TR-2

SUITABILITY OF POWER TRANSFORMATION BASED GUMBEL EV-I DISTRIBUTION FOR FLOOD FREQUENCY ANALYSIS

SATISH CHANDRA

DIRECTOR

STUDY GROUP

S M SETH

R D SINGH

M PERUMAL

NATIONAL INSTITUTE OF HYDROLOGY

JAL VIGYAN BHAVAN

ROORKEE-247667(U.P.)

1986-87

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Estimation of a flood corresponding to a specified recurrence interval at a particular point on a river system is the most common problem for the engineers, scientists and others involved in design and construction of water resources projects. Generally various distributions are fitted with the limited data availablde at the gauging site of the river system and then a suitable distribution is considered for estimating the floods of required recurrence intervals. Instead of selecting a suitable distribution among the various possible conditions of distributions for fitting the given set of data, it will be more appropriate if the data is transformed to a particular distribution of known characteristics for the purpose of flood frequency analysis. Box-Cox transformation is one of the powerful procedures for transforming the data series to near normalization, which has been employed in flood frequency analysis.

The Gumbel EV-I distribution, is one of the most popular distributions, used for flood frequency analysis and it is more amenable to theoretical analysis. In this study, an attempt has been made to develop methodology for transforming the annual peak flood series to follow Gumbel EV-I distribution using Box-Cox transformation. The exponent λ of the Box-Cox transformation has been estimated by trial and error using the method of maximum likelihood(MML) and method of probability weighted moment(MPWM), so as to obtain nearly the same estimates of log likelihood functions by both the methods. This methodology has been tested using 1000 samples of various sample sizes of randomly generated synthetic "flood" series which follow the Pearsen type III distribution. The statistical estimates of the reduced variates of the Gumbel EV-I distributed transformed series, viz., mean and standard deviation, have been found nearer to 0.5772 and 1.2825 as required from theoretical considerations and thus verifying the applicability of the proposed methodology for transforming to Gumbel

EV-I distribution. It is seen that the population estimates are satisfactorily reproduced by using the proposed method of frequency analysis.

1.0 INTRODUCTION

The main objective of flood frequency analysis is to estimate the flood corresponding to a specified recurrence interval from the limited sample data. An assumption is generally made of a theoretical frequency distribution which fits the population events and the parameters of the assumed distribution are computed from the sample data. However, data arising from various situations form their own distributions. As such instead of assuming that the data follows a particular distribution, it is more appropriate to transform the data to a particular distribution of known characteristics for the purpose of frequency analysis.

Since the properties of a normal distribution are completely defined, it is a common procedure to transform the given data series to a normally distributed series. The various normalization procedures generally used in practice are log transformation, log transformation based on the theoretical relationship between original and log domain statistical estimates(Chow 1954), inverse Pearson type III transformation, inverse log Pearson type III transformation (Beard, 1967), square root transformation(Richardson,1978) and Cube root transformation(Stidd,1953). All these transformations have been shown to be special cases of power transformation(Box and Cox,1964).

Since ;Gumbel EV-I distribution can be more easily handled by theoretical analysis, the transformation of data to this distribution would be more suitable for flood frequency analysis. In this study, an attempt has been made to develop procedure for use of the Box-Cox transformation for transforming synthetic data of flood series to Gumbel EV-I distribution.

2.0 REVIEW

Box and Cox(1964) suggested the power transformation for normalization. Hinkley(1977) gave a procedure to estimate the exponent λ required for the power transformation. In this procedure the sample mean, standard deviation and median of the transformed data, $\bar{Z}$, S_z and Z_m respectively, are computed for $\lambda = 0, \pm 0.5, \pm 1.0$, and then the appropriate value of λ is found by interpolation. At this value of λ, the quantity$(\bar{Z}-Z_m)/S_z$ would be minimum. Chander et al(1978) have found this transformation suitable for flood frequency analysis and gave a maximum likelihood estimator of λ. Recently, Kuczera(1983) has used this transformation method for regional flood frequency analysis.

Chow(1954) introduced the log transformation based on the theoretical relationship between original and log domain statistical estimates to normalise the data series. Beard(1967) proposed inverse Pearson type III and inverse log Pearson type III transformations for normalization. The inverse Pearson transformation formula assumes that the original or log transformated data follows Pearson type III or log Pearson type III distribution and they are near normally transformed using Beard's formula(1967). Richardson(1978) and Stidd(1953) suggested the use of square root transformation and cube root transformation respectively for the purpose of normalization.

Another transformation, referred to as the SMEMAX(Small, Medium, Maximum) transformation, has been suggested by Bethlahmy(1977) for normalising skewed frequency distributions. This approach is based upon the trignometrical solution of a right angled trianagle as shown in figure 1.

$$R_1 = X_m - X_{min} \qquad \text{...(1 a)}$$

$$R_2 = X_{max} - X_m \qquad \text{...(1 b)}$$

where X_{min}, X_m and X_{max} are the minimum, median and maximum values

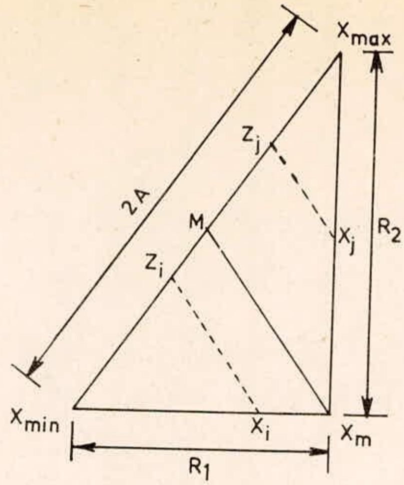

Fig. 1 The SMEMAX transformation

within the sample data. The point M(fig 1) disects the hypotenuse, whose length is 2A. The points Z_i and Z_j on the hypotenuse are the transformed values of the original observations, X_i and X_j, where $X_i Z_i$ and $X_j Z_j$ are drawn parallel to the line MX_m. As demonstrated by Venugopal(1980), the transformed values, as measured from the Vertex X_{min}, may then be obtained directly from the geometry of the triangle as:

$$Z_i = (X_i - X_{min})(A/R_1) \quad X_i \leqslant X_m \qquad ...(2a)$$

$$Z_i = A + (X_i - X_m)(A/R_2) \quad X_i > X_m \qquad ...(2b)$$

Chander et al (1978) pointed out that SMEMAX transformations transform the series having the difference between the largest value and the median value equal to that between the median value and the smallest value which is a necessary but not sufficient property to normalize the transformed series as the resulting series can still have appreciable skewness or kurtosis.

3.0 PROBLEM DEFINITION

It is required to develop a methodology for transforming the given random data series to Gumbel EV-I distribution and verify the effectiveness of the transformation technique by reproducing the population characteristics using this procedure.

4.0 PROPOSED METHODOLOGY

A given data series X is transformed to series Z of Gumbel EV I distribution using Box-Cox transformation as follows:

$$Z_i = \frac{(X_i)^{\lambda} - 1}{\lambda} \qquad \text{When } \lambda \neq 0 \qquad \ldots(3)$$

$$Z_i = \ln(X_i) \qquad \text{When } \lambda \rightarrow 0$$

where

λ = an exponent used for transformation and its value is found by trial and error, such that the series Z is approximately Gumbel EV-I distributed.

The series Z_i is approximately Gumbel EV-I distributed, as the moments of the transformed series higher than second order will not be strictly verified for their closeness with the theoretical values in the proposed methodology. It can also be seen that when $\lambda \rightarrow 0$, the transformed series follows the log Gumbel distribution.

The two methods of parameters estimation, the method of maximum likelihood(MML) and the method of probability weighted moments(MPWM) (Landwehr et al.,1979), are linked together in order to estimate λ of equation (3) which transforms the annual peak series into Gumbel EV-I distributed series. The probability density function of the Gumnbel EV distribution is given as:

$$P(Z_i) = \frac{1}{\alpha} \left[e^{-\left(\frac{Z_i - u}{\alpha}\right)} \right]^{-e^{-\left(\frac{Z_i - u}{\alpha}\right)}} \qquad \ldots(4)$$

in which,

$\left(\frac{Z_i - u}{\alpha}\right)$ is the reduced variate Y_i, and α and u are the parameters of the distribution.

4.1 Method of Maximum Likelihood(MML):

The maximum likelihood method of estimating the parameters, u and α ,postulates that the parameters should be such that the probability of N individual maximum events Z_1, Z_2, Z_3 ....Z_N corresponding to N annual peaks should be a maximum The log likelihood function is given as;

$$= \ln (P(z_1). P(Z_2) \ldots P(Z_N) =$$

$$N \ln(\frac{1}{\alpha}) - \frac{1}{\alpha} \sum_{i=1}^{N} (Z_i - u) - \sum_{i=1}^{N} e^{-(\frac{Z_i - u}{\alpha})} \quad \ldots(5)$$

The solution procedure for estimating u and α as described by Kite(1978) has been adopted in the study. Using the estimated values of u and α , the log likelihood function is estimated from equation(5).

4.2 Method of Probability Weighted Moments(MPWM)

The method of probability weighted moments for parameter estimation was introduced by Greenwood et al. (1979) for distributions that might be written in inverse from like Z= Z(F). In this form, the variable Z can be expressed as a function of corresponding probability of non-excedence. Landwehr et al(1979) used this method for parameter estimation of Gumbel EV-I distribution.

A random variable Z is said to be Gumbel EV-I distributed, if

$$F = \exp(-\exp(-(\frac{Z-u}{\alpha}) \quad \ldots(6)$$

where F= P(Z < z), the probability of non exceedence and Z, the variate. The inverse form of the distribution is defined as

$$Z = u - \alpha \ln(-\ln F) \quad \ldots(7)$$

The parameters of the Gumbel EV-I distribution are estimated using the technique of probability weighted moments on the sample data. The probability weighted moments are expressed as follows:

$$M_K = \frac{1}{N} \sum_{i=1}^{N} Z_i(F)(1-F_i)^K \qquad K = 0,1,2 \quad \ldots(8)$$

where N = Number of Z_i variates

Landwehr et al.(1979) substituted the formula

$\binom{n-i}{K} / \binom{n-1}{K}$ in place of $(1-F_i)^K$ and expressed the parameters of the distribution u and α using MPWM as:

$$\hat{\alpha} = (M_o - 2M_1) / \ln(2) \qquad ...(9)$$

$$\hat{u} = M_o - 0.5772\,\hat{\alpha} \qquad ...(10)$$

where, $\hat{u}$ and $\hat{\alpha}$ are the sample estimates of u and α .

4.3 Linking MML and MPWM:

Landwehr et al (1979) have found that the estimates of $\hat{u}$ and $\hat{\alpha}$ obtained using MPWM are unbiased. However, they have shown that the estimates of MML has lesser variance than that of MPWM. In the proposed method, the estimates $\hat{u}_1$ and $\hat{\alpha}_1$ have been obtained first by using MML in the transformed Series Z_i for a given λ , and using these estimates the probability of non-exceedance F_i is determined corresponding to each Z, by the following expression:

$$F_i = \exp\left(- \exp\left(\frac{Z_i - \hat{u}_1}{\hat{\alpha}_1}\right)\right) \qquad ...(11)$$

This expression for F_i is used in equation(8) to estimate the probability weighted moments M_o and M_1. Equations(9) and (10) have been used to estimate $\hat{u}_2$ and $\hat{\alpha}_2$, where $\hat{u}_2$ and $\hat{\alpha}_2$ are the estimates of method of probability weighted moments. The log likelihood function is calculated using the parameters $\hat{u}_2$ and $\hat{\alpha}_2$. The ratio R of log likelihood functions of MML and MPWM is computed as:

$$R = \frac{\ln\,(p(Z_1)\, p(Z_2) ...p(Z_N)\,)\ \text{MML}}{\ln\,(PZ_1)\, p(Z_2)\ ...\ p(Z_N))\ \text{MPWM}} \qquad ...(12)$$

The ratio R would be nearer to unity corresponding to that trial value of λ which yields the values of $\hat{u}_1$ and $\hat{\alpha}_1$ nearly the same as that of $\hat{u}_2$ and $\hat{\alpha}_2$ respectively. For this condition, the coefficient of variation of reduced variate $Y_i = (Z_i - u)/\alpha$ would also be nearer to 2.222 as required from theoretical considerations.

4.4 Estimation of λ :

The expondnent λ of equation(3) can not be determined in closed form. It is determined by trial and error method using Newton-Raphson technique. The steps to be followed for the estimation of λ are:

(i) Assume two different values for λ i.e. λ_1 and λ_2.

(ii) Use equation 3 considering λ_1 in place of λ, to transform the annual maximum series, X_i, to series $Z1_i$.

(iii) Fit EV-I distribution with the transformed data and estimate the parameters $\hat{u}_1$ and $\hat{\alpha}_1$ using the method of maximum likelihood as described in section 4.1

(iv) Estimate the values of F_i using the parameters $\hat{u}_1$ and $\hat{\alpha}_1$ in equation(11).

(v) Estimate the probability weighted moments M_o and M_1 from the equation (8) using the values of F_i obtained from step(iv)

(vi) Estimate $\hat{u}_2$ and $\hat{\alpha}_2$ using the equations(9) & (10).

(vii) Compute the reduced variate series $Y1_i$ using the equation:

$$Y1_i = \frac{(Z1_i - \hat{u}_2)}{\hat{\alpha}_2} \qquad \ldots(13)$$

(viii) Compute the co-efficient of variation (CoV1) for the reduced variate series obtained from step(vii).

(ix) Repeat step(ii) to(viii) for λ_2 in order to estimate the co-efficient of variation(COV2) for the corresponding reduced variate series.

(x) Estimate the new trial value of λ using (Newton-Raphson) the equation:-

$$\lambda new = \lambda_2 - (\frac{FCN}{FPN}) \qquad \ldots(14)$$

where FCN = (COV 2-2.222) ...(15)

$$FPN = \frac{((COV2-3.222) -(COV1-2.222))}{\lambda_2 - \lambda_1} \text{ or } FPN = \frac{(COV2-COV1)}{\lambda_2 - \lambda_1} \qquad \ldots(16)$$

(ix) Consider the next trial values of λ_1 and λ_2 as:

$\lambda_1 = \lambda_2$(the previous trial value)

$\lambda_2 = \lambda_{new}$ (Evaluated at Step (x))

(xii) Repeat Step(ii) to (xi) till the convergence criteria are not satisfied. The convergence criteria are:(i) the objective function should be less than or equal to 0.01 or (ii) the total no. of trials should not exceed the maximum specified limit(say 100).

(xiii) The last trial value of λ_2, which satisfies one of the above stated convergence criteria, gives an estimate of required λ.

Estimate of λ ,obtained from the above steps, transform the annual maximum series to a Gumbel EV1 distributed series using Box-Cox transformation. For this condition, the ratio of log likelihood functions of the MML and MPWM is nearer to unity and also the co-efficient of variation of the reduced variate series y_i is nearer to 2.222 as required from theoretical considerations.

5.0 ANALYSIS AND RESULTS

5.1 Monte Carlo Experiments:

In order to judge the applicability of this methodology, Monte Carlo experiments were used. For this purpose, the skewed distribution were generated using the following approximate relationship(Clarke, 1973) between the reduced variate of Pearson type III distribution and the co-efficient of skewness;

$$Y = p\left(1 - \frac{1}{9p} + \frac{t}{3\sqrt{p}}\right)^3 \qquad \text{...(17)}$$

where

$$p = \frac{4}{g^2} \qquad \text{...(18)}$$

g being the given co-efficient of skewness and t, the standard normal deviate.

The frequency factor, K of Pearson type-III distribution is written as (Kite, 1978):

$$K = -\frac{2}{g} + \frac{gY}{2} \qquad \text{..(19)}$$

The Pearson type-III distributed synthetic "floods" were generated using the following equation(Chow, 1964):

$$x = \mu + \sigma K \qquad \text{...(20)}$$

where, μ and σ are the given population mean and standard deviation respectively.

Two sets of 100,000 numbers which follow the pearson type-III distribution approximately with co-efficient of Skewness 1.0 and 0.5 were generated using the above described procedure. For both the sets, population mean and standard deviation were supplied as 500 and 200 respectively, although any other mean and standard deviation could have been used.

Since, the relationship between the reduced variate and the co-efficient of skewness given by equation(17) is only approximate the co-efficient of skewness of 100,000 generated numbers need not be nearly equal to the supplied population co-efficient of skewness. Therefore, two synthetically generated sets of 100,000 numbers were considered as representative of population and their respective computed mean, standard deviation and co-efficient of skewness were treated as population parameters. The mean, standard deviation and co-efficient of skewness thus computed for the two sets of synthetic "floods" series are given below:

Series	Mean	Standard deviation, σ	co-efficient of skewness γ
Case-I	528	289	1.336
Case-II	520	275	0.750

Population estimates of "floods" for return periods of 20,50,100,200,500 and 1000 years were obtained using the population mean, standard deviation and co-efficient of skewness as given above. Using 100,000 generated numbers of each of the two series considered, 1000 samples for each of the following sample sizes were studied using the proposed methodology:

N = 10, 20, 30, 40, 50, 60, 70, 80, 90, and 100

For each of the sample size considered, 1000 sample estimates for return periods of 20,50,100,200,500 and 1000 years were computed. The estimates in the original domain were estimated from those in the transformed domain using the following inverse transformation relationship:

$$X_T = (Z_T\lambda + 1)^{1/\lambda} \quad ...(21)$$

These estimates were subjected to the performance tests adopting the criteria discussed below:

5.2 Performance Criteria:

The performance indices used for judging the accuracy of the suggested methodology are bias(δ), co-efficient of variation(CV) and root mean square error

(RMSE) as defined below. These indices were normalized by the population values for the purpose of comparison(Lettenmaier and Burges,1982).

(i) Bias:-

It is the tendency to overestimate or underestimate a given event level corresponding to the population estimate. A positive bias indicates the overestimation and a negative bias indicates the under-estimation. It is measured as:

$$\text{Bias} \quad \delta = \frac{E(\hat{X}_T) - X_T}{X_T} \qquad \text{...(22)}$$

where

$E(\hat{X}_T)$ = mean of the estimates of $\hat{X}_T$ for a given sample size.

$\hat{X}_T$ = the estimate of flood corresponding to T-year recurrence interval

(ii) Co-efficient of variation(cv):

The co-efficient ov variation is a measure of. the precision of estimation or scatter of the estimates derived from many samples of the same sample size. It is measured as:

$$CV = \frac{(E|\hat{X}_T - E(\hat{X}_T)|^2)^{1/2}}{X_T} \qquad \text{...(23)}$$

(iii) Root mean square error(RMSE):

RMSE is a common statistical measure which combines the effects of bias and variability. It measures the accuracy of the suggested methodology in fitting the population estimate.

It is estimated as:

$$RMSE = (\delta^2 + CV^2)^{1/2} = \frac{E(\hat{X}_T - X_T)^2}{X_T} \qquad \text{...(24)}$$

5.3 Discussion of Results:

Table 1 shows the expected values of mean, standard deviation, and co-efficient of skewness of the reduced variate series of the Gumbel EV-I distribution obtained by using the suggested method of transformation on 1000 samples of various sample sizes derived from Case-I and Case-II series. It can also be seen that the expected values of mean computed over the reduced variates of 1000 samples for all the sample sizes of both the series considered yield a value of 0.5772 which corresponds to the theoretical value for Gumbel EV-I distributed reduced variate series. Similarly, the expected values of the standard deviations computed over the reduced variates of 1000 samples for all the sample sizes of both the series considered, yield a value close to 1.2825 which is the theoretical value for the Gumbel EV-I distributed reduced variate series. However, the expected value of the co-efficient of skewness of the reduced variates of case-II series are nearer to 1.139 (the theoretical value of Gumbel EV-I distribution), indicating that case II series could be transformed more close to Gumbel EV-I distribution as compared to Case I series.

Figure 2-13 illustrate the bias and RMSE, in percentages, of the estimates of different return period "floods" obtained by using the proposed methodology on different sample sizes with 1000 samples each considered from Case I and Case II series. Since the value of RMSE and CV do not differ much, only RMSE has been considered for analysis.

From figure 2-7 of case I series, it can be inferred that the bias in the estimates are less than 5% except in the estimates corresponding to the return periods of 500 and 1000 years which are computed as 16% and 28% respectively for sample size equal to 10(not shown in the figures). It can be seen from the variation of bias values that the lower return period floods are under estimated and higher return periods floods are over estimated. The RMSE of the estimates reduces as the sample size increases for a given return period and it increases with return periods. This has resulted in very high values of the RMSE as 115% and 185% respectively for the estimates of 500 and 1000 year return periods obtained using sample size 10. The reduction in RMSE of the estimates of all return periods is quite significant when the sample size increases from 10 to 20 indicating the importance of sample size in the frequency analysis using the proposed methodology.

It can be seen from figures 8-13 that the variation of bias and RMSE of the estimates obtained using Case-II series also indicate somewhat similar behaviour as for Case I series, but with smaller variations. It can be seen that the "floods" are generally under estimated. Also the magnitudes of bias and RMSE of the estimates are generally less than that of the estimates arrived by using Case-I series. This difference in bias and RMSE obtained using Case-I and Case-II series may be attributed to the difference in skewness of the generated population series.

Figures 14 and 15 describe the variation characteristics of the exponent used for transformation of 1000 samples of each of the different sample sizes considered from both case-I and case-II series. It can be seen from figure 14, that the mean values of λ obtained by transforming 1000 samples of the different sample sizes arrived from case-II series are less than the correspondingn mean λ values of Case-I series indicating the effect of co-efficient of skewness of the population series. Similarly, as it can be seen from figure 15, the co-efficient of variation of λ values corresponding to case-II series are less than of case-I series for all sample sizes explaining the possible higher RMSE values in the estimates of Case-I series than that of Case II Series.

Table 1

Expected Values of Mean, Standard deviation and co-efficient of Skewness of the reduced variate series obtained using the suggested transformation procedure on 1000 samples

Sample size	Case-I Series $y=E(\bar{y})$	$y=E(S_y)$	$\gamma_y=E(g_y)$	Case-II Series $u=E(\bar{y})$	$y=E(S_y)$	$\gamma_y = E(g_y)$
10	0.5772	1.2813	0.9108	0.5772	1.2816	0.9004
20	0.5772	1.2810	0.9893	0.5772	1.2817	1.0150
30	0.5772	1.2809	0.9971	0.5772	1.2812	1.058
40	0.5772	1.2805	1.0084	0.5772	1.2813	1.0867
50	0.5772	1.2807	1.0352	0.5772	1.2816	1.0996
60	0.5772	1.2807	1.0489	0.5772	1.2815	1.1109
70	0.5772	1.2808	1.0617	0.5772	1.2815	1.1137
80	0.5772	1.2809	1.0767	0.5772	1.2820	1.1207
90	0.5772	1.2809	1.0781	0.5772	1.2820	1.1186
100	0.5772	1.2808	1.0818	0.5772	1.2816	1.1209

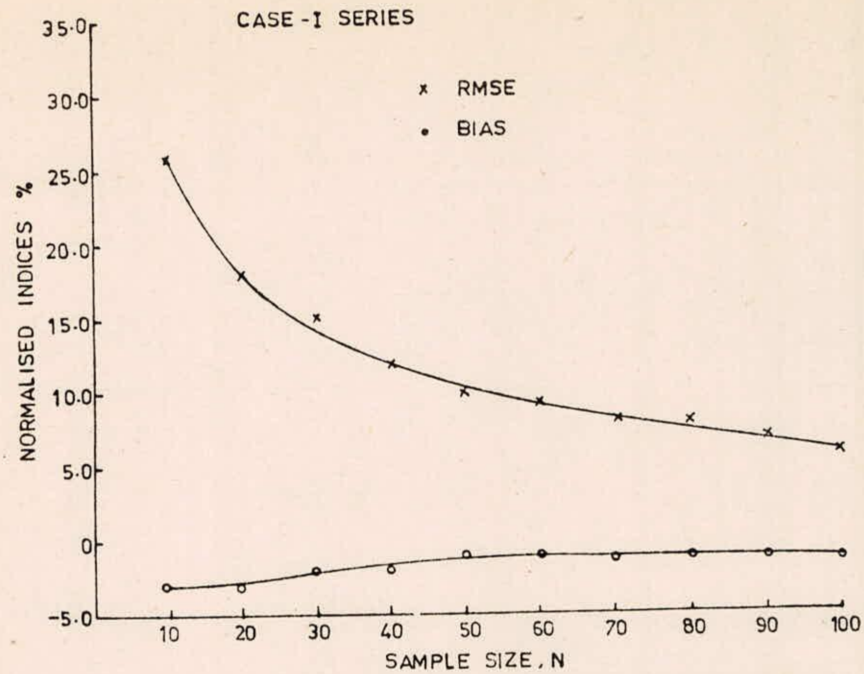

FIG. 2 PERFORMANCE TESTS FOR THE SUGGESTED TRANSFORMATION PROCEDURE FOR RETURN PERIOD T = 20 YEARS

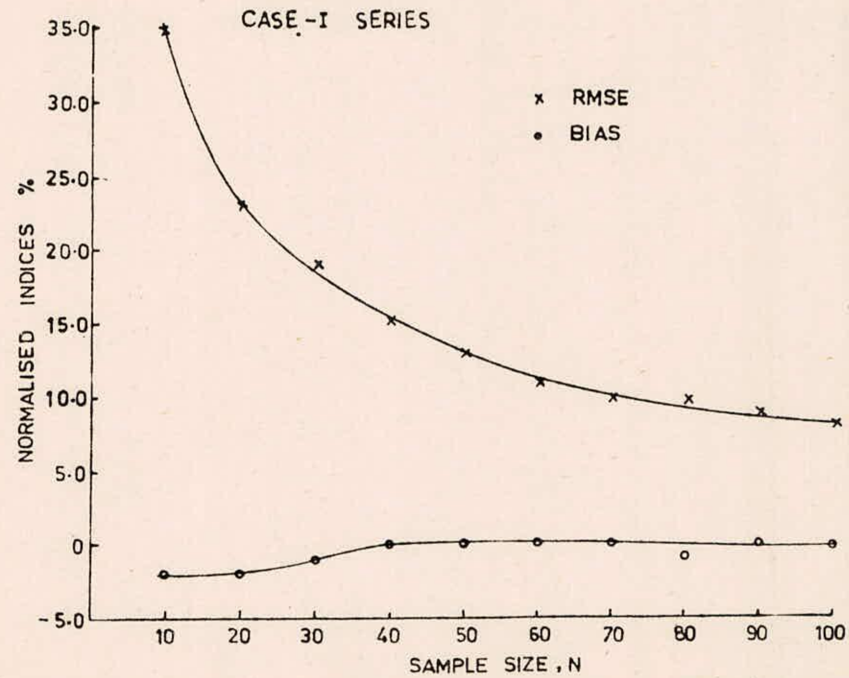

FIG. 3 PERFORMANCE TESTS FOR THE SUGGESTED TRANSFORMATION PROCEDURE FOR RETURN PERIOD T = 50 YEARS

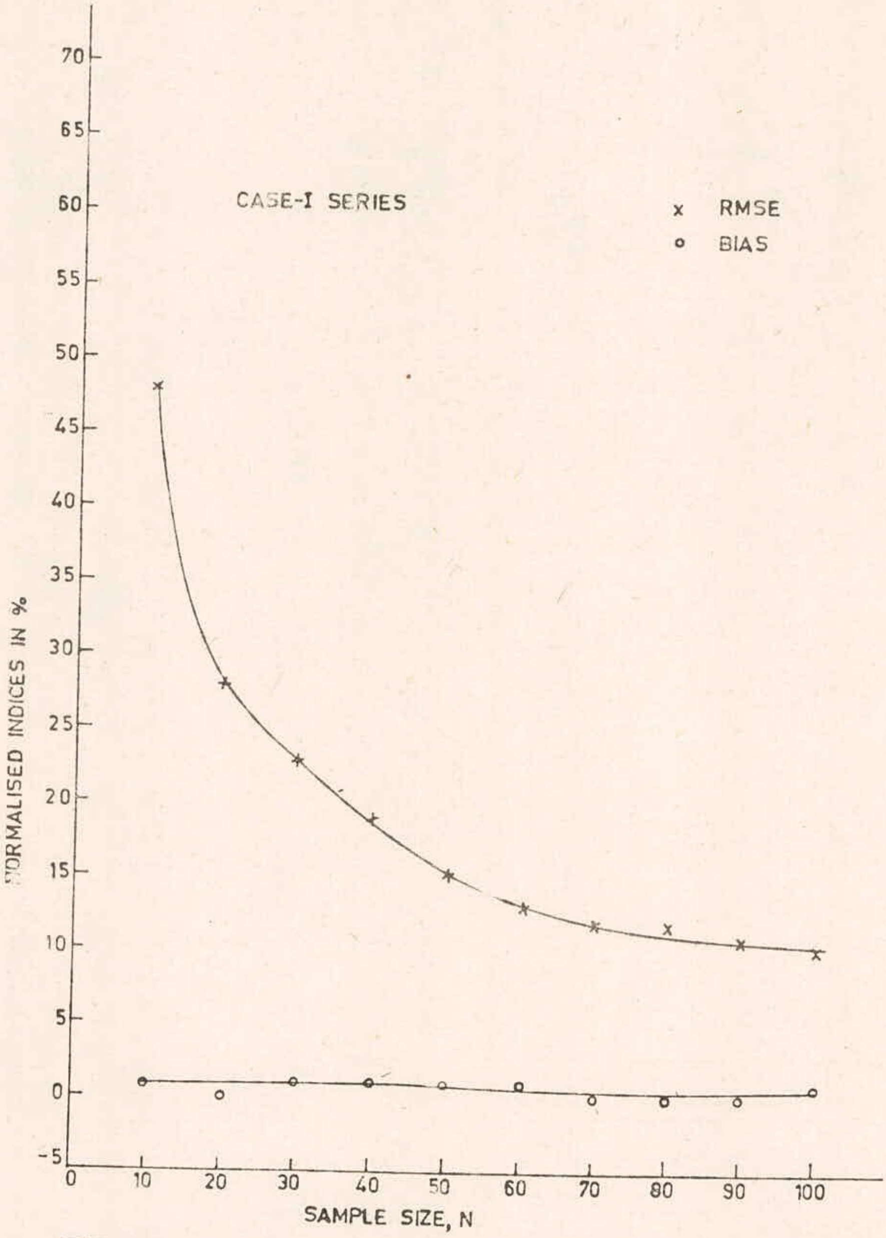

FIG. 4 PERFORMANCE TESTS FOR THE SUGGESTED TRANSFORMATION PROCEDURE FOR RETURN PERIOD T=100 YEARS

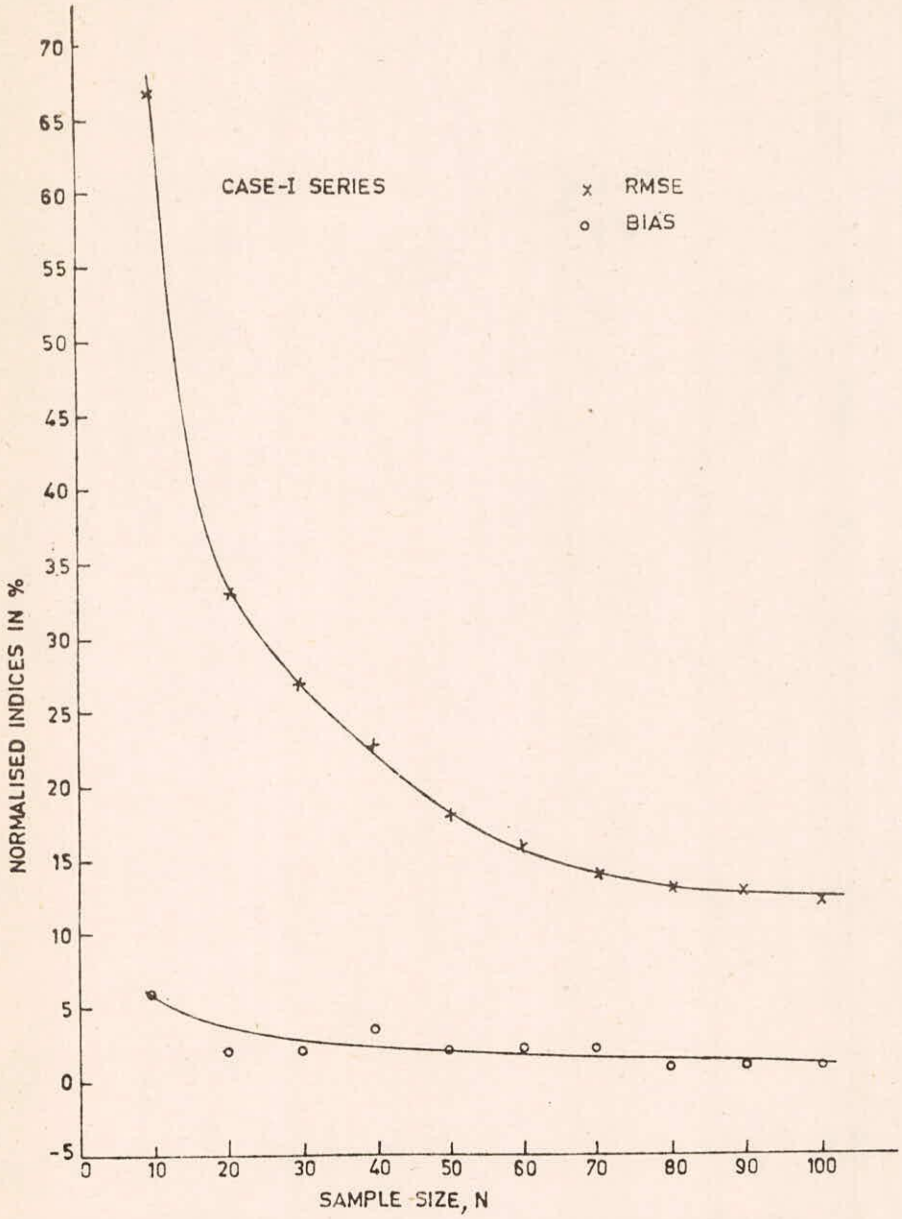

FIG. 5 PERFORMANCE TESTS FOR THE SUGGESTED TRANSFORMATION PROCEDURE FOR RETURN PERIOD T=200 YEARS

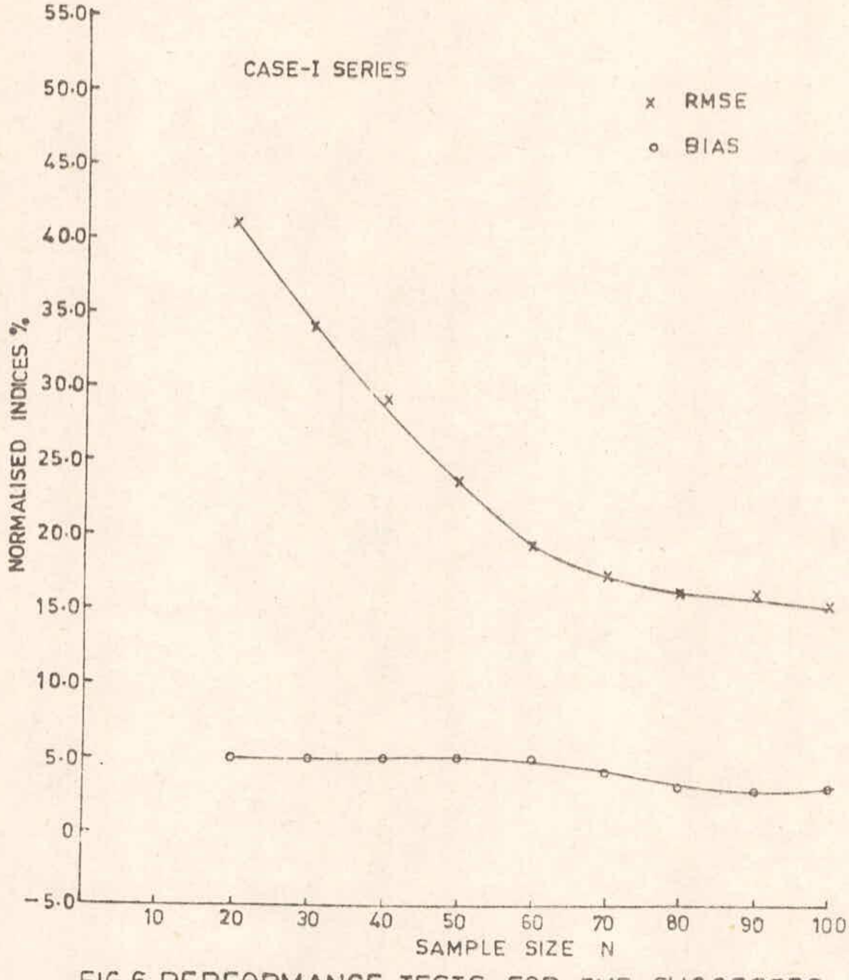

FIG.6 PERFORMANCE TESTS FOR THE SUGGESTED TRANSFORMATION PROCEDURE FOR RETURN PERIOD T=500 YEARS

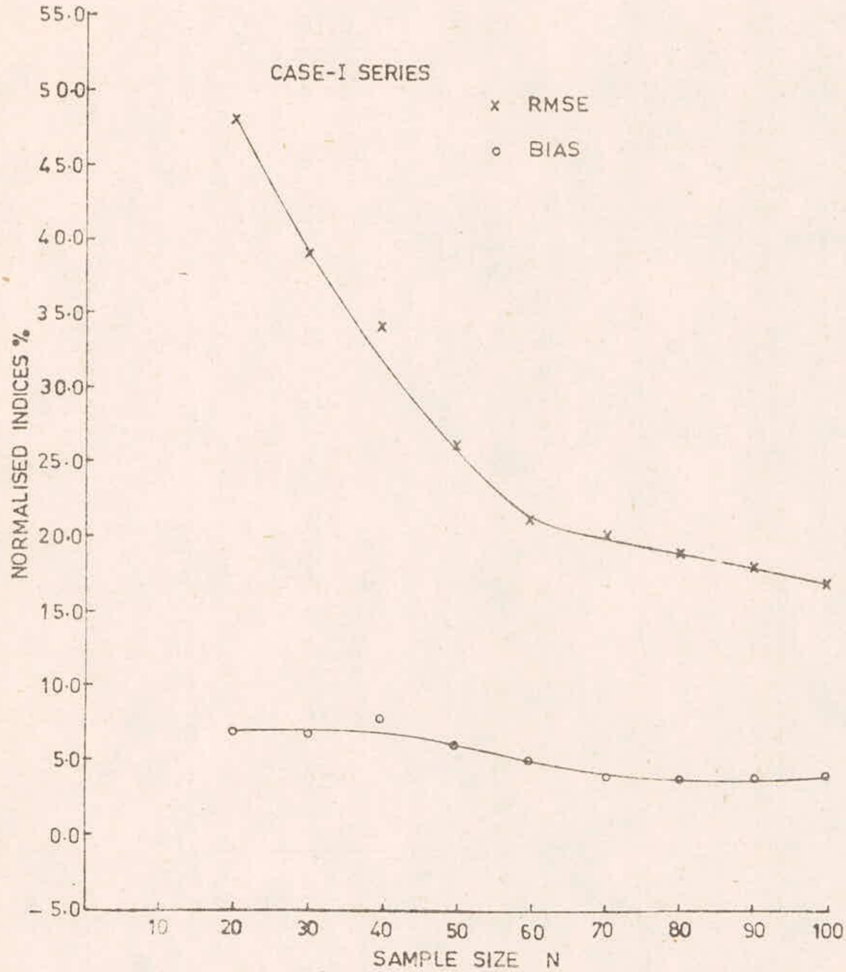

FIG.7 PERFORMANCE TESTS FOR THE SUGGESTED TRANSFORMATION PROCEDURE FOR RETURN PERIOD T = 1000 YEARS

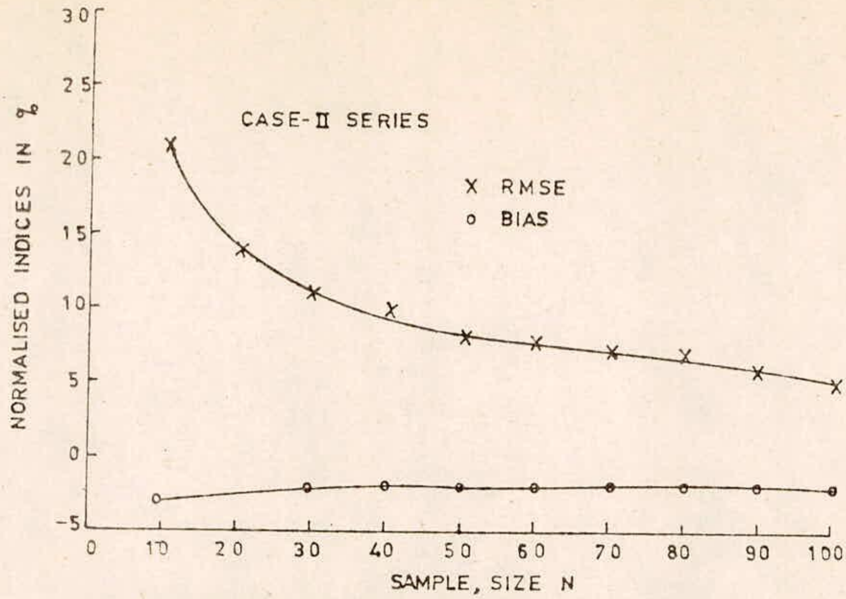

FIG.8 · PERFORMANCE TESTS FOR THE SUGGESTED TRANSFORMATION PROCEDURE FOR RETURN PERIOD T=20 YEARS

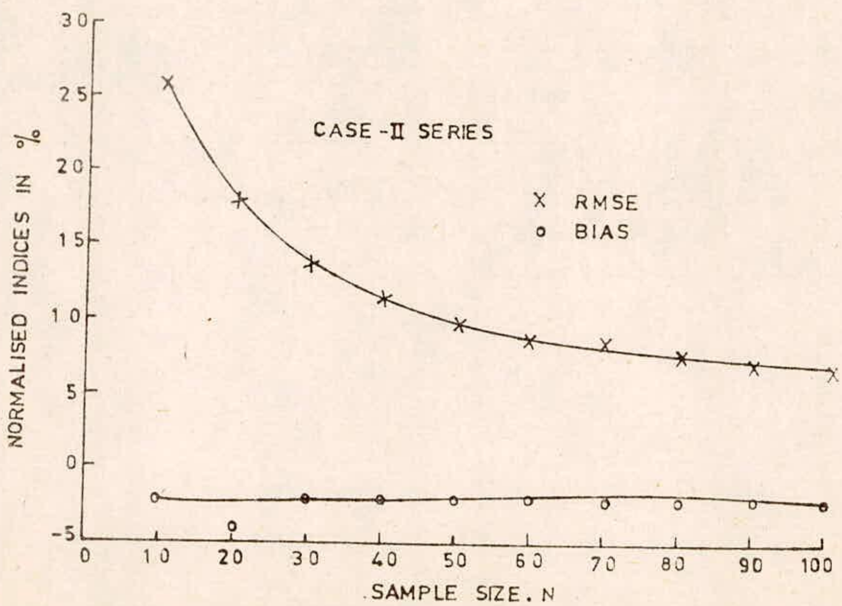

FIG.9 -PERFORMANCE TESTS FOR THE SUGGESTED TRANSFORMATION PROCEDURE FOR RETURN PERIOD T=50 YEARS

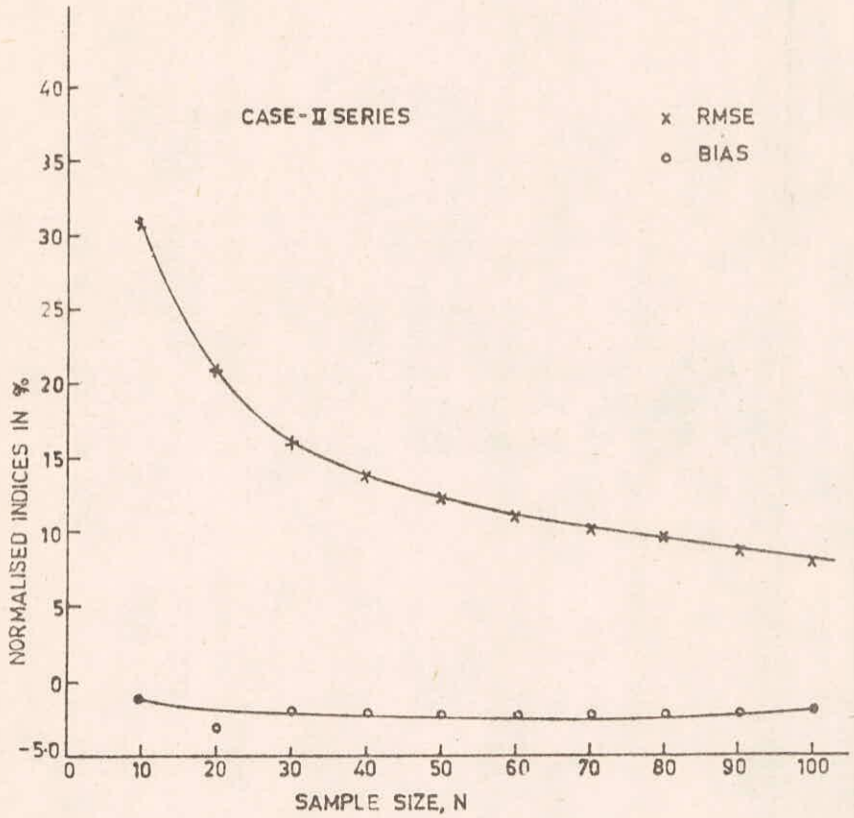

FIG.10 PERFORMANCE TESTS FOR THE SUGGESTED TRANSFORMATION PROCEDURE FOR RETURN PERIOD T=100 YEARS

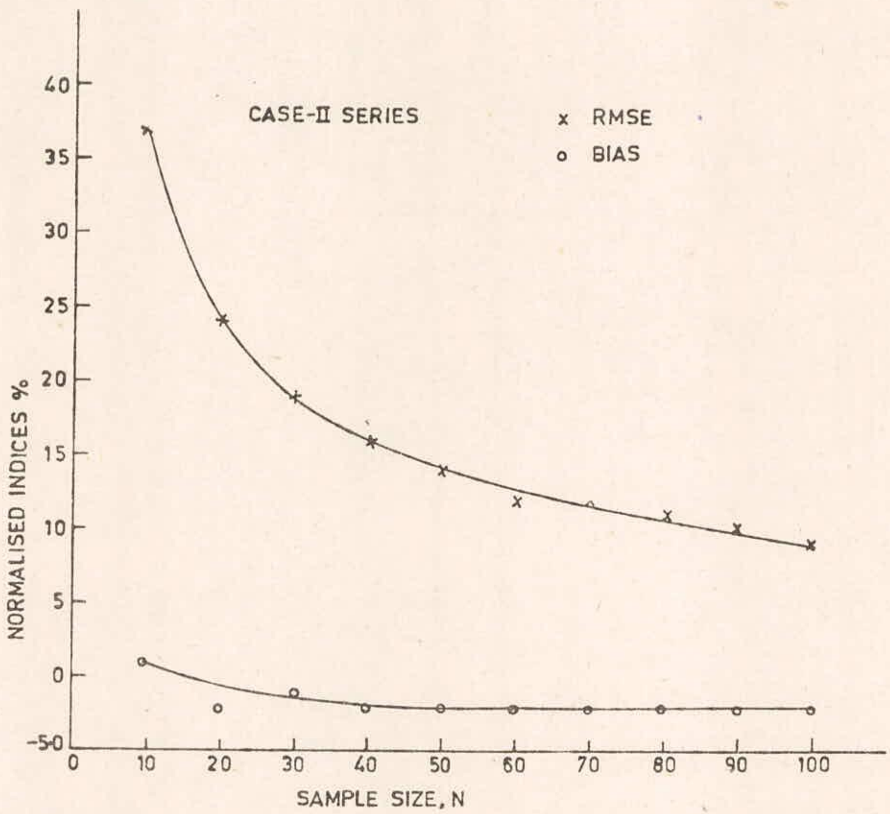

FIG. 11 PERFORMANCE TESTS FOR THE SUGGESTED TRANSFORMATION PROCEDURE FOR RETURN PERIOD T=200 YEARS

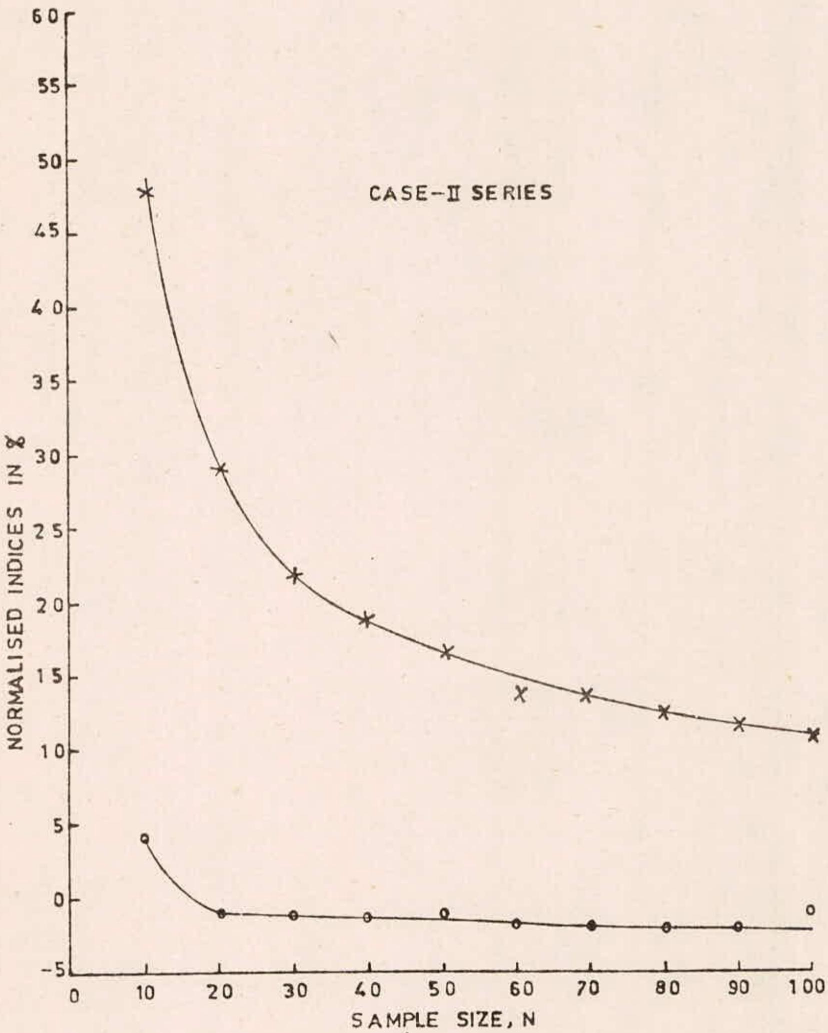

FIG.12 PERFORMANCE TESTS FOR THE SUGGESTED TRANSFORMATION PROCEDURE FOR RETURN PERIOD T=500 YEARS

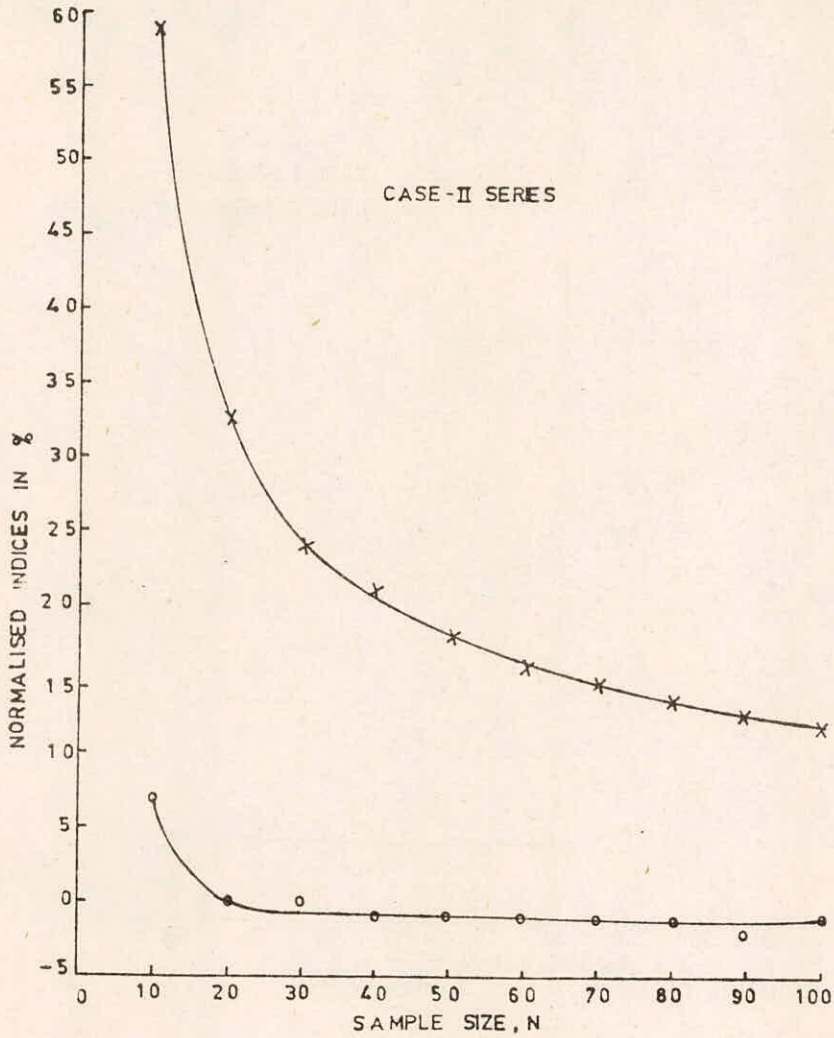

FIG.13 - PERFORMANCE TESTS FOR THE SUGGESTED TRANSFORMATION PROCEDURE FOR RETURN PERIOD T=1000 YEARS

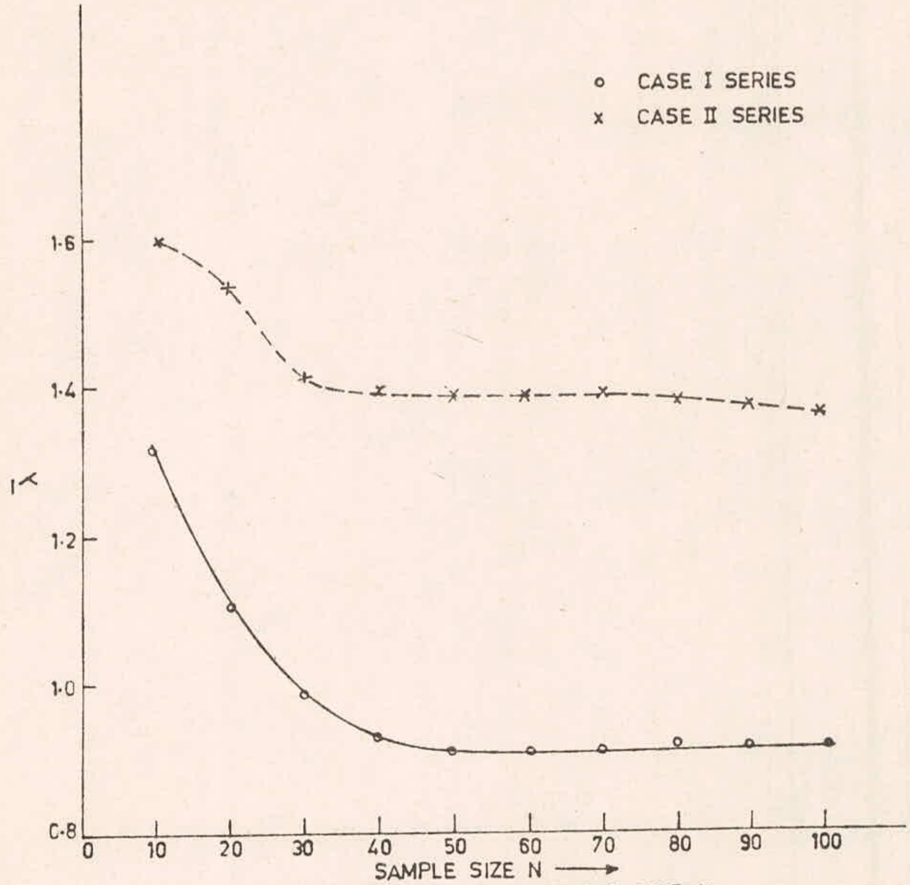

FIG.14 - VARIATION OF MEAN λ WITH SAMPLE SIZE N

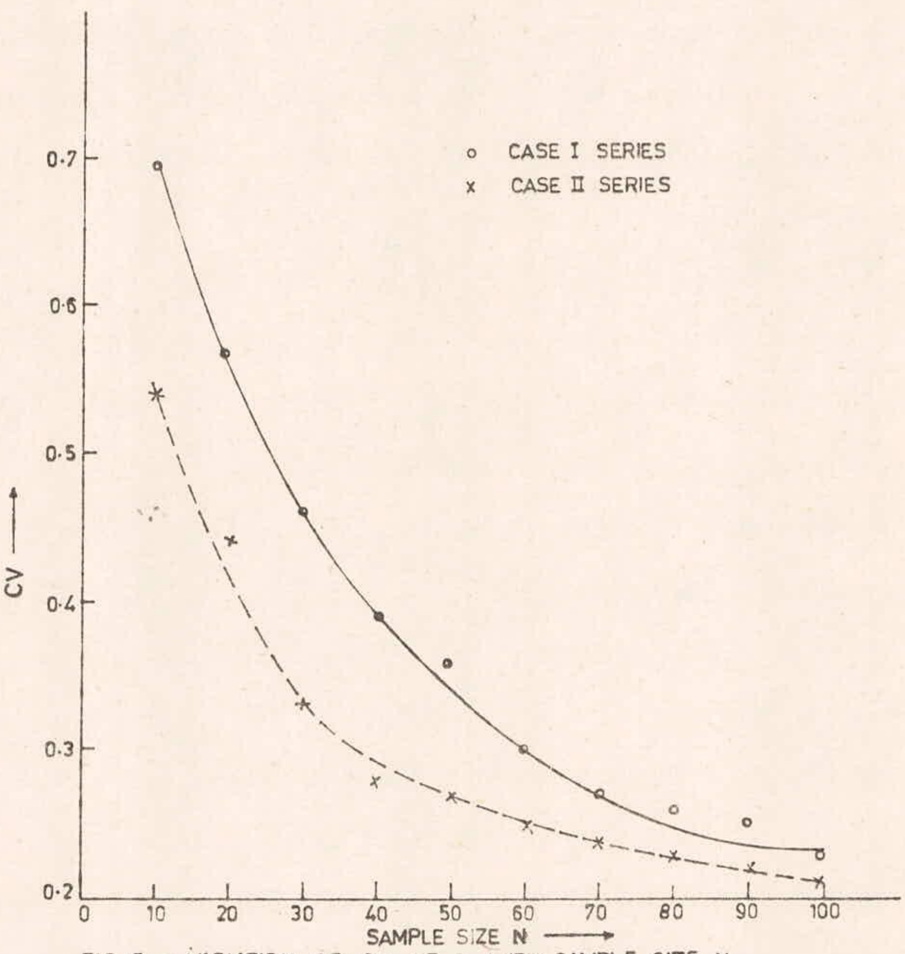

FIG.15 - VARIATION OF CV OF λ WITH SAMPLE SIZE N

6.0 CONCLUSIONS

From the present study carried out with Monte Carlo experiments involving two different cases of generated series, the following conclusions are made:-

(i) The expected values of mean and standard deviation of the reduced variate series of the Gumbel EV-I distribution, obtained by using the proposed methodology for both cases close to the theoretical values 0.5772 and 1.2825 respectively for the Gumbel EV-I distributed reduced variate series. However, the expected value of the co-efficient of skewness of the EV-I reduced variates of case-II series is nearer to the theoretical value of the co-efficient of skewness i.e. 1.139 as compared to the Case-I series. It indicates that the Case-II series is transformed somewhat more closely to Gumbel EV-I distribution in comparison to the Case-I series.

(ii) The bias and root mean square errors are quite high for all return period floods computed from a sample of size 10. This indicates the importance of sample size in flood frequency analysis.

(iii) The mean values of exponent λ used for transformation have a systematic variation with sample size. Similar variations with different sample size are also observed. for the values of co-efficient of variation of λ .

(iv) Inspite of the effect of co-efficient of skewness on the estimation of floods as indicated in the study, the methodology suggested can be used satisfactorily for transforming given data series to Gumbel EV-I distribution.

REFERENCES

1. Benson,M.A.(1968)'Uniform Flood Frequency estimating methods for Federal Agencies, Water Resour.Res.,Vol.4,pp.891-908.

2. Bethlahmy, N.(1977),'Flood analysis by SMEMAX transformation', Proc. Am.Soc. Civ.Engrs.,J.Hydraul.Div.Vol.103,(HY1) pp.69-78.

3. Box ,G.E.P. and D.R.Cox (1964),'An analysis of transformation', Journal of the Royal Statistical Society, Vol.B 26,pp.211-252.

4. Chander,S., S.K.Spolia and Arun Kumar(1978),'Flood Frequency Analysis by Power Transformation', Journal of the Hydraulics Division, Proc.A.S.C.E., Vol.104, No.HY11, Nov.,pp.1495-1504.

5. Chow,V.T.(1954),'The log probability and its engineering application, proc. ASCE,Vol.80, Paper No.536, Nov.pp.1-25.

6. Chow,V.T.(1964),'Statistical and Probability Analysis of Hydrologic Data' Part I.Frequency Analysis in : V.T.Chow(Editor),Handbook of Applied Hydrology, McGraw Hill Book Co., New York.

7. Clarke,R.T.(1973) 'Mathematical Models in Hydrology', Food and Agriculture Organisation of the United Nations, Rome.

8. Greenwood,J.A., J.M.Landwehr,N.C.Matalas and J.R.Wallis(1979)' Probability Weighted Moments,' Water Resources Research, Vol.15, No.5, October.

9. Hinkley,D.(1977),'On quick choice of power transformation' Appl.Stati.,26, pp.67-69.

10. Kite,G.W.(1978) 'Frequency and Risk Analysis in Hydrology', Water Resources Publications, Fort Collins, Colorado.

11. Kuczera, G.(1983)'A Bayesian Surrogate for Regional Skew in Flood Frequency Analysis' Water Resources Research, Vol.19, No.3, June,pp.821-832.

12. Landwehr,J.M.,N.C.Matalas and J.R.Wallis(1979)'Probability weighted Moments compared with some Traditional Techniques in Estimating Gumbel Parameters and qualities' Water Resources Research, Vol.15, No.5, October pp.1055-1064.

13. Lettenmaier,D.P. and S.J.Burges(1982) 'Gumbel's Extreme Value I Distribution: A New Look' Journal of Hydraulics Division, Proc.A.S.C.E.,Vol. 108, No.HY4, April,pp.502-514.

14. Richardson,C.W.(1977),'A model of Stochastic structure of daily precipitation over an area', Hydrology paper No.91, CSU, Fort Collins, Colorado July.

15. Stedinger,J.R.(1980)'Fitting log normal distributions to hydrologic data', Wat.Resourc.Res.16,pp.481-490.

16. Stidd,C.K.(1953),'Cube root normal precipitation distributions', Trans. Am.Geophy., Union, Vol.34, No.1, pp.31-35.

17. Venugopal, K(1980),'Flood analysis by SMEMAX transformation- a review', Proc.Am.Soc.Civ.Engrs.J.Hydraul.Div.Vol.106, No.HY2, pp.338-340.