International Journal of Scientific Research in Mathematical and Statistical Sciences
Research Paper
Volume-7, Issue-5, pp.53-59, October (2020)
E-ISSN: 2348-4519

Comparison of Estimators of Five Probability Distributions for Estimation of Peak Flood

N. Vivekanandan

Central Water and Power Research Station, Pune, Maharashtra, India

Author's Mail: anandaan@rediffmail.com, Tel: +91-20-24103367

Available online at: www.isroset.org

Received: 04/Sept/2020, Accepted: 20/Oct/2020, Online: 31/Oct/2020

Abstract—Estimation of Peak Flood (PF) for a given return period is of utmost importance for planning, design and management of civil and hydraulic structures. Depending on the design-life of the structure, the estimated PF with a desired return period is used. This can be achieved through Flood Frequency Analysis (FFA) that involves fitting probability distribution to the series of observed Annual Peak Flood (APF) data. In this paper, a study on comparison of estimators of 2-parameter Log Normal, Log Pearson Type-3 (LP3), Generalized Extreme Value (GEV), Extreme Value Type-1 and Extreme Value Type-2 distributions adopted in FFA for river Tapi at Bhusawal and Savkheda sites is carried out. Based on the intended applications and the variate under consideration, standard parameter estimation procedures such as Method of Moments (MoM), Maximum Likelihood Method (MLM) and L-Moments (LMO) are used for determination of parameters of the probability distributions. The adequacy of probability distributions applied in FFA is quantitatively assessed by Goodness-of-Fit (GoF) tests viz., Chi-square and Kolmogorov-Smirnov, and diagnostic test using D-index. The paper presents the GEV (LMO) distribution is better suited for estimation of PF at Bhusawal whereas LP3 (MLM) for Savkheda.

Keywords—Chi-Square, D-index, Generalized Extreme Value, Kolmogorov-Smirnov, L-Moments, Log Pearson Type-3, Maximum Likelihood Method, Peak Flood

I. INTRODUCTION

Flooding is the most common hazard among the environmental hazards that impose damage in commercial buildings, roads, bridges, water supply systems and sewage disposal and agricultural lands. One of the most important objectives in Flood Frequency Analysis (FFA) is to estimate a design flood for a region. These flood estimates have been used to design the hydraulic structures viz., dams, culverts, channels and other structures that are used in flood management [1]. This can be achieved through Flood Frequency Analysis (FFA), which involves fitting probability distribution to the series of observed Annual Peak Flood (APF) data.

A number of probability distributions viz., Normal (NOR), 2-parameter Log Normal (LN2), 3-parameter Log Normal (LN3), Pearson Type-3 (PR3), Log Pearson Type-3 (LP3), Generalized Logistic (GLO), Generalized Extreme Value (GEV), Extreme Value Type-1 (EV1) and Extreme Value Type-2 (EV2) are widely applied in FFA [2]. However, by considering the purpose of the study and data availability, appropriate distribution will be applied for arriving at a design flood. Based on the intended applications and the variate under consideration, standard parameter estimation procedures viz., Method of Moments (MoM), Maximum Likelihood Method (MLM) and L-Moments (LMO) are used for determination of parameters of the distributions wherever applicable.

This research paper is arranged in the following manner. Section I contains the introduction on the essentiality of estimation of peak flood. The studies conducted by numerous researchers on flood estimation are presented in Section II. The methodology adopted in FFA is presented in Section III. Section IV describes the results of the data analysis and discussions made thereof. The conclusions and recommendations made from the study are presented in Sections V and VI respectively. The Section VII concludes the research work with future direction.

II. LITERATRUE REVIEW

A number of studies have been carried out by different researchers on FFA. Kjeldsen et al. [3] carried out the study on Regional FFA (RFFA) using LMO for KwaZulu-Natal Province of South Africa. Study by Kumar et al. [4] revealed that the GEV distribution is better suited for FFA for eight gauging sites of Middle Ganga Plains. Yue and Wang [5] determined the parameters of the distributions using LMO for modelling of annual stream flow in different climatic regions of Canada. Kumar and Chatterjee [6] employed the LMO to define homogenous regions within 13 gauging sites in Brahmaputra, India. Study by Modarres [7] revealed that the LMO approach is a suitable tool for Regional Frequency Analysis (RFA) for Isfahan Province. Bhuyan et al. [8] found that the RFFA based on the GEV distribution using level one LH-moment gives better results over LMO. Malekinezhad et al. [9] reported that the GEV (LMO) is better suited for modelling APF of three different regions in Iran. Sarhadi and Heydarizadeh [10] studied the RFA and temporal pattern characteristics of dry spells in Iran using LMO. They also described that the GLO, GEV and PR3 distributions are the best-fit for most of the regions. Markiewicz et al. [11] applied the generalized exponential and inverse Gaussian distributions in frequency analysis of annual maximum flows for Polish rivers. Kossi et al. [12] carried out the RFFA for Volta River Basin (VRB) using LMO of five probability distributions. They found that the GEV and GPA distributions are better suited to yield accurate flood quantiles in VRB. Mohammed and Azhar [13] derived the hydrometeorological approach to estimate the design flood at Kol Dam in the Satluj River Basin using Snyder's probable maximum flood hydrograph. Suhartano et al. [14] applied the NOR, LN2, LP3 and EV1 distributions to analyse the design flood by FFA in Lesti sub watershed. Ul Hassan et al. [15] adopted the GEV, GLO, LN3, PR3 and EV1 distributions in estimating the flood at five gauging sites of Torne River. Parvez and Inayathulla [16] analyzed the rainfall data using LN2 distribution and developed the intensity-duration-frequency curves for Upper Cauvery, Karnataka. However, when different distributions are adopted in FFA, a common problem that arises is how to determine which distribution fits best for a given set of data. This can be answered by quantitative and qualitative assessments; and the results are quantifiable and reliable [17]. For the quantitative assessment on APF data within the observed range, Goodness-of-Fit (GoF) tests viz., Chi-square (χ^2) and Kolmogorov-Smirnov (KS) are applied. In addition to GoF tests, a diagnostic test viz., D-index is used for the selection of best suitable probability distribution for estimation of PF. Qualitative assessment is made from the fitted curves of the estimated PF.

The literature thus presents a varied spectrum of applications of probability distributions and the parameter estimation methods that are used in FFA. In this paper, a study on comparison of MoM, MLM and LMO estimators of five probability distributions (viz., LN2, LP3, GEV, EV1 and EV2) adopted in FFA is presented with illustrative example for selection of an appropriate distribution for estimation of PF.

III. METHODOLOGY

The procedures involved in FFA are: (i) prepare the APF data series from the daily stream flow data series; (ii) determination of parameters of the distributions using MoM, MLM and LMO, and estimate the PF for different return periods; (iii) conduct quantitative and qualitative assessments; and (iv) analyse the results and made the suggestions thereof. Table 1 presents the Cumulative Distribution Function (CDF) and quantile estimator (x(T)) of probability distributions adopted in FFA.

Table 1. CDF and quantile estimator of probability distributions

Distribution	**CDF**	**Quantile Estimator (x(T))**
LN2	$F(x)=\phi\left(\frac{\ln(x)-\mu(y)}{\sigma(y)}\right)$	$x(T)=\exp\ (\mu(y)+\sigma(y)\,\phi^{-1}(P))$
LP3	$F(x)=\begin{cases} G\left(k,\frac{\ln(x)-\xi}{\alpha}\right) & ,\alpha>0 \\ 1-G\left(k,\frac{\ln(x)-\xi}{\alpha}\right) & ,\alpha<0 \end{cases}$	No explicit expression of the quantile function is available
GEV	$F(x)=\exp\left(-\left[1-\frac{k(x-\xi)}{\alpha}\right]^{1/k}\right)$, $\alpha>0$	$x(T)=\xi+\frac{\alpha}{k}\left(1-\left[-\ln(P)\right]^{k}\right)$
EV1	$F(x)=\exp\left(-\exp-\left(\frac{x-\xi}{\alpha}\right)\right)$, $\alpha>0$	$x(T)=\xi-\alpha\ln\left[-\ln(P)\right]$
EV2	$F(x)=\exp\left(-\left(\frac{x}{\alpha}\right)^{-k}\right)$, $\alpha>0$	$x(T)=\alpha\left[-\ln(P)\right]^{-(1/k)}$

In Table 1, ξ is the location parameter, α is the scale parameter, k is the shape parameter, $\mu(y)$ and $\sigma(y)$ are the average and standard deviation of the log transformed series of the observed data (i.e., y=ln(x)), F(x) is the CDF of x (i.e., Annual Peak Flood), ϕ^{-1} is the inverse of standard normal distribution function and defined by $\phi^{-1}=\left(P^{0.135}-(1-P)^{0.135}\right)/0.1975$ where in P is the probability of exceedance, G(...) is the incomplete gamma integral and x(T) is the estimated PF for a return period (T). A relation between F, P and T is defined by F(x)=1-P=1-1/T. The procedures involved in FFA viz., determination of the parameters, estimation of PF for different return periods and computation of Standard Error (SE) on the estimated PF using LN2, LP3, GEV, EV1 and EV2 distributions are briefly described in the text book titled 'Flood Frequency Analysis' by Rao and Hamed [18].

Estimation of Confidence Interval Limits

The lower and upper confidence limits (LCL and UCL) of the estimated PF at 95% level are computed from:
LCL=Estimated PF-1.96(SE) and
UCL= Estimated PF+(1.96)SE
wherein SE is the Standard Error of the estimated PF.

Goodness-of-Fit Tests

The theoretical descriptions of GoF tests (viz., χ^2 and KS) applied in checking the adequacy of fitting probability distributions to the APF data series are given as below:

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

where, $O_j(x)$ is the observed frequency value of x for j^{th} class, $E_j(x)$ is the expected frequency value of x for j^{th} class and NC is the number of frequency classes.

$$KS = \operatorname*{Max}_{i=1}^{N}\left(F_e(x_i) - F_D(x_i)\right) \quad \text{... (2)}$$

where, $F_e(x_i)=i/(N+1)$ is the empirical CDF of x_i wherein 'i' indicates the rank assigned to the observed data arranged in ascending order (i.e., $x_1<x_2<x_3<.....<x_N$), x_1 and x_N are the lowest highest values in the observed data series, $F_D(x_i)$ is the computed CDF of x_i by the probability distribution and N is the number of samples. The acceptance regions of χ^2 and KS tests statistic at the desired significance level (η) are given by $\chi^2_C \le \chi^2_{1-\eta,NC-m-1}$ and $KS_C \le KS_{1-\eta,N}$ respectively. Here, m denotes the number of parameters of the distribution, χ^2_C is the computed value of χ^2 statistic and KS_C is the computed value of KS statistic. *Test criteria*: If the computed values of GoF tests statistic given by the distribution are less than its theoretical values at the desired significance level then the distribution is acceptable for FFA at that level [19].

Diagnostic Test

The GoF tests results may not offer a conclusive inference in sometimes and thus a diagnostic test is applied for the selection of suitable probability distribution for estimation of PF. The theoretical description of D-index [20] is given as below:

$$D-\text{index} = \frac{1}{N}\sum_{i=1}^{6}\left|x_i - x_i^*\right| \quad \text{... (3)}$$

where, x_i is the observed value of i^{th} sample and x_i^* is the estimated value of i^{th} sample. For computation of D-index, the first six highest values in the observed APF data series and its corresponding estimated APF will be considered. The distribution with minimum D-index is identified as better-suited distribution in comparison with other distribution (or method) for estimation of PF.

Application

In this paper, a study on comparison of MoM, MLM and LMO estimators of five probability distributions (viz., LN2, LP3, GEV, EV1 and EV2) for estimation of PF for river Tapi at Bhusawal and Savkheda gauging sites is carried out. The Tapi river basin is situated in the northern part of the Deccan Plateau and extends over an area of 65.145 km that is about 2% of the total geographical area of the country and nearly 80% of the basin lies in the State of Maharashtra. The river basin lies between the longitudes of 72^o 38′ to 78^o 17″ E and latitudes of 20^o 05′ to 22^o 03″ N. The Bhusawal gauging site is located between the latitude of 75^0 46′ 56″ E and longitude of 21^o 03′ 54″ N with catchment area of 32478 km^2. The Savkheda gauging site is located between the latitude of 75^0 30′ 54″ E and longitude of 21^o 08′ 53″ N with catchment area of 48136 km^2. Daily stream flow data observed at Bhusawal for the period 1987 to 2016 and Savkheda for the period 1991 to 2016 (with some gaps) are used. The APF data series is extracted from the daily stream flow data series and used in FFA. From the scrutiny of the stream flow data of Savkheda, it is found that the observed data for the period of six years (viz., 1991, 2001, 2009, 2011, 2012 and 2015) are not available. However, by considering the importance of the observed extreme events at Savkheda, the data for the missing years are ignored while carrying out FFA. For Bhusawal site, there are no missing values in the observed data series. The descriptive statistics viz., Average, Standard Deviation (SD), Coefficient of Skewness (CS) and Coefficient of Kurtosis (CK) of the observed and log-transformed series of APF data are presented in Table 2.

Table 2. Descriptive statistics of APF data

Gauging site	Average (cumecs)	SD (cumecs)	CS	CK
Bhusawal	5625.7 (8.489)	3350.8 (0.548)	1.746 (0.028)	(3.474) 0.430
Savkheda	884.9 (6.427)	921.8 (0.829)	2.623 (0.538)	8.321 (-0.183)
Numbers given within the brackets indicates the descriptive statistics of the log transformed values of APF data				

IV. RESULTS AND DISCUSSIONS

By applying the procedures of FFA, as described above, the parameters of LN2, LP3, GEV, EV1 and EV2 distributions were determined by MoM, MLM and LMO, and are used for estimation of PF. The PF estimates for different return periods from 1.01-year to 100-year are presented in Tables 3 and 4 while the plots are shown in Figures 1 and 2.

For Bhusawal, it is found that the MLM is not feasible (NF) for determining the parameters of LP3 and hence the estimated PF using LP3 (MLM) is not presented in Table 3. Likewise, the estimated PFs using LP3 (LMO) are also not presented in Tables 3 and 4 due to non-existence of LMO for determination of parameters of LP3 distribution. From Tables 3 and 4, it is noted that the estimated PFs using EV2 (MLM) distribution for a return period from 20-year to 100-year are higher than those values of other probability distributions adopted in FFA. From Figures 1 and 2, it is noticed that the fitted curves of the PF estimates using LP3, GEV and EV2 distributions are in the form of exponential while the EV1 and LN2 curves are in the form of linear.

Analysis of Results Based on GoF Tests

For Bhusawal and Savkheda, the GoF tests statistic values of five probability distributions (using MoM, MLM and LMO) were computed and the results are presented in Tables 5 and 6. In the present study, number of frequency class was considered as 5 while computing the χ^2 test statistic. The theoretical value of χ^2 at 5% significance level was determined based on the degrees of freedom, viz., one for 3-parameter distributions (viz., LP3 and GEV) and two for 2-parameter distributions (viz., LN2, EV1 and EV2). From the analysis, it is found that the LP3 (using MLM and LMO) for Bhusawal while LP3 (LMO) for Savkheda are not feasible (NF) for FFA and hence the GoF tests results of LP3 (using MLM and LMO) are not presented in Tables 5 and 6. From GoF tests results, it is noticed that the computed values of GoF tests statistic using MoM, MLM and LMO estimators of five probability distributions (viz., LN2, LP3, GEV, EV1 and EV2) adopted in FFA are less than its corresponding theoretical values at 5% significance level, and at this level, all five distributions are found to be adequate for FFA.

Table 3. Estimated Peak Flood (cumecs) using LN2, LP3, GEV, EV1 and EV2 distributions for Bhusawal

Return period (year)	LN2			LP3			GEV			EV1			EV2		
	MoM	MLM	LMO	MoM	MLM	LMO	MoM	MLM	LMO	MoM	MLM	LMO	MoM	MLM	LMO
1.01	1372.5	1357.6	1362.8	1355.0	NF	NF	763.0	1309.0	1396.1	420.1	918.9	422.4	2663.3	1637.3	2155.6
2	4854.6	4859.8	4859.8	4859.8	NF	NF	4946.6	4848.0	4753.4	5075.1	5061.5	5105.3	4795.1	4513.0	4458.5
5	7667.0	7708.5	7697.9	7708.5	NF	NF	7779.3	7527.2	7461.0	8037.4	7538.2	7905.0	6815.3	8274.2	6884.5
10	9735.8	9810.7	9790.1	9810.6	NF	NF	9806.3	9583.0	9623.0	9998.7	9177.9	9758.6	8601.5	12360.1	9179.1
20	11859.0	11972.5	11940.3	11972.5	NF	NF	11872.2	11796.8	12025.4	11880.0	10750.8	11536.7	10753.3	18164.2	12095.8
25	12560.5	12687.6	12651.3	12687.5	NF	NF	12553.7	12553.1	12862.7	12476.8	11249.8	12100.7	11542.5	20523.5	13202.2
50	14807.3	14980.5	14930.2	14980.4	NF	NF	14735.6	15061.3	15696.2	14315.2	12786.8	13838.2	14356.7	29897.0	17288.0
75	16329.8	16536.2	16475.8	16377.0	NF	NF	16063.3	16652.0	17536.1	15383.8	13680.2	14848.1	16297.9	37204.0	20221.2
100	17169.0	17394.2	17328.2	17394.7	NF	NF	17030.4	17841.4	18932.3	16140.0	14312.5	15562.8	17828.3	43430.9	22593.3

Table 4. Estimated Peak Flood (cumecs) using LN2, LP3, GEV, EV1 and EV2 distributions for Savkheda

Return period (year)	LN2			LP3			GEV			EV1			EV2		
	MoM	MLM	LMO	MoM	MLM	LMO	MoM	MLM	LMO	MoM	MLM	LMO	MoM	MLM	LMO
1.01	87.6	89.9	86.1	125.0	149.9	NF	292.7	286.7	265.8	629.7	472.0	414.8	121.3	151.3	180.4
2	620.9	618.6	618.6	574.4	559.1	NF	673.4	587.6	574.6	733.4	726.5	754.9	687.2	538.8	530.5
5	1261.0	1242.8	1262.3	1205.6	1161.8	NF	1402.9	1063.3	1151.5	1548.3	1263.6	1454.2	1082.7	1151.6	1010.8
10	1826.1	1789.8	1832.7	1855.8	1810.0	NF	1961.9	1484.4	1723.4	2087.9	1619.2	1917.2	1462.8	1904.0	1548.9
20	2479.3	2418.8	2493.5	2712.6	2697.6	NF	2563.1	1992.2	2478.1	2605.5	1960.4	2361.3	1952.2	3084.1	2332.7
25	2710.3	2640.6	2727.5	3042.0	3047.5	NF	2768.3	2178.6	2771.2	2769.6	2068.6	2502.2	2139.4	3593.9	2656.2
50	3497.8	3394.8	3526.2	4267.1	4384.2	NF	3448.4	2843.6	3880.5	3275.4	2401.9	2936.2	2836.5	5757.8	3962.9
75	4070.9	3942.0	4108.3	5149.0	5376.7	NF	3879.3	3302.7	4699.3	3569.4	2595.7	3188.5	3341.8	7572.3	5000.5
100	4399.7	4255.4	4442.5	5861.4	6194.7	NF	4201.3	3664.8	5373.3	3777.4	2732.8	3367.0	3752.9	9192.5	5895.2

Table 5. Theoretical and computed values of GoF tests statistic of LN2, LP3, GEV, EV1 and EV2 distributions for Bhusawal

Probability distribution	Theoretical value at 5% level		Computed values of GoF tests statistic					
			χ^2			KS		
	χ^2	KS	MoM	MLM	LMO	MoM	MLM	LMO
LN2	5.990	0.231	0.667	0.667	0.667	0.094	0.094	0.094
LP3	3.840	0.231	0.667	---	---	0.089	---	---
GEV	3.840	0.231	0.667	0.667	0.667	0.089	0.089	0.089
EV1	5.990	0.231	1.667	1.333	1.333	0.095	0.086	0.081
EV2	5.990	0.231	1.667	0.667	0.667	0.150	0.115	0.090

Table 6. Theoretical and computed values of GoF tests statistic of LN2, LP3, GEV, EV1 and EV2 distributions for Savkheda

Probability distribution	Theoretical value at 5% level		Computed values of GoF tests statistic					
			χ^2			KS		
	χ^2	KS	MoM	MLM	LMO	MoM	MLM	LMO
LN2	5.990	0.280	1.200	1.200	1.200	0.125	0.113	0.113
LP3	3.840	0.280	4.800	1.750	---	0.096	0.087	---
GEV	3.840	0.280	4.800	1.800	1.600	0.150	0.123	0.098
EV1	5.990	0.280	4.800	1.800	1.600	0.178	0.186	0.193
EV2	5.990	0.280	4.800	2.221	2.000	0.278	0.198	0.124

Analysis of Results Based on Diagnostic Test

In addition to GoF tests, for identifying a best suitable distribution for estimation of PF amongst five probability distributions adopted in FFA, second line of action, say, diagnostic test (i.e., D-index) was applied. For Bhusawal and Savkheda, the D-index values for LN2,LP3, GEV, EV1 and EV2 distributions were computed and are presented in Table 7.

Table7. D-index values given by LN2, LP3, GEV, EV1 and EV2

Probability distribution	Bhusawal			Savkheda		
	MoM	MLM	LMO	MoM	MLM	LMO
LN2	1.377	1.457	1.417	2.458	2.884	2.675
LP3	1.376	---	---	2.263	2.481	---
GEV	1.435	1.438	1.315	2.483	4.152	2.899
EV1	1.523	1.956	1.617	2.838	3.505	2.683
EV2	2.138	2.639	1.375	4.120	2.074	3.868

From diagnostic test results, it is noticed that the D-index value of GEV (LMO) is minimum when compared to the corresponding values of other distributions for Bhusawal. For Savkheda, it is noticed that the D-index values of EV2 (MLM), LP3 (MoM), LN2 (MoM) and LP3 (MLM) are the first, second, third and fourth minimum in the order of the magnitude when compared with the D-index values of other distributions. The diagnostic test results of LP3 (using MLM and LMO) are not presented in Table 7 because of non-feasibility for fitting LP3 (MLM and LMO) to APF data of Bhusawal while LP3 (LMO) for Savkheda.

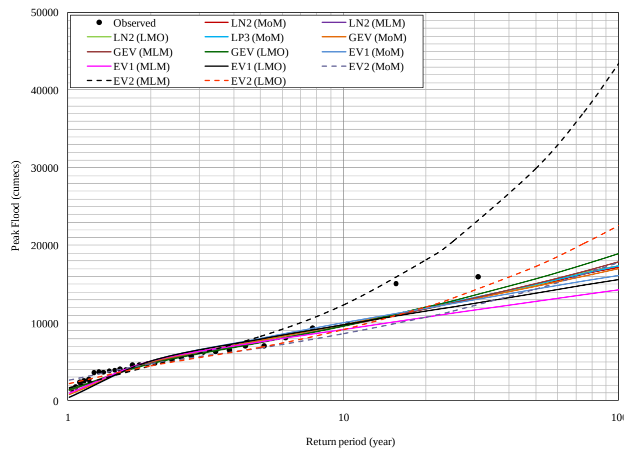

Figure 1. Plots of estimated PF using LN2, LP3, GEV, EV1 and EV2 distributions with observed APF for Bhusawal

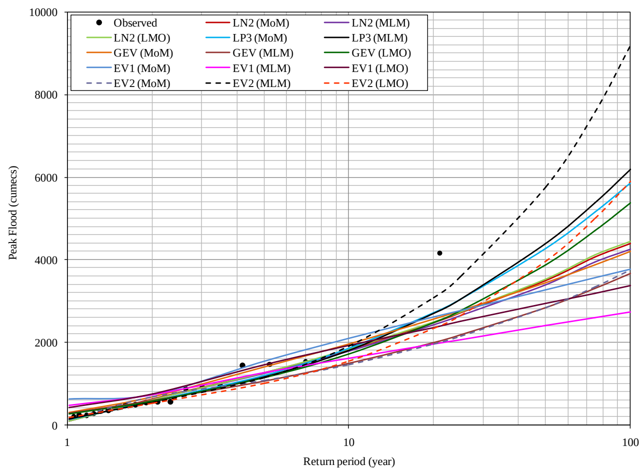

Figure 2. Plots of estimated PF using LN2, LP3, GEV, EV1 and EV2 distributions with observed APF for Savkheda

Selection of Probability Distribution

Based on FFA results obtained from quantitative assessment by using GoF and diagnostic tests, it is observed that the analysis offered diverging inferences and thus called for qualitative assessment using the plots of the estimated PFs (Figures 1 and 2). Hence, the best fit for estimation of PF is re-assessed through fitted curves of the estimated PFs together with D-index values; and accordingly final selection is made.

i) For Bhusawal, it is noted that the estimated PF using GEV (LMO) is higher than those values of GEV (using MoM and MLM).
ii) For Savkheda, it is noted that the estimated PF using LP3 (MoM) is less than those values of LP3 (MLM) while LN2 (LMO) gave higher estimates than LN2 (MoM). This indicates that the MoM estimators are less accurate when compared with MLM and LMO when different distributions are applied in FFA.
iii) Also, for Savkheda, it is noted that the fitted line (Figure 2) of the estimated PFs using EV2 (MLM) are well above the observed data and thus there is no good correlation between the observed and estimated values.
iv) So, after eliminating the D-index values of EV2 (MLM), LP3 (MoM) and LN2 (MoM) from the selection, it is identified that the D-index value of LP3 (MLM) is the next minimum in the array of D-index values and hence LP3 (MLM) is considered as the best choice for estimation of PF for Savkheda.

Hence, the qualitative assessment (plots of FFA results) of the outcomes is weighed with D-index values and accordingly GEV (LMO) distribution is found to be best fit for estimation of PF at Bhusawal and LP3 (MLM) for Savkheda. The estimated PF with 95% confidence limits by the selected probability distribution (i.e., GEV (LMO) for Bhusawal and LP3 (MLM) for Savkheda) for different return periods from 1.01-year to 100-year are presented in Table 8 while the plots are shown in Figures 3 and 4. The percentages of the observed APF falls within the confidence limits of the estimated PFs using viz., GEV (LMO) for Bhusawal and LP3 (MLM) for Savkheda are about 97% and 95% respectively

Table 8. Estimated PFs with 95% confidence limits for different return periods using GEV (LMO) for Bhusawal and LP3 (MLM) for Savkheda

Return period (year)	Bhusawal			Savkheda		
	Estimated PF (cumecs)	95% confidence limits		Estimated PF (cumecs)	95% confidence limits	
		Lower	Upper		Lower	Upper
1.01	1396.1	275.4	2516.8	149.9	50.0	249.8
2	4753.4	3200.1	6306.7	559.1	350.2	768.0
5	7461.0	5525.3	9396.7	1161.8	749.9	1573.7
10	9623.0	7000.4	12245.6	1810.0	1050.3	2569.7
20	12025.4	8249.0	15801.8	2697.6	1399.7	3995.5
25	12862.7	8652.1	17073.3	3047.5	1575.4	4519.6
50	15696.2	9651.5	21740.9	4384.2	2249.8	6518.6
75	17536.1	10248.9	24823.3	5376.7	2850.5	7902.9
100	18932.3	10749.6	27115.0	6194.7	3250.1	9139.3

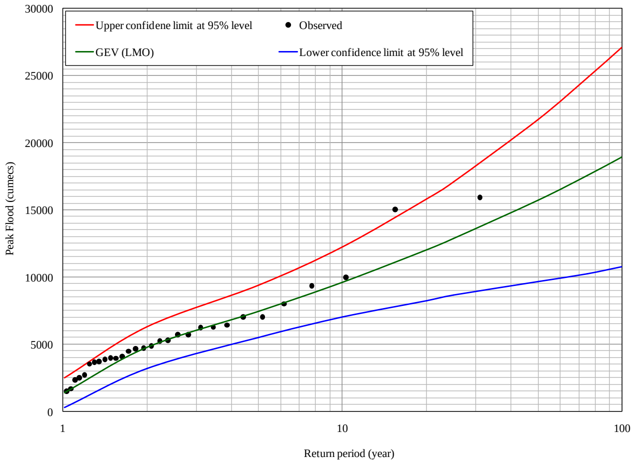

Figure 3. Plots of estimated PF with 95% confidence limits using GEV (LMO) and observed APF for Bhusawal

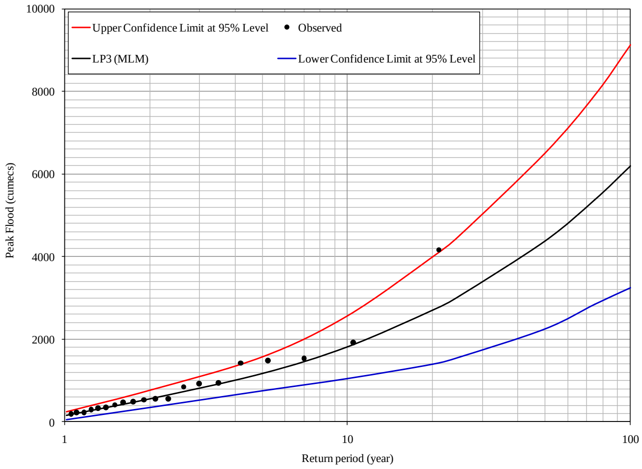

Figure 4. Plots of estimated PF with 95% confidence limits using LP3 (MLM) and observed APF for Savkheda

V. CONCLUSIONS

This paper described a study on comparison of MoM, MLM and LMO estimators of LN2, LP3, GEV, EV1 and EV2 distributions adopted in FFA for estimation of PF for river Tapi at Bhusawal and Savkheda sites. The adequacy of fitting probability distributions was checked by quantitative (viz., GoF and diagnostic tests) and qualitative (viz., fitted curves of the estimated values) assessments. Based on the results of the data analysis, the following conclusions were drawn from the study:

i) χ^2 and KS test results supported the use of LN2, LP3, GEV, EV1 and EV2 (using MoM, MLM and LMO) for FFA for Bhusawal and Savkheda.
ii) Qualitative assessment (plots of the estimated values) of the outcomes was weighed with D-index values and accordingly GEV (LMO) distribution was found to be better suited for estimation of PF for Bhusawal whereas LP3 (MLM) for Savkheda.
iii) The qualitative assessment also indicated that about 97% of the observed APF of Bhusawal and 95% of APF of Savkheda were within the confidence limits of the estimated PFs using GEV (LMO) and LP3 (MLM) respectively.

VI. RECOMMENDATIONS

The study recommended that the PFs given by GEV (LMO) for Bhusawal and LP3 (MLM) for Savkheda could be used for design purposes. By considering the data length (i.e., 30 years for Bhusawal and 20 years for Savkheda) of the observed APF, the study indicated that the estimated PFs for a return period beyond 100-year may be cautiously used due to uncertainty in higher order return periods.

VII. FUTURE WORK

The study presented in this paper could be further enhanced by adopting Extreme Value, Normal and Gamma families of probability distributions with voluminous data to arrive at a suitable PF for the design purposes of civil and hydraulic structures in Bhusawal and Savkheda sites.

ACKNOWLEDGMENT

The author is grateful to the Director, Central Water and Power Research Station, Pune, for providing the research facilities to carry out the study.

REFERENCES

[1] B. Naghavi, F.X. Yu and V.P. Singh, "Comparative evaluation of frequency distributions for Louisiana extreme rainfall", *Water Resources Bulletin*, Vol.**29**, No.**2**, pp.**211-219**, **1993**.

[2] N. Vivekanandan, "A comparative study on Gumbel and LP3 probability distributions for estimation of extreme rainfall", *International Journal of Water Resources Engineering*, Vol. **6**, No.**1**, pp. **21-33**, **2020**.

[3] T.R. Kjeldsen, J.C. Smithers and R.E. Schulze, "Regional flood frequency analysis in the KwaZulu-Natal province, South Africa, using the index-flood method", *Journal of Hydrology*, Vol.**255**, Nos.**1-4**, pp.**194–211**, **2002**.

[4] R. Kumar, C. Chatterjee, S. Kumar, A.K. Lohani and R.D. Singh, "Development of regional flood frequency relationships using L-moments for Middle Ganga Plains Subzone 1(f) of India", *Water Resources Management*, Vol.**17**, No.**4**, pp. **243–257**, **2003**.

[5] S. Yue and C.Y. Wang, "Possible regional probability distribution type of Canadian annual stream flow by L-moments", *Water Resources Management*, Vol.**18**, No.**5**, pp. **425–438**, **2004**.

[6] R. Kumar and C. Chatterjee, "Regional flood frequency analysis using L-Moments for north Brahmaputra region of India", *Hydrologic Engineering*, Vol.**10**, No.**1**, pp. **1–7**, **2005**.

[7] R. Modarres, "Regional precipitation climates of Iran", *Journal of Hydrology (NZ)*, Vol.**45**, No.**1**, pp.**13-27**, **2006**.

[8] A. Bhuyan, M. Borah and R. Kumar, "Regional flood frequency analysis of North-Bank of the River Brahmaputra by using LH-Moments", *Water Resources Management*, Vol.**24**, No.**9**, pp.**1779-1790, 2010**.

[9] H. Malekinezhad, H.P. Nachtnebel and A. Klik, *Regionalization approach for extreme flood analysis using L-moments, Agricultural Science and Technology (Iran)*, Vol.**13** (Supplementary Issue), pp.**1183–1196, 2011.**

[10] A. Sarhadi and M. Heydarizadeh, "Regional frequency analysis and spatial pattern characterization of Dry Spells in Iran", *International Journal of Climatology*, Vol.**34**, No.**3**, pp. **835–848**, **2014**.

[11] I. Markiewicz, W.G. Strupczewski, E. Bogdanowicz and K. Kochanek, "Generalized exponential distribution in flood frequency analysis for Polish Rivers", *PLoS ONE*, Vol.**10**, No.**12**, pp.**1-15**, **2015**.

[12] K. Kossi, A.A. Barnabas, D. Bernd and C.C.H. Fabien, "Regional flood frequency analysis in the Volta river basin, West Africa", *Hydrology Journal*, Vol.**3**, No.**1**, pp.**1-15**, **2016**.

[13] S. Mohammed and H. Azhar, "Estimation of design flood at Kol dam using hydrometeorological approach", *International Journal of Environmental Sciences and Natural Resources*, Vol.**4**, No.**1**, pp.**5-10**, **2017**.

[14] E. Suhartanto, M.L. Lily, N. Dina, I.H. Febri and A.K. Dwi, "Estimation of design flood with four frequency analysis distributions", *Asian Journal of Applied Science and Technology*, Vol.**2**, No.**1**, pp.**13-27**, **2018**.

[15] M. Ul Hassan, O. Hayat and Z. Noreen, "Selecting the best probability distribution for at-site flood frequency analysis; a study of Torne River", *SN Applied Science*, Vol.**1**, No.**12**, pp.**1-10**, **2019**.

[16] M.B. Parvez and M. Inayathulla, "Statistical Analysis of Rainfall for Development of Intensity-Duration-Frequency Curves for Upper Cauvery, Karnataka by Log-Normal Distribution", International Journal of Scientific Research in Mathematical &Statistical Sciences, Vol.**6**, No.**5**, pp.**12-33**, **2019**.

[17] P.E. Charles Annis, "Goodness-of-Fit tests for statistical distributions", **2009**.

[18] A.R. Rao and K.H. Hamed, "Flood frequency analysis", 2000.

[19] J. Zhang, "Powerful goodness-of-fit tests based on the likelihood ratio", *Journal of the Royal Statistical Society (Statistical Methodology)*, Vol.**64**, No.**2**, pp.**281-294**, **2002**.

[20] USWRC, "Guidelines for determining flood flow frequency", United States Water Resources Council (USWRC), Bulletin No. 17B (Revised), Washington, DC, New York, **1982**.

AUTHOR PROFILE

N. Vivekanandan has post graduated in Mathematics from Madurai Kamaraj University. In addition, he has received M.E.(Hydrology) from University of Roorkee, M.Phil.(Mathematics) from Bharathiar University, and M.B.A.(Human Resources) from Manonmaniam Sundaranar University. He has published more than 50 technical reports and 225 research papers in national and international journals & conferences. His main research work focuses on hydrology, water resources, irrigation management, soft computing, etc. He has 2 years of industrial experience, 1 year of teaching experience and 27 years of research experience.