the areas enclosed between the horizontal axis and the different curves to the right of any specified ordinate(s) will also differ. Since the reciprocals of these areas constitute the RETURN PERIODS (in years) of attaining or exceeding the corresponding peak flood magnitudes, the return period measures as estimated from the different curves will be also different (fig 8). The disparity feature between the estimates which is quite evident for flood magnitudes of the order of nine lakh cfs keeps on steadily increasing for higher magnitudes. It is easy to verify the converse feature also namely of the curves yielding non-unique estimates of return periods of occurrence for identical peak flood magnitudes. In other words, the use of the frequency curves despite their adequately fitting the frequency histogram of the observed annual



peak floods data seems rather inappropriate particularly for the purpose of estimating the peak magnitudes of longer return periods of occurrence. It is an inherent limitation of all statistical methodologies and/or derivations.

## 7. THEORETICAL DISTRIBUTION OF EXTREME VALUES

All the same, there are numerous occasions when such estimation problems have to be solved and the subject has been making rapid progress to provide reasonable approaches from within the realms of statistical methodologies. It was probably under these contingencies that Gumbel\*(1941) developed his Extreme Value distribution function

$$P(x) dx = ae^{-a(x-u)} e^{-e^{-a(x-u)}} dx \quad \dots (3)$$

for analysing the data consisting of the extreme (largest or smallest) members ( $x$ ) of several large samples (where  $a$  and  $u$  are parameters pertaining to the basin and  $e$  is the Napierian Constant = 2.71828).

Fisher\*\* and Tippett(1928) had derived the expression earlier in their exposition on "Limiting Forms of the Frequency Distribution of the Largest or Smallest Member of a Sample".

The limiting form(3) is known to hold for most forms of frequency distributions of the parent population of  $x$  so long as the sample size  $m$  is sufficiently large. In the present case  $m$  equals 365 corresponding to the number of days in a year. The value of  $n$  for the number of years whose peak flow data have been observed is also similarly required to be sufficiently large for obtaining better reliable estimates of  $a$  and  $u$  in(3). Assuming for simplicity the values of  $m$  and  $n$  obtaining

---

\* E J Gumbel : 'The Return Period of Flood Flows', Annals of Mathematical Statistics, Vol XII pp 163-190, 1941

\*\* R A Fisher & L H C Tippett : Proceedings of the Cambridge Philosophical Society, Vol XXIV, Pt 2, pp 180-190, 1928.

in most hydrological data to be tolerably adequate in practice, different processes have been availed for evaluating  $\underline{a}$  and  $\underline{u}$  in(3). On integrating (3) the probability of an year's peak flow  $\underline{x}$  not exceeding  $\underline{x}_0$  is obtained as

$$P_0 = e^{-e^{-a(x_0 - u)}} \quad \dots (4)$$

#### 8. HYPOTHETICAL PLOTTING POSITIONS

In terms of the average return period measure T(in years) of just exceeding  $\underline{x}$  once in T years(4) can be re-expressed as

$$x = u - (1/a) \log_e \log_e T/(T - 1) \quad \dots (5)$$

While there is no dispute up to this stage from a theoretical point of view, different methods have been availed on different occasions for evaluating the two parameters  $\underline{a}$  and  $\underline{u}$  from a set of any observed data. Thus, Gumbel offered the method of equating the first two moments or alternatively their variants, the mode and the mean deviation to the respective entities of distribution(3), which is a perfectly legitimate procedure. Some other procedures, apparently presumed to be capable of yielding even more powerful estimates of  $\underline{a}$  and  $\underline{u}$  by directly fitting expression(5) to the data by the method of least squares have also been availed. For this, the different approximate concepts of the return period measure as postulated by Hazen, Gumbel and V T Chow have been availed for evaluating the return period measures of the observed n x's. It is here possibly that the studies have landed in a region of fallacious approach. Gumbel\* is, no doubt, aware of the possible incorrectness of the method of the plotting position's definition, but has still availed it as the most plausible method under the circumstances

---

\*E J Gumbel : Statistics of Extremes, pp 29-37, Columbia University Press, New York(1958)

of the case. The method consists of arranging the  $n$  years' observed annual floods data in descending order as  $x_1 \dots x_n$  and assigning their respective return periods as

$$\begin{array}{cccccccc}
 & T_1 & T_2 & T_3 & T_4 \dots & T_m \dots & T_n & \\
 \text{Hazen} & \frac{2n}{1} & \frac{2n}{3} & \frac{2n}{5} & \frac{2n}{7} \dots & \frac{2n}{2m-1} \dots & \frac{2n}{2n-1} & \\
 \text{Gumbel and} & \frac{n+1}{1} & \frac{n+1}{2} & \frac{n+1}{3} & \frac{n+1}{4} \dots & \frac{n+1}{m} \dots & \frac{n+1}{n} & \\
 \text{V T Chow} & & & & & & & 
 \end{array} \left. \vphantom{\begin{array}{cccccccc} T_1 & T_2 & T_3 & T_4 \dots & T_m \dots & T_n & \end{array}} \right\} \dots (6)$$

Taking  $n = 50$  for instance, these magnitudes obtain as

$$\begin{array}{cccccccc}
 & T_1 & T_2 & T_3 & T_4 \dots & T_{50} & & \\
 \text{Hazen} & 100 & 33.3 & 20 & 14.3 \dots & 1.01 & & \\
 \text{Gumbel and} & 51 & 25.5 & 17 & 12.75 \dots & 1.02 & & \\
 \text{V T Chow} & & & & & & & 
 \end{array} \left. \vphantom{\begin{array}{cccccccc} T_1 & T_2 & T_3 & T_4 \dots & T_{50} & & \end{array}} \right\} \dots (7)$$

On plotting the  $n$  points  $(T_m, x_m)$  so obtaining on specially ruled Extreme Value Probability paper, a least square line through them is drawn and the values of the two parameters  $a$  and  $u$  determined. Thus different values will be obtained according as Hazen's or the other series of  $T$  values are availed, resulting in higher return period measure for any specific large flood magnitude by Hazen's process than by the other.

Actually neither sets of the apriori  $T$  values in(6) nor any other similar sets can claim to constitute the true average return period measure of the  $x$ 's. In fact, they do not debar other definitions being formulated for evaluating  $T$  in different other ways\*. Thus, the design flood estimates derived there-from by fitting the least square lines through the  $n$  points  $(T_m, x_m)$  will also differ accordingly. Even

---

\*B F Kimball - "On the choice of plotting Positions on Probability Paper" - Journal of Am Stat Assoc Vol 55 No 291, Sept '60

R K Linsley, M A Kohler and J L H Paulhus - "Hydrology for Engineers", pages 247 - 249, McGraw Hill, New York, 1958.

when the different sets of the plotting positions may be covering much common ground in the domain of the observed data, their average lines may not be exactly coalescent. Thus on account of any such incidental slope differences, the different mean lines through the different sets may find themselves considerably apart from each other, especially when extended outside the observed domain of the data.

#### 9. MAXIMUM LIKELIHOOD ESTIMATION

The feature only reflects the inherent limitations of analysing small samples data. For, the set of the  $n$  years' observed data form but a small part only of the unending series of the annual floods' data. The small part is thus fortuitously quite capable of coinciding with a series of very low flood years or of very-high-flood years or of a balanced or other mixture of the two. In fact no law, cyclic or otherwise of any assignable number of years duration of periodic regularity or otherwise has so far been ascertained from observed data or inferable other evidence. The utmost that is known of the observed series of the annual floods data is that each separate value comprises the largest or the 'extreme member' of the yearly samples of 365 daily flood measures each. Such data are yet known to subscribe to the only Distribution Law(3):

Availing therefore(3) alone as our main prop we can evaluate  $a$  and  $u$  by maximising the probability of the observed  $n$  floods. IN OTHER WORDS, WE ONLY ASK WHAT VALUES OF  $a$  AND  $u$  WOULD MAKE THE CHANCE A MAXIMUM OF OUR HAVING OBSERVED THE MAGNITUDES  $x_1, \dots, x_n$  AS  $n$  ANNUAL HIGHEST FLOODS? For this, we do not need to postulate any subjective concepts for the return period measures of the observed  $n$   $x$ 's. Since no sequential law of the annual floods has been observed or is expected,

the individual peak flood values all must obviously be having respective return period measures quite independently of each other. Thus, the return period measures of the  $n$  observed peaks may be all larger than  $n$  years, all smaller than  $n$  years, or mixed. But each peak is an annual peak value, and must all have occurred under the most congenial circumstances inducing their occurrence in the different respective years. In other words, the probability of the  $n$  magnitudes having occurred as  $n$  annual highest floods must be a maximum. Now by the multiplication theorem for the joint occurrence of  $n$  independent events

$$\begin{aligned} P(x_1 \dots x_n) &= P(x_1) P(x_2) \dots P(x_n), \text{ which with the help of (3)} \\ &= a^n e^{-\sum a(x_i - u)} e^{-\sum e^{-a(x_i - u)}} \dots (8) \end{aligned}$$

In order to maximise (8) which is the same thing as maximising

$$\begin{aligned} L &= \log P(x_1 \dots x_n) \\ &= n \log a - \sum a(x_i - u) - \sum e^{-a(x_i - u)} \end{aligned}$$

$$\text{we put } \frac{dL}{du} = 0 \quad \text{and} \quad \frac{dL}{da} = 0 \quad \dots (9)$$

$$\begin{aligned} \text{Now } \frac{dL}{du} &= na - a \sum e^{-a(x_i - u)} \\ &= na - ae^{au} \sum e^{-ax_i} \end{aligned}$$

on equating which to zero, we obtain

$$e^{au} \sum e^{-ax_i} = n \quad \text{ie} \quad e^{-au} = \frac{1}{n} \sum e^{-ax_i} \dots (10)$$

$$\text{Next, } \frac{dL}{da} = \frac{n}{a} - \sum (x_i - u) + \sum (x_i - u) e^{-a(x_i - u)} \dots (11)$$

Substituting from (10) in the last term of (11) and rewriting,

$$\begin{aligned} \frac{dL}{da} &= \frac{n}{a} - n\bar{x} + nu + e^{au} \sum x_i e^{-ax_i} - nu \\ &= \frac{n}{a} - n\bar{x} + e^{au} \sum x_i e^{-ax_i} \dots (12) \end{aligned}$$

Utilising (10) for a second time in (12) and transposing, it can be rewritten as

$$\frac{dL}{da} = \frac{n}{\sum e^{-ax_i}} \sum x_i e^{-ax_i} - n(\bar{x} - 1/a)$$

on equating  $\frac{dL}{da}$  to zero therefore, we obtain

$$f(a) \equiv \sum x_i e^{-ax_i} - (x - 1/a) \sum e^{-ax_i} = 0 \quad \dots (13)$$

yielding a unique value for  $\underline{a}$ , simply deduced from the  $\underline{n}$  observed magnitudes of the annual peak floods  $x_1 \dots x_n$  without any hypothetical assumptions whatsoever about their return periods. The requisite value of  $\underline{u}$  also is then uniquely determinable from (10).

#### 10. SCHEMATISATION OF COMPUTATIONS

But the form of expression (13) despite its containing a single unknown  $\underline{a}$  is rather complicated to yield its unique solution explicitly. It does not offer any further simplification of form. The neighbourhood of the expected solution therefore, has to be intuitively assessed first, whereafter a series of nearer approximations to the exact solution of  $\underline{a}$  have to be successively tried for bringing  $f(a)$  nearer to zero. The amount of numerical work involved which has been called as "prohibitive" by Gumbel is no doubt very cumbersome. But the obtaining results very often more than repay for all the trouble and labour by the sheer virtue of the absolute freedom from any subjective or hypothetical assumptions accompanying the estimation procedures.

The methodical steps described below help reduce the tedium. Making an intelligent start thus from

$$a = a_1 = \frac{\pi}{\sqrt{6}} \frac{1}{s} \quad \text{where } s = \text{standard deviation of the } x\text{'s,}$$

the solution obtained by the method of moments (mentioned above) as the first approximate solution, the value of  $f(a_1) \neq 0$  is computed. Then, utilising Taylor's expansion for

$$f(a_2) = f(a_1 + h_1) = 0$$

upto the first power of  $h_1$ , we obtain  $h_1 = -\frac{f(a_1)}{f'(a_1)} \quad \dots (14)$

where  $f'(a) = \frac{d}{da} f(a) = -\sum x_i^2 e^{-ax_i} + (\bar{x} - 1/a) \sum x_i e^{-ax_i} - \sum e^{-ax_i}/a^2$ .

But the approximation may not suffice and most likely,  $f(a_2)$  either may not be actually equal to zero. We again repeat the steps and determine successively the necessary values of  $h_2, h_3$  etc until the requisite smallness of value of  $f(a_k)$  is reached. In most cases a value of  $k = 3$  or 4 will more than suffice in practice for yielding a negligible  $f(a_k)$ . An illustrative example is given below to help clarify the computational steps better. The example of the Sabarmati at Dharoi(1935-52) is of course especially chosen for the simple reason of its small size data to suit the presentation of the detailed procedures in print. The routines of the computational procedure otherwise, will be exactly similar for rivers with longer years data which no doubt are always definitely more desirable for reliable design flood estimation.  $n$  should in general be preferably larger than about 20.

The preliminary computations upto

$$a_1 = .048045$$

are all shown and are simple and self explanatory. The steps which are found useful for reaching the successive approximation stages are also presented in suitably schematised form for adoption for similar computations elsewhere with other data. Thus starting from

$$f(a_1) = 6.338863 \quad \text{for} \quad a_1 = .048045$$

we soon come down to

$$f(a_4) = .001448 \quad \text{for} \quad a_4 = .049837$$

The successive reductions in the values of  $h$  are largely found useful in guiding where to call a halt to the process of bringing  $f(a)$  nearer to zero.

Computing the value of  $u_k$  from (10) at the stage reached of  $a_k$  we have for the estimation expression of the expected annual peak flood magnitudes of any long (or short) term return period (T) of occurrences in the Sabarmati at Dharoi as .

$$\text{for } k = 1 \quad x_T = 22.982208 - \frac{1}{.048045} \log_e \log_e \frac{T}{T-1}$$

$$\text{for } k = 2 \quad x_T = 22.704838 - \frac{1}{.049737} \log_e \log_e \frac{T}{T-1}$$

$$\text{for } k = 3 \quad x_T = 22.687887 - \frac{1}{.049835} \log_e \log_e \frac{T}{T-1}$$

$$\text{for } k = 4 \quad x_T = 22.687353 - \frac{1}{.049837} \log_e \log_e \frac{T}{T-1}$$

It is not difficult to see the fair stability attained of the estimation expression at the third approximation stage of  $k = 3$ . It may be useful to mention that a fast computer with the help of an ordinary calculating machine and a book of logarithmic tables at hand, should not take more than 6 to 7 hours of a working day to reach the kind of stability even for observed annual peak floods data on record extending upto 50 years. The long term peak flood magnitude derived from the four expressions for different values of T are also tabulated. The stability obtaining of the expected peak flood magnitudes of different return periods is abundantly evident for  $k = 2$  even, as the incremental improvements accruing to the estimates thereafter do not add to more than .2% in the present example at least. Another remarkable feature of the method is that the percentage improvements to the estimates at every successive approximation stages remain nearly comparable for all Ts from 20 to about 500 years and possibly beyond.



*Computational Procedure with data of Sabarmati at Dharoi (1935-52)*

Year	Peak flow rate (000 cfs)	$x^2$	$e^{-a_1x}$	$e^{-a_2x}$	$e^{-a_3x}$	$e^{-a_4x}$
1935	27	729	.2733	.2611	.2604	.2604
1936	1	1	.9530	.9515	.9515	.9515
1937	44	1936	.1208	.1121	.1116	.1116
1938	28	784	.2605	.2484	.2477	.2477
1939	80	6400	.0214	.0187	.0186	.0186
1940	32	1024	.2149	.2036	.2030	.2030
1941	15	225	.4864	.4742	.4736	.4735
1942	44	1936	.1208	.1121	.1116	.1116
1943	68	4624	.0381	.0340	.0338	.0338
1944	87	7569	.0153	.0132	.0131	.0131
1945	66	4356	.0420	.0375	.0373	.0373
1946	16	256	.4635	.4512	.4505	.4505
1947	20	400	.3825	.3698	.3691	.3691
1948	7	49	.7144	.7060	.7055	.7055
1949	17	289	.4419	.4293	.4286	.4286
1951	36	1296	.1773	.1669	.1663	.1663
1952	2	4	.9084	.9053	.9051	.9051
Total	590	31878	5.6345	5.4949	5.4873	5.4872

$$\frac{\Sigma(x - \bar{x})^2}{n-1} = \frac{\Sigma x^2 - (\Sigma x)^2/n}{n-1}$$

$$\frac{31878 - 20476.4706}{16}$$

$$712.595588$$

$$s = 26.694486$$

$$a_1 = \frac{\pi}{\sqrt{6}} \times \frac{1}{s} = .048045 \text{ (1st approx to } a)$$

$$\bar{x} = \frac{\Sigma x}{n} = \frac{590}{17} = 34.7059$$

*Computations for further values of  $a_{k+1}$*

Step	$k = 1$	2	4	5
1. $a_k \log_{10} e$	.020866	.021601	.021643	.021644
2. $1/a_k$	20.8138	20.1058	20.0662	20.0654
3. $1/a_k^2$	433.2143	404.2432	402.6524	
4. $\bar{x} - 1/a_k$	13.8921	14.6001	14.6397	14.6405
5. $\Sigma e^{-a_k x_i}$	5.6345	5.4949	5.4873	5.4872
6. $\Sigma x_i e^{-a_k x_i}$	84.6139	80.5551	80.3383	80.3368
7. $\Sigma x_i^2 e^{-a_k x_i}$	2481.3331	2317.4409	2309.0391	
8. $f(a_k) = (6) - (4) \times (5)$	6.338863	.329011	.005874	.001448
9. $f'(a_k) = -(7) + (4) \times (6) - (3) \times (5)$	-3746.8143	-3362.6043	-3342.3850	
10. $h_k = \frac{f(a_k)}{f'(a_k)}$	.001692	.000098	.000002	
11. $a_k + h_k$	$a_2 = .049737$	$a_3 = .049835$	$a_4 = .049837$	

*Computations for values of  $u_k$*

From (10)

$$u_k = \log_e 10 \times \frac{1}{a_k} \times \log_{10} (n) - \log_e 10 \times \frac{1}{a_k} \times \log_{10} (\Sigma e^{-a_k x_i})$$

$$= \quad \quad \quad A \quad \quad - \quad \quad B$$

	$k = 1$	2	3	4
$A$	58.967577	56.961742	56.849551	56.847285
$\log_{10} (\Sigma e^{-a_k x_i})$	.750855	.739960	.739359	.739351
$B$	35.985369	34.256904	34.161664	34.159932
$u_k = A - B$	22.982208	22.704838	22.687887	22.687353

*Respective improvements from successive approximations to Peak flow estimates of  $T$  years return period ( $x_T$ ) from*

$$x_T = u_k - \frac{1}{a_k} \log_e \log_e \frac{T}{T-1}$$

$x_T$  (000 cfs) for

$k$	$x_T$ (000 cfs) for									
	$T = 20$	% im- provements	$T = 50$	% im- provements	$T = 100$	% im- provements	$T = 200$	% im- provements	$T = 500$	% im- provements
1	84.80		104.20		118.73		133.21		152.31	
		-2.807		-2.917		-2.982		-3.025		-3.066
2	82.42		101.16		115.19		129.18		147.64	
		- .158		- .168		- .165		- .178		- .186
3	82.29		100.99		115.00		128.95		147.37	
		0		- .010		- .009		0		0
4	82.29		100.98		114.99		128.95		147.37	

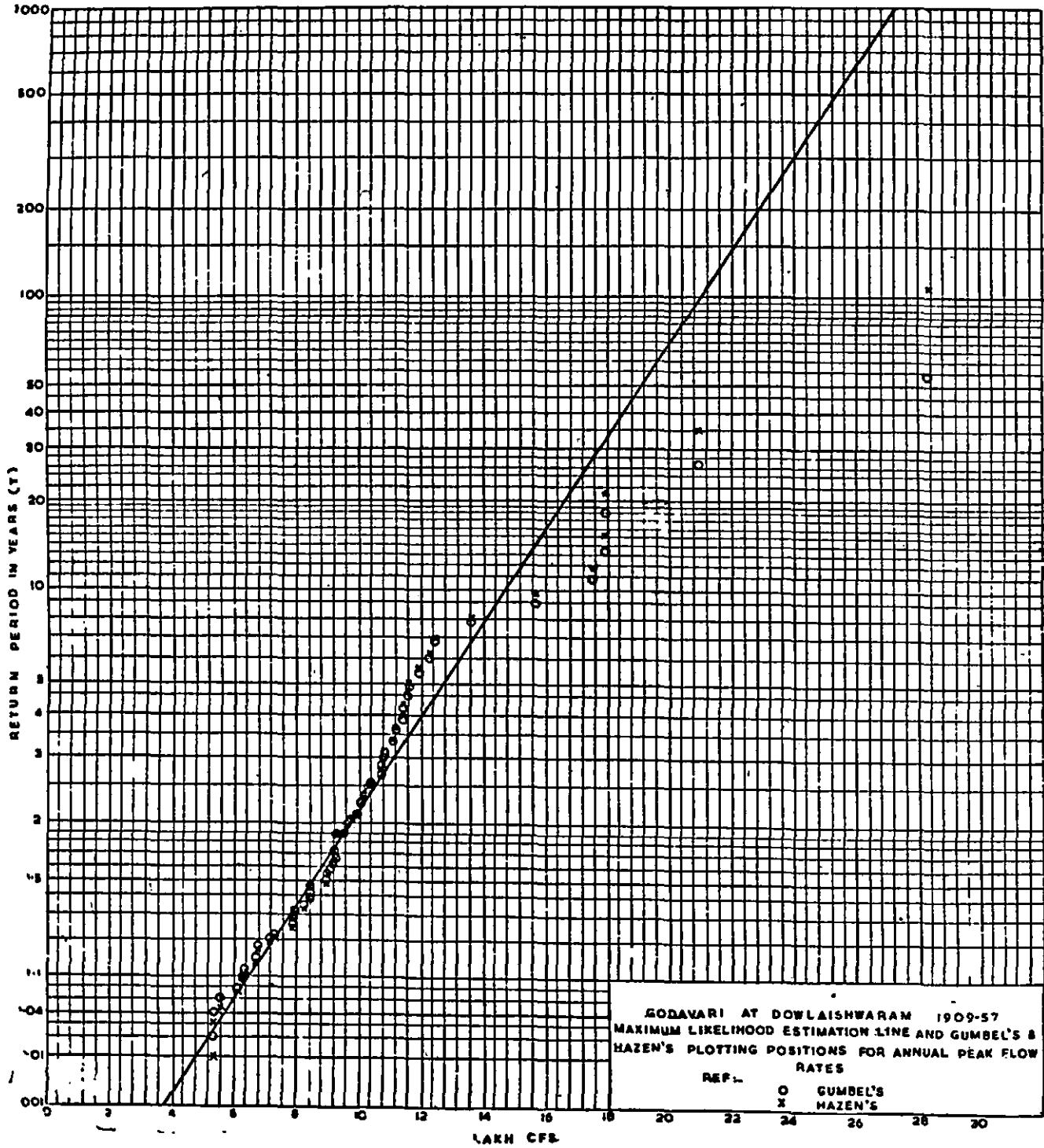


Fig 9

## 11. EXTREME VALUE PROBABILITY PAPER

The expressions can be conversely also utilised for estimating uniquely the average return period of occurrences of any big or small annual peak flood magnitudes. In fact, they can be presented graphically by means of straight lines on the specially ruled Extreme Value probability paper as in fig 9. The magnitudes of the expected peak flow rates can be then directly read off for any desired return period value from the corresponding points along the line, or conversely the expected return periods for any postulated peak flow magnitudes. Any other hypothetical plotting positions for the observed annual peak flow data of the past thus have to be termed as purely tendentious or otherwise fallacious. But the other two methods mentioned in passing earlier, utilised by Gumbel namely of equating the first two moments or their variants the mode and mean deviation of the observed  $n$  annual peak flow data to the corresponding entities of (3) for determining the values of  $a$  and  $u$  also possess the likewise advantages of freedom from any subjective bias or limitations of arbitrarily defining the plotting positions. They also yield similar straight lines for reading off the expected peak flow magnitudes equally uniquely for any return periods magnitudes and conversely.

The question, therefore, arises immediately: Which of the three methods yields the better efficient estimates? In particular, where lies the superiority of the method just described with all its entailing huge computational labour, unless the obtaining estimates are better precise? The expression for computing the standard errors of the estimates by the two methods previously described by Gumbel has also

been given by him as

$$SE(x_T) = \left[ \frac{e^{a(x_T - u)}}{a\sqrt{n}} \right] \left[ e^{-a(x_T - u)} - 1 \right]^{\frac{1}{2}}$$

Appendix I shows the derivation of the expression obtaining for the standard error of the estimates utilising the values of  $a$  and  $u$  as deduced by the method of maximum likelihood. It is

$$SE(x_T) = \sigma = \frac{1}{a\sqrt{n}} \left[ 1 + \frac{6}{\pi^2} (1 - C - \log_e \log_e \frac{T}{T-1}) \right]^{\frac{1}{2}} \dots (15)$$

While the forms of above expressions by themselves do not offer the ready means of comparing their relative magnitudes, Table I reproduces the expected annual peak flood magnitudes of various long term return periods of occurrences as deduced by the different methods together with their respective standard error measures for different rivers observed data. They provide the abundant incontrovertible evidence thus obtaining of the greater precise measures invariably accompanying the estimates by the method of maximum likelihood over the others.

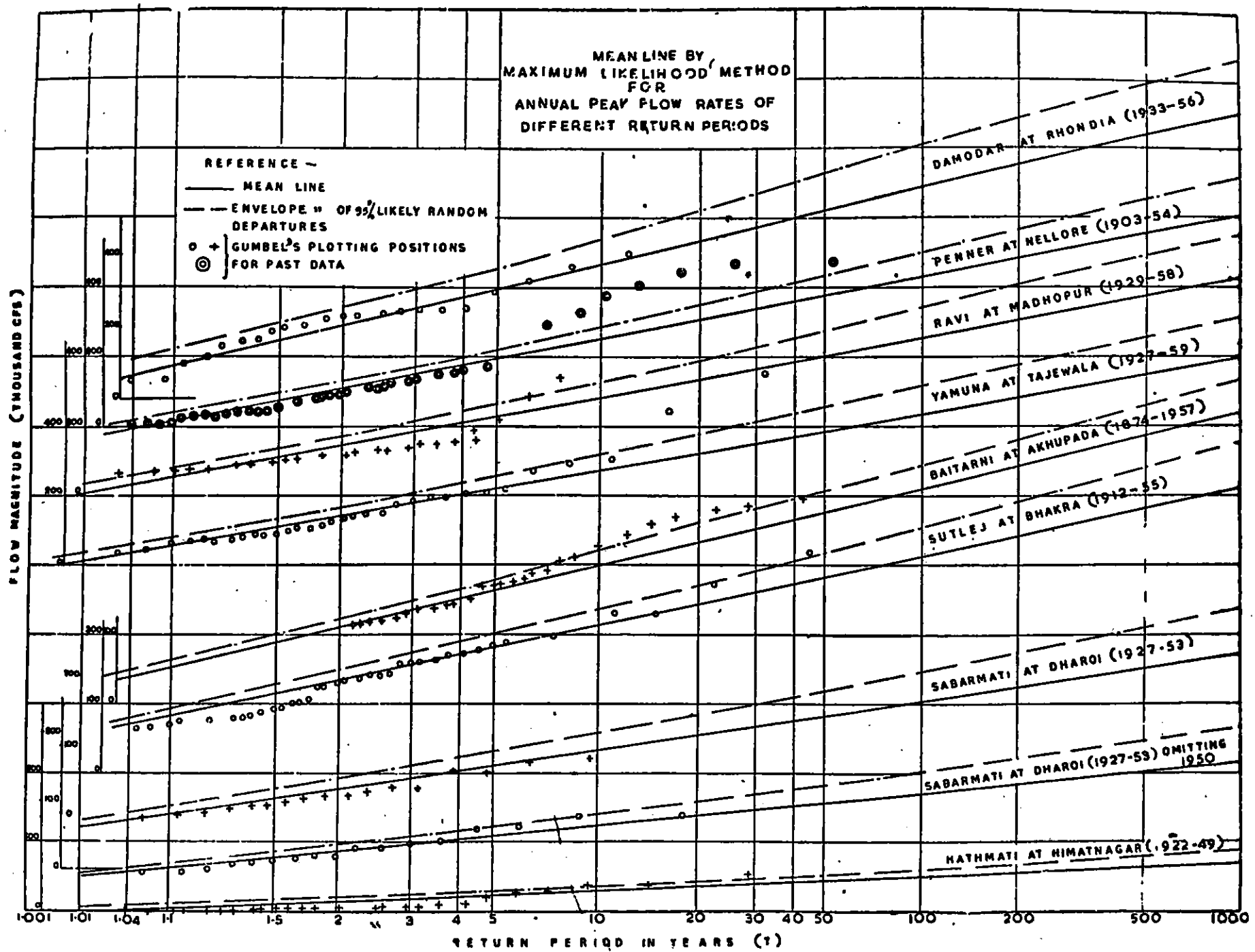
## 12. STABILITY OF PEAK FLOW ESTIMATES

The Table also reproduces the magnitudes of the corresponding peak floods as deduced by availing the alternative methods utilising the plotting positions concept. While their comparison with the estimates otherwise deduced by the method of maximum likelihood may be irrelevant, fig 9 shows the different degrees of their conformities between the estimates by the various methods which again may differ fortuitously from one river to another. Whenever they may be found to conform closely to the estimates by the method of maximum likelihood also they do not provide any firm evidence about the correctness of the plotting positions method(s). For, casual additions of new data or deletions of a few doubtful observations are likely to leave their impress very

differently on the two methods particularly when such data correspond to very high flood years. Thus the estimates values (Table I) by the various methods by including and excluding alternatively one or two high years data show the relative greater inertness (smaller susceptibility in other words) of the method of maximum likelihood to such omissions or additions to the main bulk of the data. The greater susceptibility of the plotting positions method in fact springs from the greater importance the least square process for formulating the line, assigns to the positions of the end points. In other words, the fitted line is more inclined to follow their lie. Thus any errors, big or small, in the plotting positions of the very high observed annual peak flows are likely to be duly reflected in the consequent derived estimates of long term return periods. With the best intentions therefore, for deriving precise estimates of the expected long term peak flow magnitudes, systematic errors of unknown magnitudes may be most unknowingly creeping into them. An illustration will perhaps clarify better.

In the 27 years' (1926-53) annual highest observed peak flows magnitudes data on record for the Mahanadi at Sambalpur, the following four years with the respective flow magnitudes shown against each stand out as the first to the fourth in descending order.

Year	(000 cfs)	<u>Return period (T years) estimated by method of</u>				
		Hazen	Gumbel	Maximum Likelihood	Moments	Mode and Mean deviation
1947	965	54	28	9.3	14.5	11.2
1946	950	18	14	8.4	12.8	10.1
1948	948	10.8	9.3	8.3	12.6	10.0
1952	848	7.7	7	4.6	5.8	5.1



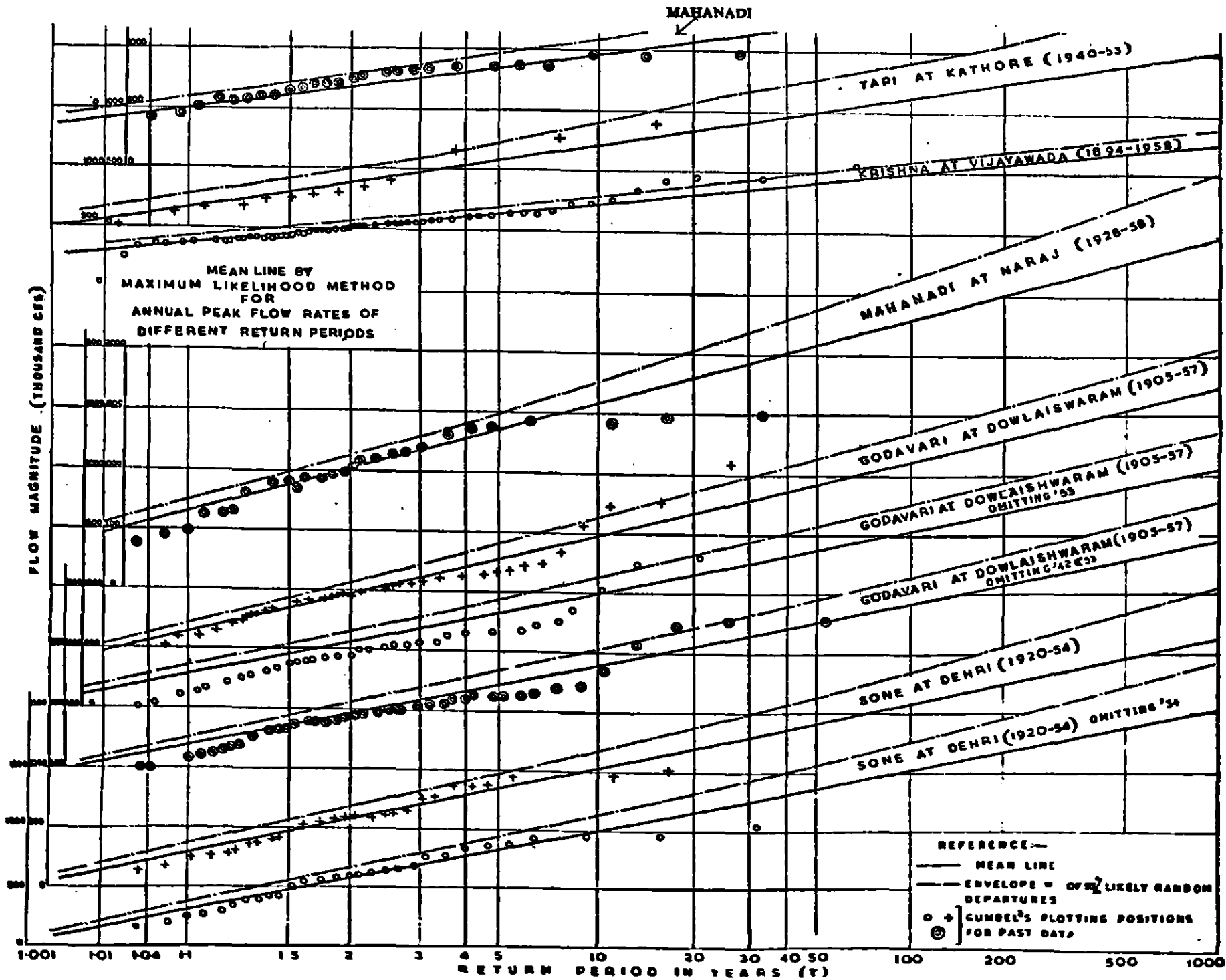


Fig 10



It is doubtful if the three highest magnitudes observed can be at all termed as different one from the other. Even currentmeter observations of the flow velocities are ordinarily qualified as subjected to  $\pm 3\%$  errors. But if the annual observed flow magnitudes are each to be assigned definite ranks and hypothetical return period measures dependent on them, the three nearly equal magnitudes above of the peak flows will each have vastly different values of T. Prima facie they will appear very unrealistic. Unlike the adhoc hypothetically assigned T values, their most likely measures as deduced from the 27 years observed annual peak flows data themselves by the method of maximum likelihood are all very rationally close to each other as they should be. Thus it is reasonable to hope that any derived peak flow estimates of long term return periods will be also better realistic (conforming to actualities) by the method of maximum likelihood than by the estimates derived by utilising the hypothetical plotting positions concept.

### 13. UPPER CONFIDENCE LIMITS

Having reached ~~this~~ to the happy solution of an otherwise so far tough problem it may be worth pondering for a while over the pithy note of R A Fisher, the greatest of all Statistical wizards cited at the outset. The paper describes a precise way for evaluating uniquely the expected estimates of long term return periods of occurrences in any stream(s) but appends an expression(15) yet for computing the standard error measures. This has been perhaps done rightly also. For, who can correctly foresee the infinite future which is yet to unfold? A peep into the recorded limited past can at best provide a semblance only toward any approximate guidance. Any mathematical law that may be nearly discovered to follow the course of events will be an approximation and

may not fully coalesce with all the occurrences exactly. The standard error device accordingly computed for any formula fitting the past data thus provides a yardstick measure of their closeness to the actual occurrences in the past and plausibly of their expected deviations in the future, so far as the other causes, interfering with their course, purely arise from the chance or such other/random phenomena. Thus, despite the uniqueness of the average estimates obtaining from (5) with the values substituted of  $\underline{a}$  and  $\underline{u}$  as derived by the method of maximum likelihood from the limited observed data of the past, any like estimates deduced from another set of  $\underline{n}$  years observed data may be slightly different also, the differences ordinarily not exceeding  $1.64 \sigma$  where the standard error  $\sigma$  is given by (15).

In order to provide reasonable safeguards therefore, for meeting 95% of such chance variations it is usual ordinarily to take the upper limiting value of  $x_T + 1.64 \sigma$  for all design flow estimates. The straight lines representing the expressions for computing the average peak flow magnitudes of different return periods as deduced by the method of maximum likelihood for various observed rivers data together with the lines for the corresponding 95% upper limiting values are shown in figs 10. [The original data of the various rivers are presented separately in Appendix II and fig 11]. The hypothetical return period values according to Gumbel are also plotted therein and conform to lines which may at times vastly differ from the lines fitted by the method of maximum likelihood. Thus it is not always safe to accept their line as reasonably accurate particularly when long term design flow estimates are desired.

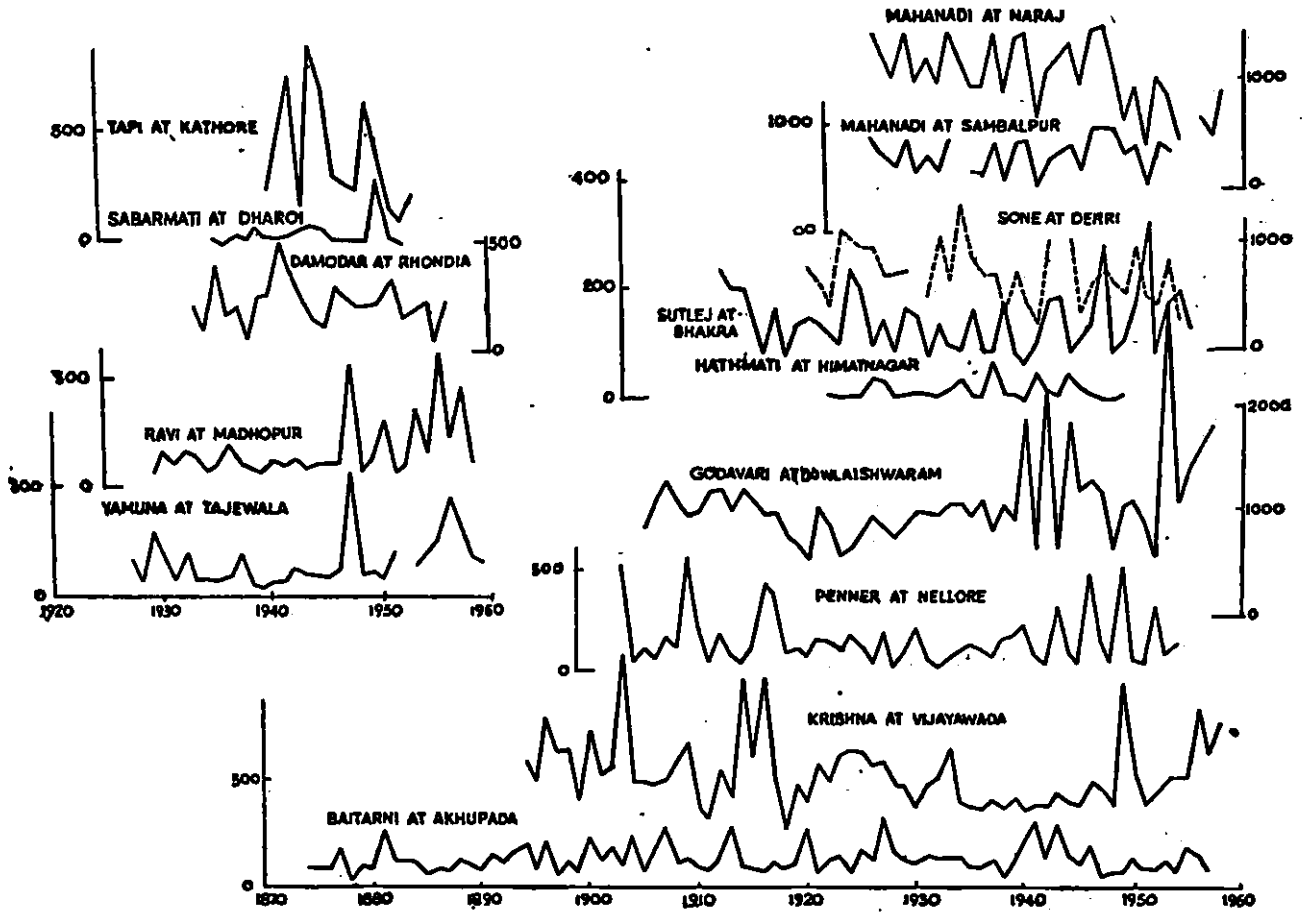


Fig 11 : Annual highest flow rates

#### 14. RESUME

When data of the directly or indirectly observed annual maximum flow measures for a reasonably sufficient number of years are available, the method of maximum likelihood has been shown to be better successful for deriving both precise as well as consistent estimates of the design flow measures of desired long term return periods. At the outset, the virtual impossibility has been postulated of framing any apriori definitions for evaluating the plausible return period measures of the different observed annual peak flow rates data on record. Instead, the method of maximum likelihood, with no hypothetical assumptions whatsoever, has been called to aid for evaluating the parameters of the extreme value distribution law governing the data. Any design flow estimates so derived can all thus claim absolute freedom from subjective biases in the analytical procedures etc.

The increased precision of the estimates as well as their larger insusceptibility to withstand effectively small additions or deletions of doubtfully high values or such other material to and from the basic data have been demonstrated.

A schematized method is presented of speedily and systematically sieving through the complicated computations for arriving at the estimation expression for the average peak flow magnitude of different return periods. The procedure outlined rids it of the tedium usually attendant of the computations in all such estimation problems, and thus commends as readily acceptable to most design engineers equipped only with routine calculating machines and ordinary tables of logarithms.

#### ACKNOWLEDGEMENTS

The preparation of the Technical Memorandum in its above form was undertaken primarily at the instance of Dr K'L Rao for whose almost challenging encouragement the authors remain highly indebted.

The authors owe grateful acknowledgements to Dr A R Kamat, S C Desai, C V Gole, and S V Chitale for kindly going through the paper critically. V T Jumde and M V Bagade carried out most ungrudgingly the large volume of computational work necessary for the presentation of the results.

TABLE 1

Long Term Estimates ( 000 cfs ) of Peak Flows

Period and site	Return Period ( years )	Deduced with Parameters using							
		Maximum likelihood method		Moments		Mode and mean deviation		Plotting position concept due to	
		3	4	5	6	7			
1. Baitarni at Akhupada ( 1874-1957 )	20	237 ± 14	248 ± 25	246 ± 25	257	249			
	50	280 18	295 40	293 39	306	295			
	100	313 21	331 56	328 55	343	331			
	200	345 23	368 80	365 78	383	368			
	500	388 27	411 126	408 123	429	411			
2. Krishna at Vijayawada ( 1894-1958 )	20	795 41	826 70	821 69	848	826			
	50	903 50	943 111	936 109	973	943			
	100	983 58	1031 157	1022 154	1066	1031			
	200	1064 66	1125 223	1114 218	1166	1125			
	500	1170 76	1232 352	1219 344	1281	1233			
3. Godavari at Dowlaishwaram ( 1905-57 )	20	1673 104	1812 197	1722 169	1849	1797			
	50	1922 129	2108 311	1976 267	2156	2087			
	100	2108 148	2330 439	2167 378	2386	2304			
	200	2293 167	2567 621	2371 535	2631	2535			
	500	2538 193	2838 983	2604 846	2912	2800			
4. - do - ( omitting 1953 )	20	1597 96	1625 159	1608 148	1675	1625			
	50	1825 120	1863 251	1828 233	1928	1862			
	100	1997 137	2040 355	1993 330	2118	2039			
	200	2167 155	2230 505	2169 467	2320	2228			
	500	2392 178	2447 795	2371 738	2552	2444			
5. - do - ( omitting 1942 and 1953 )	20	1549 92	1537 143	1546 136	1582	1537			
	50	1765 114	1748 226	1748 216	1808	1748			
	100	1928 132	1906 320	1899 305	1978	1905			
	200	2089 149	2075 452	2060 431	2159	2074			
	500	2803 172	2268 715	2245 682	2366	2267			
6. Penner at Nellore ( 1903-54 )	20	306 30	365 60	342 58	378	359			
	50	376 37	455 95	428 91	472	446			
	100	428 42	522 134	493 129	542	511			
	200	481 48	594 190	561 182	617	587			
	500	550 55	676 301	640 288	703	661			
7. Sutlej at Bhakra ( 1912-55 )	20	242 19	249 31	239 32	261	251			
	50	282 23	292 49	283 51	308	294			
	100	313 26	324 69	316 72	343	326			
	200	343 30	358 98	351 102	380	361			
	500	383 34	397 155	391 161	423	400			
8. Mahanadi at Naraj ( 1926-58 )	20	1787 153	1599 187	1696 205	1654	1580			
	50	2072 190	1818 295	1936 324	1890	1791			
	100	2286 219	1982 418	2116 458	2066	1949			
	200	2495 247	2157 591	2308 648	2255	2117			
	500	2781 285	2357 934	2527 1025	2471	2310			
9. Yamuna at Tajewala ( 1927-59 )	20	325 35	380 71	354 68	400	378			
	50	390 43	463 112	434 108	490	460			
	100	438 50	525 159	494 152	556	522			
	200	487 56	591 224	557 216	628	587			
	500	550 64	667 354	631 341	710	663			

TABLE 1 (Contd)

Period and site	Return Period (years)	Deduced with Parameters using					Plotting position concept due to	
		Maximum likelihood method	Moments		Mode and mean deviation		Gumbel	Hazen
			3	4	5	6	7	
10. Sone at Dehri (1920-54)	20	1196±117	1126±163	1193±174	1188	1126		
	50	1413 145	1317 258	1397 276	1400	1376		
	100	1576 166	1461 365	1550 390	1558	1459		
	200	1738 188	1613 516	1714 552	1727	1611		
	500	1952 217	1788 817	1901 872	1921	1785		
11. - do - (omitting 1933)	20	1154 113	1074 155	1132 166	1126	1065		
	50	1361 141	1252 245	1323 262	1321	1239		
	100	1517 162	1386 346	1466 370	1467	1369		
	200	1672 183	1529 489	1618 524	1623	1508		
	500	1877 211	1692 774	1793 828	1801	1668		
12. Ravi at Madhopur (1929-58)	20	336 37	445 92	418 85	456	429		
	50	403 46	549 145	514 134	562	526		
	100	453 53	626 205	586 189	641	598		
	200	502 60	709 289	663 268	726	676		
	500	568 69	804 458	751 424	823	765		
13. Hatmati at Himat Nagar (1922-49)	20	35 5	43 10	43 10	46	43		
	50	44 6	55 16	55 17	59	54		
	100	50 7	63 23	63 23	68	62		
	200	57 8	72 33	72 33	78	71		
	500	65 9	82 52	83 52	89	82		
14. Mahanadi at Sambalpur (1926-53)	20	1086 83	1006 104	1049 119	1043	1001		
	50	1227 103	1118 165	1177 188	1168	1111		
	100	1333 118	1202 233	1273 266	1261	1193		
	200	1438 133	1292 330	1376 377	1360	1281		
	500	1577 154	1395 522	1493 596	1474	1381		
15. Damodar at Rhondia (1933-56)	20	455 52	432 75	434 72	468	440		
	50	539 65	508 118	508 114	556	519		
	100	602 75	565 167	563 162	622	577		
	200	665 84	626 236	621 228	693	640		
	500	748 97	695 374	688 361	773	712		
16. Sabarmati at Dharoi (1935-52)	20	117 20	172 54	139 43	174	159		
	50	145 25	219 86	176 67	221	202		
	100	166 29	255 121	204 95	256	234		
	200	187 32	293 171	234 135	293	268		
	500	215 37	336 270	268 213	336	306		
17. - do - (omitting 1950)	20	82 14	85 23	84 24	95	86		
	50	101 17	104 36	105 38	118	105		
	100	115 20	118 50	120 54	135	120		
	200	129 22	134 71	137 77	153	135		
	500	147 26	151 113	156 122	174	153		
18. Tapi at Kathore (1940-53)	20	833 136	882 236	903 266	992	891		
	50	1000 169	1065 373	1108 420	1210	1076		
	100	1126 194	1202 527	1262 594	1374	1214		
	200	1251 220	1347 745	1427 840	1548	1361		
	500	1416 253	1514 1178	1615 1328	1749	1530		

## APPENDIX I

### *Standard Error of Peak Flow Estimates by Maximum Likelihood*

Since from (5),  $x_T = u - \frac{1}{a} \log_e \log_e \frac{T}{T-1} \equiv u - \frac{1}{a} R$

$$\text{Variance } (x_T) = \text{Variance } (u) + R^2 \text{ Variance } \left(\frac{1}{a}\right) - 2R \text{ Covariance } \left(u, \frac{1}{a}\right) \dots (16)$$

or, denoting Variances and Covariances by Var and Cov

$$\text{Var}(x_T) = \text{Var}(u) + R^2 \text{Var} \left(\frac{1}{a}\right) - 2R \text{Cov} \left(u, \frac{1}{a}\right) \dots \dots (16)$$

Now the Variance-Covariance matrix of  $u$  and  $a$  viz

$$\begin{vmatrix} \text{Var}(u) & \text{Cov}(u, a) \\ \text{Cov}(u, a) & \text{Var}(a) \end{vmatrix} \dots (17)$$

is the reciprocal\* of the matrix

$$-n \begin{vmatrix} E \left( \frac{d^2 \log f}{du^2} \right) & E \left( \frac{d^2 \log f}{da du} \right) \\ E \left( \frac{d^2 \log f}{da du} \right) & E \left( \frac{d^2 \log f}{da^2} \right) \end{vmatrix} \dots (18)$$

where 'f' stands for the distribution function

$$\frac{-a(x-u)}{ae} - e^{-a(x-u)}$$

and  $E$  for the expected values of the second order partial derivatives of  $\log f$ ,

$$\text{e.g. } E \left( \frac{d^2 \log f}{du^2} \right) = \int_{-\infty}^{+\infty} \frac{d^2 \log f}{du^2} f dx$$

On evaluating and substituting the expected values, (18) becomes

$$n \begin{vmatrix} a^2 & (1-C) \\ (1-C) & \frac{1}{a^2} \left[ \frac{\pi^2}{6} + (1-C)^2 \right] \end{vmatrix} \dots (19)$$

( $C$  being Euler's Constant = 0.577216).

---

\* M G Kendall; "The Advanced Theory of Statistics" Vol II, pp 36-42.

Reciprocal of (19) is

$$\frac{1}{n} \left\| \begin{array}{cc} \frac{1}{a^2} \left[ 1 + \frac{6(1-C)^2}{\pi^2} \right] & \frac{-6(1-C)}{\pi^2} \\ \frac{-6(1-C)}{\pi^2} & \frac{6a^2}{\pi^2} \end{array} \right\| \quad \dots \quad (20)$$

Comparison of (20) with (17) gives

$$\text{Var}(u) = \frac{1}{na^2} \left[ 1 + \frac{6(1-C)^2}{\pi^2} \right], \quad \text{Var}(a) = \frac{6\sigma^2}{n\pi^2}$$

$$\text{Cov}(u, a) = \frac{-6(1-C)}{n\pi^2}$$

wherefrom we get\*

$$\text{Var}\left(\frac{1}{a}\right) = \left(-\frac{1}{a^2}\right)^2 \text{Var}(a) = \frac{6}{n\pi^2 a^2}$$

$$\text{and Cov}\left(u, \frac{1}{a}\right) = -\frac{1}{a^2} \text{Cov}(u, a) = \frac{6(1-C)}{n\pi^2 a^2}$$

substituting which in (16) gives

$$\begin{aligned} \text{Var}(x_T) &= \frac{1}{na^2} \left[ 1 + \frac{6(1-C)^2}{\pi^2} \right] + R^2 \frac{6}{n\pi^2 a^2} - 2R \frac{6(1-C)}{n\pi^2 a^2} \\ &= \frac{1}{na^2} \left\{ 1 + \frac{6}{\pi^2} \left[ (1-C)^2 + R^2 - 2R(1-C) \right] \right\} \\ &= \frac{1}{na^2} \left[ 1 + \frac{6}{\pi^2} (1-C-R)^2 \right] \end{aligned}$$

Therefore Standard Error  $(x_T) = \sqrt{\text{Var}(x_T)}$

$$= \frac{1}{\sqrt{n} a} \left[ 1 + \frac{6}{\pi^2} (1-C-R)^2 \right]^{\frac{1}{2}}$$

---

\* M G Kendall, "The Advanced Theory of Statistics", Vol 1, pp 208, Charles Griffin & Co Ltd, London (1943)



APPENDIX II

Annual Peak Flows ( 000 cfs )

Year	Baitarni at Akhupada	Krishna at Vijayawada	Penner at Nellore	Godavari at Dowlaishwaram	Sutlej at Bhakra	Sone at Dehri	Hatmati at Himat Nagar	Mahanadi at Sambalpur	Mahanadi at Naraj	Yamuna at Tajewala	Ravi at Madhopur	Damodar at Rhondia	Sabarnati at Dharoi	Tapi at Kathore
1905	56	483	82	795										
06	172	463	48	1078										
07	275	495	144	1228										
08	101	572	105	1036										
09	122.5	653	473	901										
10	96	370	206	965										
1911	82	320	34	1144										
12	125	543	152	1186	237									
13	290	418	63	969	199									
14	92.5	949	29	1141	199									
15	83	578	98	1066	146									
16	65	955	382	915	75									
17	101	498	324	936	160									
18	85	254	72	718	70									
19	101	456	96	678	133									
20	263	392	69	528	145	674								
1921	65	558	142	998	136	580								
22	127	499	127	824	110	309	6.8							
23	140	624	91	555	93	1028	0.9							
24	52.5	634	151	629	237	965	5.9							
25	173	613	111	796	191	892	2.7							
26	125	552	30	933	96	871	34.9	841	1410					
27	325	560	158	845	141	600	29.7	698	1164	189.5				
28	172	461	17	725	85	NA	3.7	595	991	81				
29	125	458	67	845	162	638	7.1	840	1400	294	72			
30	102.5	370	169	972	150	NA	10.3	559	931	145	155			

Year	Baitarni	Krishna	Penner
1896	211	760	
97	50	619	
98	116	624	
99	72	405	
1900	228	719	
1901	114	513	
02	180	548	
03	96	1061	479
04	228	479	38

Year	Baitarni	Krishna
1874	87	
75	87	
1876	85	
77	183	
78	20	
79	95	
80	83	
81	250	
82	108	
83	110	
84	112	
85	54	
1886	80	
87	68	
88	112	
89	98	
90	72	
91	142	
92	112	
93	156	
94	195	578
95	75	496

1931	148	468	41	954	72	400	6.3	690	1150	82	.116			
32	146	512	6	933	135	953	2.9	566	943	213	157			
33	142	632	46	1020	93	580	15.5	847	1412	79	137	216		
34	140	384	69	1029	87	1209	31.4	NA	1185	89	77	106.5		
35	95	360	107	909	156	786	7.8	552	920	76	111	422.5	27	
36	91	350	95	1075	84	619.5	0.7	546	910	98	193	174.5	1	
37	127	393	51	785	83	614	60.1	825	1375	194	112	202	44	
38	47.5	370	143	1011	175	272	7.8	492	820	51	94	64	28	
39	117	402	158	896	83	681	7.8	816	1360	43	69	258.5	80	
40	217	354	199	1791	61	404	1.0	840	1400	70	115	266.5	32	243
1941	300	382	55	626	97	181	41.8	402	670	78	104	521.5	15	481
42	127	381	13	2108	177	959	10.8	656	1093	143	138	375.5	44	758
43	287	427	293	635	186	NA	7.7	714	1190	118	94	251	68	179
44	148	385	89	1794	81	959	45.6	793	1322	101	114.5	153	87	900
45	108	375	27	1164	109	279	20.6	589	981	97	114.5	121	66	722
46	192	475	420	1250	135	512	13.1	950	1445	129	103.5	314	16	339
47	45	443	132	1117	276	607	0.2	965	1495	563	566	260	20	291
48	64	369	15	671	83	561	0.1	948	1103	110	79	230	7	255
49	69	914	455	1119	107.5	443	9.0	701	634	126	133	230	17	662
50	134	565	33	1153	172	861		792	913	92	301	246	291	398
1951	83	360	42	846	325	421		450	394	215	71	347	36	162
52	84	406	294	526	83	356		848	1045	NA	101	168	2	115
53	108	478	55	2830	170	777		776	876	154	357	195		235
54	58	513	117	1084	199	206			450	224	152	243		
55	175	503		1364	130				NA	275	617	61		
56	142	830		1571					647	450	232	244		
57	62	592		1755					503	317	472			
58		740							931.5	199	128			
59										180				

NA denotes ' data not available. '

Year	Bairarni at Akhupada	Krishna at Vijayawada	Penner at Nellore	Godavari at Dowlaiswaram	Sutej at Bhakra	Sone at Dehri	Hatnani at Himat Nagar	Mahanadi at Sambalpur	Mahanadi at Naraj	Yamuna at Tajewala	Ravi at Madhopur	Dunodar at Rhondia	Subarnati at Dharoi	Tapl at Kathore
------	----------------------	-----------------------	-------------------	--------------------------	-----------------	---------------	------------------------	-----------------------	-------------------	--------------------	------------------	--------------------	---------------------	-----------------

## CWPRS TECHNICAL MEMORANDUMS

### SUBJECT CLASSIFICATION

Subject	Symbol
Water studies & Hydrology	HLO
Rainfall & Runoff	
Regeneration & Losses	
Quantity of Water	
Fluid Mechanics	FLM
Design of Hydraulic Structures	HYD
River Training & Flood Control	HYR
Irrigation, Drainage & Reclamation	IRD
Surface & subsurface irrigation	
Water requirements of crops	
Soil survey and crop pattern	
Drainage	
Reclamation	
Navigation	NAV
Ports, Harbours & Estuaries	
Inland Navigation	
Coastal Erosion	
Ships & Navigation Crafts	
Structural Design	STD
Soil Mechanics & Foundation Engineering	SFE
Investigations	
Stability, deformations & displacements of earth structures	
Foundation studies including geophysical	
Construction Materials & Practices	CMP
Concrete Technology	
Building & other Materials	
Instrumentation	IST
Model and Prototype comparisons	MPC
Other miscellaneous	MIS