J. Indian Water Resour. Soc.,
Vol. 41, No. 2, Apr., 2021

A STUDY ON COMPARISON OF EXTREME RAINFALL ESTIMATES USING L-MOMENTS OF SIX PROBABILITY DISTRIBUTIONS

N. Vivekanandan*

ABSTRACT

Estimation of rainfall for a given return period is utmost importance for planning, design and management of hydraulic structures. This can be achieved through Extreme Value Analysis (EVA) by fitting probability distributions (PDs) to the series of annual 1-day maximum rainfall data. In this paper, a study on comparison of extreme rainfall estimates using L-Moments (LMO) of six PDs such as Exponential, Extreme Value Type-1, Extreme Value Type-2, Generalized Extreme Value (GEV), Generalized Pareto and Log-Normal for Afzalpur and Kalaburagi sites is carried out. The adequacy of fitting six PDs adopted in EVA of rainfall is quantitatively assessed by applying the Goodness-of-Fit (GoF) (viz., Chi-square and Kolmogorov-Smirnov) and diagnostic (viz., D-index) tests, and qualitatively assessed by using the probability plots of the estimated rainfall. The EVA results of rainfall indicate the GEV (using LMO) is better suited amongst six PDs adopted for estimation of rainfall at Afzalpur and Kalaburagi sites.

Keywords: *Chi-square, D-index, Extreme Value Analysis, Generalized Extreme Value, Kolmogorov-Smirnov, L-Moments, Probability distribution, Rainfall*

INTRODUCTION

Rainfall is one of the most important natural input resources that help us to study about the temporal and spatial variations in nature. One of the important problem in hydrology deals with the interpreting past records of hydrological event in terms of future probabilities of occurrence. In this context, Extreme Value Analysis (EVA) of rainfall is generally considered as one of the important tools to determine the expected rainfall at various chances (Subudhi, 2007). Such information can be used to prevent floods and droughts, and also considered for planning, design and management of hydraulic structures (Mujere, 2011).

A number of Probability Distributions (PDs) viz., Exponential (EXP), Extreme Value Type-1 (EV1), Extreme Value Type-2 (EV2), Generalized Extreme Value (GEV), Generalized Pareto (GPA) and 2-parameter Log-Normal (LN2) (Naghavi et al., 1993) are widely applied for EVA of rainfall. Generally, Method of Moments (MoM) is used in determining the parameters of the distributions. Sometimes, it is difficult to assess exact information about the shape of a PD that is conveyed by its third and higher order moments. Also, when the sample size is small, the numerical values of sample moments can be very different from those of the distribution from which the sample was drawn. It is also reported that the estimated parameters of the distributions fitted using MoM are often less accurate than those obtained by other parameter estimation procedures. To address these shortcomings, the application of alternative approach, namely, the L-Moments (LMO) adopted in EVA is discussed in this paper.

* *Scientist-B, Central Water and Power Research Station, Pune, Maharashtra, India*
Corresponding author Email: anandaan@rediffmail.com

Manuscript No. 1561
Received 23 March 2019; Accepted 20 April 2021

Sharma and Singh (2010) analyzed the daily rainfall data to identify the best fit distribution for Adilabad district of Telangana. They expressed that the LN2 was found as the best fit distribution for the annual season while gamma for monsoon season. AlHassoun (2011) carried out a study on developing empirical formula to estimate the rainfall intensity in Riyadh region using EV1 (commonly known as Gumbel), LN2 and Log Pearson Type-3 (LP3). He concluded that the LP3 gives better accuracy amongst three distributions studied in estimation of rainfall intensity. Esteves (2013) applied the Gumbel distribution to estimate the extreme rainfall depths at different rain-gauge stations in southeast United Kingdom. Vivekanandan (2014) applied the Gumbel distribution for modelling the seasonal and annual rainfall for Krishna and Godavari river basins. Rasel and Hossain (2015) applied the Gumbel distribution for development of intensity-duration-frequency curves for seven divisions in Bangladesh. Afungang and Bateira (2016) applied the Gumbel distribution to estimate the maximum amount of rainfall for different periods in the Bamenda mountain region, Cameroon. Studies carried out by Sasireka et al. (2019) indicated that the extreme rainfall for various return periods obtained from Gumbel distribution could be used for design purposes by considering the risk involved in the operation and management of hydraulic structures in Tiruchirappalli region.

Thus, the studies reported didn’t suggest applying a particular PD for EVA of rainfall for different region or country. This apart, when different PDs are used for rainfall estimation, a common problem is encountered as regards the issue of best model fits for a given set of data. This can be answered by formal statistical procedures involving Goodness-of-Fit (GoF) and diagnostic tests; and the results are quantifiable and reliable. Qualitative assessment is made from the plots of the observed and estimated extreme (i.e., 1-day maximum) rainfall. For quantitative assessment on rainfall estimation within in the observed range, Goodness-of-Fit (GoF) (viz., Chi-square (χ^2) and Kolmogorov-Smirnov (KS)) and diagnostic (viz., D-index) tests are applied (Zhang, 2002; USWRC, 1981). The paper presents the methodology adopted in EVA of rainfall using LMOs of six PDs with an objective to identify the best suitable distribution for rainfall estimation through quantitative and qualitative assessments with an illustrative example and the results obtained thereof.

This research paper is arranged in the following manner. The studies conducted by numerous researchers in estimating the rainfall are discussed in the section on 'Introduction'. The procedures adopted in EVA of rainfall is presented in the section on 'Methodology'. The discussion on the results obtained from the study are presented in the section on 'Results and Discussion'.

METHODOLOGY

The procedures involved in carrying out EVA include (i) determination of parameters of the six PDs (viz., EXP, EV1, EV2, GEV, GPA and LN2) using LMO; (ii) estimation of extreme (i.e., 1-day maximum) rainfall for different return periods; (iii) evaluation of the PDs adopted in EVA using quantitative and qualitative assessments; and (iv) analyze the results and made discussions thereof.

Theoretical Description of L-Moments

LMOs are analogous to ordinary moments, which provide measures of location, dispersion, skewness, kurtosis and other aspects of the shape of the PDs or data samples (Hosking, 1990). But, LMOs are computed from linear combinations of the probability weighted moments. LMO can be used as the basis of a unified approach to the statistical analysis adopting PDs (CWC, 2010).

Let x_1, x_2,.....,x_N be a conceptual random sample of size N and $x_{1N} \le x_{2N} \le ..... \le x_{NN}$ denote the corresponding order statistics. The r+1th LMOs defined by Hosking and Wallis (1993) are given as below:

$$\lambda_{r+1} = \sum_{k=0}^{r} \frac{(-1)^{r-k}(r+k)!}{(k!)^2(r-k)!} b_k \quad \dots (1)$$

where, $\lambda_{r+1}$is the r+1th sample moment and b_k is an unbiased estimator with

$$b_k = \frac{1}{N} \sum_{i=k+1}^{N} \frac{(i-1)(i-2).....(i-k)}{(N-1)(N-2).....(N-k)} x_{iN} \quad \dots (2)$$

The first two LMOs (λ_1 and λ_2) are expressed by:

$$\lambda_1 = b_0 \text{ and } \lambda_2 = 2b_1 - b_0 \quad \dots (3)$$

Table 1 presents the quantile functions and the parameters of six PDs considered in the study.

In Table 1, F(x) (or F) = P = 1-(1/T) is the Cumulative Distribution Function (CDF) of the variable x (i.e., Annual 1-day Maximum Rainfall (AMR)), μ is the average of the log-transformed series of AMR, σ is the standard deviation of the log transformed series of AMR, ξ is the location parameter, α is the scale parameter, k is the shape parameter, K_P is the frequency factor corresponding to probability of exceedance (P) and x_T is the estimated rainfall for a given return period (T) (Bobee and Ashkar, 1991).

Goodness-of-Fit Tests

GoF tests are essential for checking the adequacy of fitting PD to the annual maximum series of rainfall data, which is used for EVA. Out of a number GoF tests available, the widely accepted GoF tests are χ^2 and KS (Zhang, 2002), which are used in the study. The theoretical descriptions of GoF tests statistic are given as below:

χ^2 test statistic is defined by:

Table 1: Quantile functions and parameters of six PDs

Distribution	Quantile function (x_T)	Parameters by LMO
EXP	$x_T = \xi - \alpha \ln(1-F)$	ξ=λ_1-2λ_2 and 2λ_2=α
EV1	$x_T = \xi - \alpha \ln(-\ln(1-F))$	ξ=λ_1-(0.5772157)α and λ_2=α(ln2)
EV2	$x_T = \alpha e^{(-\ln(-\ln(F))/k}$	By using the logarithmic transformation of the observed data, the parameters of EV1 are initially obtained by LMO. These parameters are used to determine the parameters of EV2 from $\alpha = e^{\xi}$ and k=1/(scale parameter of EV1).
GEV	$x_T = \xi + (\alpha(1-(-\ln F)^k)/k)$	$z = (2/(3+t_3) - (\ln 2/\ln 3)$; $t_3 = (2(1-3^{-k})/(1-2^{-k})) - 3$ $k = 7.8590z + 2.9554z^2$; $\alpha = \lambda_2 k/(1-2^{-k})\Gamma(1+k)$ $\xi = \lambda_1 + (\alpha(\Gamma(1+k)-1)/k)$
GPA	$x_T = \xi + (\alpha(1-(1-F)^k)/k)$	$\xi = \lambda_1 + \lambda_2(k+2)$; $t_3 = (1-k)/(3+k)$ $k = (4/(t_3+1)) - 3$; $\alpha = (1+k)(2+k)\lambda_2$
LN2	$x_T = \exp(\mu + K_P \sigma)$	λ_1=μ and λ_2=σ/3.14286; $\lambda_2 = 2b_1 - b_0$ $\lambda_1 = b_0 = \frac{1}{N}\sum_{i=1}^{N}\ln(x_i)$ and $b_1 = \frac{1}{N(N-1)}\sum_{i=2}^{N}(i-1)\ln(x_i)$

$$\chi^2 = \sum_{j=1}^{NC} \frac{\left(O_j(x) - E_j(x)\right)^2}{E_j(x)} \quad \ldots (4)$$

where, $O_j(x)$ is the observed frequency value (x) of j^{th} class, $E_j(x)$ is the expected frequency value (x) of j^{th} class and NC is the number of frequency classes. The rejection region of χ^2 statistic at the desired significance level (η) is given by $\chi^2_C \geq \chi^2_{1-\eta, NC-m-1}$. Here, m denotes the number of parameters of the distribution and χ^2_C is the computed value of χ^2 statistic by PD.

KS test statistic is defined by:

$$KS = \underset{i=1}{\overset{N}{Max}} \left(F_e(x_i) - F_c(x_i)\right) \quad \ldots (5)$$

where, $F_e(x_i)=i/(N+1)$ is the empirical CDF of x_i, $F_c(x_i)$ is the computed CDF of x_i using PD, x_i is the observed AMR for i^{th} sample and N is the number of samples.

Test criteria: If the computed values of GoF tests statistic given by the PD are not greater than its theoretical values at the desired level of significance then the distribution is considered to be adequate for EVA of rainfall at that level.

Diagnostic Test

The selection of a suitable PD for estimation of rainfall is made through D-index, which is given as below:

$$D-\text{index} = \frac{1}{\bar{x}} \sum_{i=1}^{6} \left| x_i - x_i^* \right| \quad \ldots (6)$$

where, x_i is the observed AMR, x_i^* is the estimated 1-day maximum rainfall for i^{th} sample and $\bar{x}$ is the average of the observed AMR. The probability distribution with the lowest D-index value is considered as better suited distribution for estimation of rainfall (USWRC, 1981).

STUDY AREA AND DATA USED

In this paper, a study on rainfall estimation by using LMOs of six PDs for Afzalpur and Kalaburagi sites is carried out. The daily rainfall data for the period 1970 to 2018 for Afzalpur and Kalaburagi is used. The AMR data series is derived from the daily rainfall data and also used in EVA. From the scrutiny of rainfall data, it is found that the data for the year 1982 of Afzalpur while the data for the years 1975, 1993 and 2015 of Kalaburagi are not available. However, by considering the importance of the hydrological extremes, the rainfall data for the missing years are ignored and also not considered in EVA. The descriptive statistics of the derived AMR data series of Afzalpur and Kalaburagi is presented in Table 2.

Table 2: Descriptive statistics of the AMR data series

Site	Average (mm)	Standard deviation (mm)	Coefficient of skewness	Coefficient of kurtosis
Afzalpur	79.1	52.6	5.061	30.614
Kalaburagi	81.7	22.3	0.367	-0.474

RESULTS AND DISCUSSION

By applying the procedures, as described above, a computer code was developed and used for conducting EVA of rainfall. The computer code will determine the LMO estimators of the six PDs, estimate the extreme rainfall for different return periods, and compute the GoF and diagnostic tests statistic values for the data under study.

Estimation of Extreme Rainfall

The parameters of the six PDs considered in the study were determined by LMO and are used for estimation of rainfall at Afzalpur and Kalaburagi sites. The EVA results are presented in Tables 3 and 4 while the plots are shown in Figure 1.

Table 3: Extreme rainfall estimates given by LMOs of six PDs for Afzalpur

Return period (year)	1-day maximum rainfall (mm)					
	EXP	EV1	EV2	GEV	GPA	LN2
2	67.4	73.3	67.3	66.4	65.2	71.4
5	102.3	104.4	95.8	93.8	97.1	97.2
10	128.7	125.0	121.0	119.3	125.2	114.1
20	155.0	144.8	151.4	151.2	157.3	130.3
25	163.5	151.1	162.6	163.2	168.5	135.4
50	189.9	170.4	202.4	207.1	206.8	151.3
100	216.3	189.6	251.5	263.2	250.6	167.1
200	242.7	208.7	312.3	335.3	300.7	183.1
500	277.6	233.9	415.6	463.0	378.0	204.4
1000	304.0	252.9	515.7	592.0	446.2	218.4

Table 4: Extreme rainfall estimates given by LMOs of six PDs for Kalaburagi

Return period (year)	1-day maximum rainfall (mm)					
	EXP	EV1	EV2	GEV	GPA	LN2
2	73.8	77.8	75.0	79.6	79.2	78.7
5	97.2	98.7	95.2	100.2	103.7	99.9
10	114.9	112.5	111.4	112.2	114.3	113.3
20	132.6	125.8	129.5	122.6	121.0	125.6
25	138.3	130.0	135.9	125.7	122.5	129.4
50	156.0	142.9	157.4	134.7	126.1	141.1
100	173.7	155.8	182.3	142.9	128.3	152.4
200	191.4	168.6	210.9	150.3	129.7	163.6
500	214.8	185.5	255.6	159.1	130.7	178.3
1000	232.5	198.3	295.6	165.0	131.2	190.5

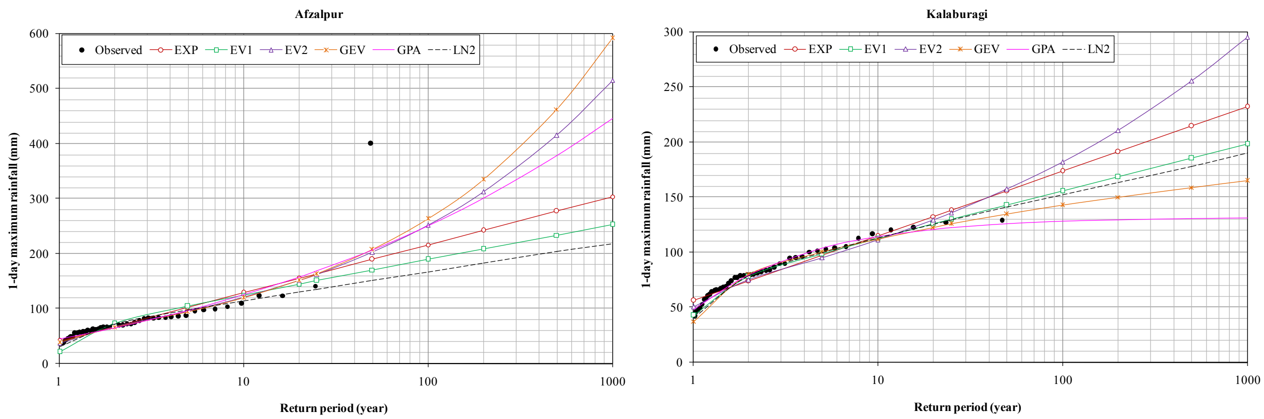

Fig 1: Plots of estimated 1-day maximum rainfall by six PDs and observed data for Afzalpur and Kalaburagi

From Tables 3 and 4, for the return period from 50-years and above, it is noticed that the estimated extreme (i.e., 1-day maximum) rainfall by GEV for Afzalpur and EV2 for Kalaburagi are higher than those values of other distributions considered in the study. For Afzalpur and Kalaburagi, it can be seen that the plots (Figure 1) of estimated rainfall using EXP, EV1 and LN2 distributions are in the form of linear whereas the plots of EV2 and GEV are in the form of exponentially growth. The EVA results obtained from GPA distribution indicated that the fitted curve is in the form of exponentially growth for Afzalpur while exponentially decay for Kalaburagi.

Analysis of Results Based on GoF Tests

The adequacy of fitting six PDs adopted in EVA of rainfall was performed by adopting GoF tests, as described earlier. In the present study, the degree of freedom (NC-m-1) was considered as two for 3-parameters distributions (viz., GEV and GPA) and three for 2-parameters distributions (viz., EXP, EV1, EV2 and LN2) while computing the χ^2 test statistic values.

From GoF tests results, as given in Table 5, the following observations were made:

i) χ^2 test results confirmed the suitability of EV2 and GEV distributions for EVA of rainfall at Afzalpur.

ii) KS test results didn't support the use of GPA and LN2 distributions for EVA of rainfall at Afzalpur.

iii) χ^2 and KS tests results supported the use of EXP, EV1, EV2, GEV and LN2 distributions for EVA of rainfall for Kalaburagi.

Analysis of Results Based on Diagnostic Test

For the selection of the best suitable distribution for rainfall estimation, the D-index values were computed by six PDs and are presented in Table 6.

From diagnostic test results, as given in Table 6, the following observations were made:

Table 5: Theoretical and computed values of GoF tests statistic by six PDs for Afzalpur and Kalaburagi

Distribution	Theoretical value at 5% level			Computed values of GoF tests statistic			
	χ^2 (Afzalpur and Kalaburagi)	KS		Afzalpur		Kalaburagi	
		Afzalpur	Kalaburagi	χ^2	KS	χ^2	KS
EXP	7.82	0.186	0.187	10.250	0.128	5.391	0.149
EV1	7.82	0.186	0.187	8.750	0.137	3.043	0.075
EV2	7.82	0.186	0.187	2.750	0.086	5.130	0.131
GEV	5.99	0.186	0.187	2.250	0.084	0.696	0.044
GPA	5.99	0.186	0.187	9.250	0.342	6.478	0.194
LN2	7.82	0.186	0.187	21.250	0.202	0.696	0.056

Table 6: D-index values given by six PDs for Afzalpur and Kalaburagi

Site	D-index					
	EXP	EV1	EV2	GEV	GPA	LN2
Afzalpur	3.860	3.708	3.340	3.216	3.582	3.381
Kalaburagi	0.571	0.369	0.633	0.319	0.249	0.323

i) The D-index values of GEV for Afzalpur while GPA for Kalaburagi is found to be the lowest when compared with those values of other PDs considered in the study.

ii) The GoF tests results indicate the GPA distribution is not adequate for EVA of rainfall for Kalaburagi though its D-index value is noted as lower than those values of other PDs considered in the study.

iii) By eliminating the D-index value of GPA from the selection of PDs adopted in EVA of rainfall, it can be seen that the D-index of GEV is noted as the second next minimum to GPA for Kalaburagi.

Selection of Probability Distribution

Based on the analysis of GoF and diagnostic tests results, it is identified that the GEV (using LMO) as better suited amongst six PDs adopted for estimation of rainfall at Afzalpur and Kalaburagi sites. The plots of estimated 1-day maximum rainfall with 95% confidence limits by GEV (using LMO) distribution and observed AMR data for Afzalpur and Kalaburagi sites are presented in Figure 2.

EVA of rainfall whereas KS test results inferred the adequacy of fitting the EXP, EV1, EV2 and GEV distributions for EVA.

iii) The χ^2 and KS tests results indicated that the EXP, EV1, EV2, GEV and LN2 distributions are acceptable for EVA of rainfall for Kalaburagi.

iv) The D-index values of GEV for Afzalpur and GPA for Kalaburagi are noted as the lowest when compared with those values of other PDs considered in EVA of rainfall.

v) Qualitative assessment of the outcomes was weighed together with D-index values and fitted curves of the estimated extreme (i.e., 1-day maximum) rainfall. Accordingly, GEV (using LMO) distribution is considered as the best choice for estimation of rainfall at Afzalpur and Kalaburagi sites.

The study suggested that the extreme (i.e., 1-day maximum) rainfall values for different return periods given by GEV (using LMO) distribution could be used for planning, design and management of hydraulic structures in Afzalpur and Kalaburagi sites.

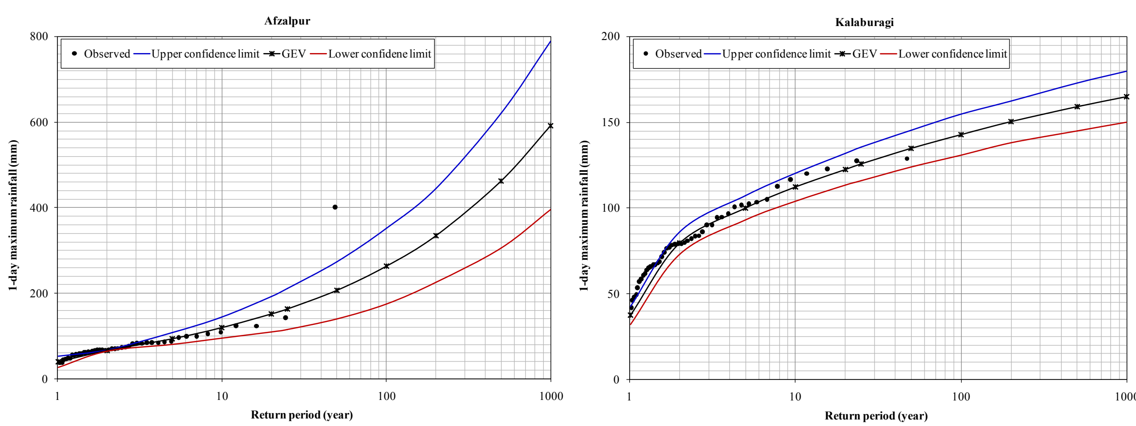

Fig 2: Plots of estimated 1-day maximum rainfall with 95% confidence limits by GEV (using LMO) distribution and observed AMR data for Afzalpur and Kalaburagi

CONCLUSIONS

The paper presents the study carried out for rainfall estimation through EVA by using LMOs of six PDs (viz., EXP, EV1, EV2, GEV, GPA and LN2) for Afzalpur and Kalaburagi. The intercomparison on EVA results of rainfall was performed and the following conclusions were drawn from the study:

i) For the return period of 50-years and above, it was found that the estimated extreme (i.e., 1-day maximum) rainfall by GEV for Afzalpur and EV2 for Kalaburagi are higher than those values of other PDs considered in the study.

ii) For Afzalpur, it could be seen that the χ^2 test results supported the use of EV2 and GEV distributions for

ACKNOWLEDGEMENTS

The author is grateful to the Director, Central Water and Power Research Station, Pune, for providing the research facilities to carry out the study. The author is thankful to India Meteorological Department, Pune, for the supply of rainfall data used in the study.

REFERENCES

1. Afungang, R., and Bateira, C., Statistical modelling of extreme rainfall, return periods and associated hazards in the Bamenda Mountain, NW Cameroon, 2016, 1(9): 5-19.

2. AlHassoun, S.A., Developing an empirical formulae to estimate rainfall intensity in Riyadh region, Journal of

King Saud University-Engineering Sciences, 2011, 23(2): 81-88.

3. Bobee, B., and Ashkar, F., The gamma family and derived distributions applied in hydrology, Water Resources Publications, USA, 1991.

4. CWC, Development of hydrological design aids (surface water) under Hydrology Project-II: State of the Art Report, Consulting Engineering Services (India) in Association with HR Wallingford, Central Water Commission (CWC), New Delhi, 2010.

5. Esteves, L.S., Consequences to flood management of using different probability distributions to estimate extreme rainfall, Journal of Environmental Management, 2013, 115 (1): 98-105.

6. Hosking, J.R.M., and Wallis, J.R., Some statistics useful in regional frequency analysis, Water Resources Research, 1993, 29 (2): 271-281.

7. Hosking, J.R.M., L-moments: Analysis and estimation of distributions using linear combinations of order statistics, Royal Statistical Society (Series B), 1990, 52(1): 105-124.

8. Mujere, N., Flood frequency analysis using the Gumbel distribution, Journal of Computer Science and Engineering, 2011, 3(7): 2774-2778.

9. Naghavi, B., Yu, F.X., and Singh, V.P., Comparative Evaluation of Frequency Distributions for Louisiana Extreme Rainfall, Water Resources Bulletin, 1993, 29(2): 211-219.

10. Rasel, M.M., and Hossain, S.M., Development of rainfall intensity duration frequency (R-IDF) equations and curves for seven divisions in Bangladesh, International Journal of Scientific and Engineering Research, 2015, 6(5): 96-101.

11. Sasireka, K., Suribabu, C.R., and Neelakantan, T.R., Extreme rainfall return periods using Gumbel and Gamma distributions, International Journal of Recent Technology and Engineering, 2019, 8(4): 27-29.

12. Sharma, M.A., and Singh, J.B., Use of probability distribution in rainfall analysis, New York Science Journal, 2010, 3(9): 40-49.

13. Subudhi, R., Probability analysis for prediction of annual maximum daily rainfall of Chakapada block of Kandhamal district in Orissa, Indian Journal of Soil Conservation, 2007, 35: 84-85.

14. United States Water Resources Council (USWRC), Guidelines for determining flood flow frequency, Bulletin No. 17B, 1981.

15. Vivekanandan, N., Modelling annual rainfall of Krishna and Godavari river basins using Extreme Value Type-1 distribution, i-manger Journal of Structural Engineering, 2014, 3(1): 7-12.

16. Zhang, J., Powerful goodness-of-fit tests based on the likelihood ratio, Journal of the Royal Statistical Society: Series B (Statistical Methodology), 2002, 64 (2): 281-294.