2015

Electronic International Interdisciplinary Research Journal (EIIRJ)

REVIEWED INTERNATIONAL JOURNAL

ISSN: 2277-8721

Impact factor: 2.085

Bi-Monthly

VOL - IV

ISSUES - V

Chief-Editor: Ubale Amol Baban

Electronic International Interdisciplinary Research Journal (EIIRJ)
Bi-monthly
Vol IV Issues V
Reviewed Journal

Impact Factor: 2.085

ISSN 2277-8721

Sep-Oct 2015

COMPARISON OF RAINFALL ESTIMATES OF EV1 AND LP3 DISTRIBUTIONS FOR ESTIMATION OF PEAK FLOOD DISCHARGE USING RATIONAL METHOD

N. Vivekanandan

Central Water and Power Research Station

Pune 411024, Maharashtra

Abstract

Estimation of Peak Flood Discharge (PFD) at a desired location on a river is important for planning, design and management of hydraulic structures. For ungauged basin, rainfall depth becomes an important input in derivation of PFD. So, rainfall depth can be estimated through Extreme Value Analysis (EVA) by fitting of probability distribution to the series of recorded rainfall data. In the present study, Extreme Value Type-1 (EV1) and Log Pearson Type-3 (LP3) distributions are adopted for estimation of rainfall at Dhaulakuan rain gauge station. Maximum Likelihood Method (MLM) is used for determination of parameters of the distributions. Goodness-of-Fit tests such as Anderson-Darling and Chi-square are applied for checking the adequacy of fitting of probability distributions to the recorded data. Root mean square error is used for the selection of most suitable distribution for estimation of rainfall. The estimated 1-day maximum rainfall obtained from EV1 distribution at Dhaulakuan station is used to compute 1-hour maximum value of distributed rainfall adopting CWC guidelines. The study suggests the estimated PFD from rational method could be used for design of flood in river protection works of Yamuna river basin.

Keywords: *Anderson-Darling, Chi-square, Extreme Value, Log Pearson, Peak flood, Rainfall*

INTRODUCTION

Estimation of Peak Flood Discharge (PFD) at a desired location on a river is required for effective flood-plain management, and for efficient design of attenuation storages, bridges, culverts, embankments, and flood-protection structures. Statistical method is effective tool for obtaining PFD and their associated probabilities on gauged stream. However, most of the small and urban catchments are classified as ungauged wherein urbanization and infra-structural development take place. In practice, most hydraulic structures to control runoff to pre-development levels are installed in small catchments which require estimation of PFD. Methods such as rational, USGS, NRCS, Unit hydrograph, FEH, FSSR, etc., are widely used for estimation of PFD. In case of large river basins, the hydrological and stream flow series of a significant duration are generally available. However, in the present study, rational method is used for estimation of PFD. For ungauged basins, more data is not available other than rainfall (NIH, 2011). The rainfall data is also of shorter duration and may pertain to a neighbouring basin. Rainfall depth thus becomes an important input in derivation of PFD (Singh et al., 2001). For arriving at such design values, Extreme Value Analysis (EVA) of rainfall is carried out.

Out of number of probability distributions, Extreme Value Type-1 (EV1) and Log Pearson Type-3 (LP3) are widely applied for EVA of rainfall and stream flow (Casas et al., 2011; Lee et al, 2012; Daneshfaraz et al., 2013). Olumide et al. (2013) applied normal and EV1 distributions for prediction of rainfall and runoff at Tagwai dam site in Minna, Nigeria. They have also expressed that the normal distribution is better suited for rainfall prediction while Log-Gumbel for runoff. Izinyon and Ajumuka (2013) carried out Flood Frequency Analysis (FFA) for three tributaries of upper Benue river basin, Nigeria adopting Log-normal, EV1 and LP3 distributions. Das and Qureshi (2014) evaluated the probability distributions of GEV, LP3 and LN2 adopted in FFA through D-index and found that the LP3 is better suited distribution for estimation of PFD for Jiya Dhol river basin. Rasel and Hossain (2015) applied EV1 distribution for development of intensity duration frequency curves for seven divisions in Bangladesh. In view of the above, EV1 and LP3 distributions are used in the present study. Parameters of the distributions are determined by Maximum Likelihood Method (MLM) and used to estimate 1-day maximum rainfall. For quantitative assessment on rainfall data within the recorded range, Goodness-of-Fit (GoF) tests such as Anderson-Darling (A^2) and Chi-square (χ^2) are applied. Diagnostic index, namely, Root Mean Square Error (RMSE) is used for the selection of most suitable distribution for estimation of rainfall. The 1-hour maximum value of distributed rainfall is computed from 1-day maximum rainfall and used as an input to estimate PFD. The methodology adopted in EVA of rainfall using probability distributions, computation of GoF tests statistic and diagnostic index, and estimation of PFD using rational formula are briefly described in the following sections.

METHODOLOGY

The study is to estimate PFD of ungauged catchments of Yamuna River Basin (YRB). Thus, it is required to process and validate the data series of 1-day maximum rainfall for various application such as (i) determine the parameters of EV1 and LP3 distributions using MLM; (ii) assess the adequacy of fitting of EV1 and LP3 distributions to the series of AMR using GoF tests; (iii) selection of most suitable probability distribution for rainfall estimation using diagnostic index; (iv) compute 1-hour maximum value of distributed rainfall from the 1-day maximum rainfall given by probability distribution; (v) estimate PFD using rational method and (v) analyse the results obtained thereof.

PDF and CDF of Probability Distributions

The PDF (f(X)) and Cumulative Distribution Function (CDF; F(X)) of EV1 distribution is given as below:

$$\left.\begin{aligned} f(X) &= \frac{e^{-(X-\alpha)/\beta} e^{-e^{-(X-\alpha)/\beta}}}{\beta} \\ F(X) &= e^{-e^{-(X-\alpha)/\beta}}, \beta > 0 \end{aligned}\right\} \quad \text{... (1)}$$

where, α and β are the location and scale parameters of the distribution (Gumbel, 1960). The parameters are computed by MLM through Equations (2) and (3), and used to estimate the rainfall (X_T) for different return periods from $X_T = \alpha + Y_T\beta$, where $Y_T = -\ln(-\ln(1-(1/T)))$.

$$\alpha = -\beta \ln\left[\sum_{i=1}^{N} \exp(-X_i/\beta) \Big/ N\right] \quad \text{... (2)}$$

$$\beta = \overline{X} - \left[\sum_{i=1}^{N} X_i \exp(-X_i/\beta) \Big/ \sum_{i=1}^{N} \exp(-X_i/\beta)\right] \quad \text{... (3)}$$

$$SE(X_T) = \left(\beta/\sqrt{N}\right)\left(1.15894 + 0.19187 Y_T + 1.1 Y_T^2\right)^{0.5} \quad \text{... (4)}$$

where, X_i is the recorded AMR of i^{th} sample and $\overline{X}$ is the average value of AMR.

The PDF and CDF of Log Pearson Type-3 (LP3) distribution is given by:

$$f(X;\alpha,\beta,\gamma) = \frac{1}{\beta X \Gamma\gamma}\left(\frac{\ln(X)-\alpha}{\beta}\right)^{\gamma-1} e^{-\left(\frac{\ln(X)-\alpha}{\beta}\right)}, \ \beta,\gamma>0 \text{ and } F(R) = \int f(X;\alpha,\beta,\gamma)dX \quad \text{... (5)}$$

where, α, β and γ are the location, scale and shape parameters of the LP3 distribution (Rao and Hameed 2000). The parameters are computed by MLM through Equation (6) and used to estimate the

Electronic International Interdisciplinary Research Journal (EIIRJ) ISSN 2277-8721
Bi-monthly
Vol IV Issues V
Reviewed Journal **Impact Factor: 2.085** **Sep-Oct 2015**

rainfall (X_T) for different return periods from $X_T = \mathrm{Exp}((\alpha+\beta\gamma)+K_P\beta\sqrt{\alpha})$. Here, K_P is the frequency factor corresponding to the probability of exceedance and coefficient of skewness based on the log transformed series of the recorded data for LP3.

$$\left.\begin{array}{l}\sum_{i=1}^{N}(\ln(X_i)-\alpha)=N\beta\gamma \\ N\psi(\gamma)=\sum_{i=1}^{N}\ln[(\ln(X_i)-\alpha)/\beta] \\ N=\beta(\gamma-1)\sum_{i=1}^{N}[1/(\ln(X_i)-\alpha)]\end{array}\right\} \quad \ldots (6)$$

where, $\psi(\gamma)=\Gamma'(\gamma)/\Gamma(\gamma)$ is called a digamma function. The SE of estimated rainfall adopting LP3 distribution (using MLM) is computed from Equation (7) and given by:

$$SE(X_T)=\frac{\beta}{N}\left\{\begin{array}{l}1+KC_S+\left(\frac{K^2}{2}\right)\left(\frac{3C_S^2}{4}+1\right)+3K\frac{\partial K}{\partial C_S}\left(\frac{C_S^3}{4}+C_S\right)+ \\ 3\left(\frac{\partial K}{\partial C_S}\right)^2\left(2+3C_S^2+5C_S+\frac{5C_S^4}{9}\right)\end{array}\right\} \quad \ldots (7)$$

where, $\frac{\partial K}{\partial C_S}\cong\left(\frac{Z^2-1}{6}\right)+\frac{4(Z^3-6Z)}{6^3}C_S-\frac{3(Z^2-1)}{6^3}C_S^2+\frac{4Z}{6^4}C_S^3-\frac{10}{6^6}C_S^4$ and Z is the standard normal variate.

The lower and upper confidence limits (LCL and UCL) of the estimated rainfall values are obtained from $LCL=ER-(1.96*SE(X_T))$ and $UCL=ER+(1.96*SE(X_T))$. Here, ER is the Estimated Rainfall (X_T) and $SE(X_T)$ is the Standard Error of the estimated rainfall.

Goodness-of-Fit Tests

GoF tests are essential for checking the adequacy of probability distribution to the recorded series of AMR for estimation of rainfall. Out of a number GoF tests available, the widely accepted GoF tests are A^2 and χ^2, which are used in the study. The theoretical descriptions of GoF tests statistic (Charles Annis, 2009) are as follows:

A^2 statistic:

$$A^2=(-N)-(1/N)\sum_{i=1}^{N}\{(2i-1)\ln(Z_i)+(2N+1-2i)\ln(1-Z_i)\} \quad \ldots (8)$$

Here, $Z_i=F(X_i)$ for i=1,2,3,…,N with $X_1<X_2<\ldots<X_N$, $F(X_i)$ is the CDF of i^{th} sample (X_i) and N is the sample size.

χ^2 statistic:

$$\chi^2 = \sum_{j=1}^{NC} \frac{\left(O_j(X) - E_j(X)\right)^2}{E_j(X)} \qquad \text{... (9)}$$

where, $O_j(X)$ is the observed frequency value of j^{th} class, $E_j(X)$ is the expected frequency value of j^{th} class and NC is the number of frequency class (Zhang, 2002). The rejection region of χ^2 statistic at the desired significance level (η) is given by $\chi_C^2 \geq \chi^2_{1-\eta, NC-m-1}$. Here, m denotes the number of parameters of the distribution and χ_C^2 is the computed value of χ^2 statistic by PDF.

Test criteria: If the computed values of GoF tests statistic given by the distribution are less than that of the theoretical values at the desired significance level then the distribution is found to be acceptable for modelling the series of rainfall data.

Diagnostic Index

Diagnostic index, RMSE is used for the selection of most suitable probability distribution for rainfall estimation (Jalal et al., 2012), which is defined by:

$$\text{RMSE}=\left(\frac{1}{N}\sum_{i=1}^{N}\left(X_i - X_i^*\right)^2\right)^{0.5} \qquad \text{... (10)}$$

where X_i^* is the estimated AMR of i^{th} sample by probability distribution. The distribution has the least RMSE value is considered as the most suitable distribution for rainfall estimation.

APPLICATION

In this paper, a study on estimation of PFD for different return periods for six ungauged catchments of YRB is carried out. The Annual 1-day Maximum Rainfall (AMR) recorded at Dhaulakuan rain gauge station for the period 1998 to 2007, as presented in Figure 1, is used. The descriptive statistics such as average rainfall ($\overline{X}$), standard deviation, coefficient of variation, coefficient of skewness and coefficient of kurtosis are computed as 174.7 mm, 64.9 mm, 37.2 % , 0.247 and -0.209 respectively.

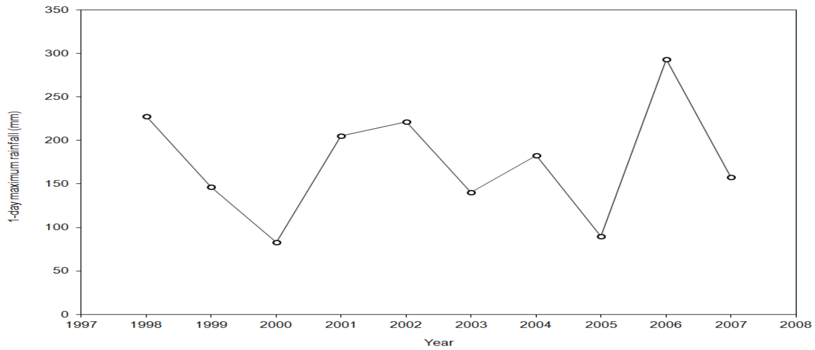

Figure 1: Plot of AMR recorded at Dhaulakuan rain gauge station

RESULTS AND DISCUSSIONS

By applying the procedures of EV1 and LP3 distributions, parameters were determined by MLM and used for estimation of 1-day maximum rainfall for different return periods. Table 1 gives the 1-day maximum rainfall estimates with confidence limits for different return periods adopting EV1 and LP3 distributions.

Table 1: 1-day maximum rainfall estimates with 95% confidence limits using EV1 and LP3 distributions for Dhaulakuan station

Return period (year)	EV1				LP3			
	ER (mm)	SE (mm)	Confidence limits		ER (mm)	SE (mm)	Confidence limits	
			LCL	UCL			LCL	UCL
2	164.6	20.3	124.7	204.4	165.6	20.5	125.5	205.7
5	226.6	31.2	165.5	287.8	224.7	30.9	164.1	285.3
10	267.7	40.0	189.3	346.1	261.3	39.1	184.7	337.8
15	290.9	45.2	202.2	379.6	280.8	43.7	195.2	366.4
20	307.1	49.0	211.2	403.1	294.6	47.0	202.6	386.6
25	319.6	51.9	218.0	421.3	304.9	49.5	207.9	401.8
50	358.2	60.9	238.8	477.5	335.6	57.1	223.8	447.5
75	380.5	66.2	250.8	510.3	354.1	61.6	233.4	474.8
100	396.4	70.0	259.3	533.5	365.1	64.4	238.8	491.4

From Table 1, it may be noted that the estimated rainfall obtained from EV1 distribution is consistently higher than the corresponding values of LP3 for return period 5-years above. By using the rainfall estimates obtained from EV1 and LP3 distributions, the rainfall frequency curves were developed and presented in Figures 2 and 3.

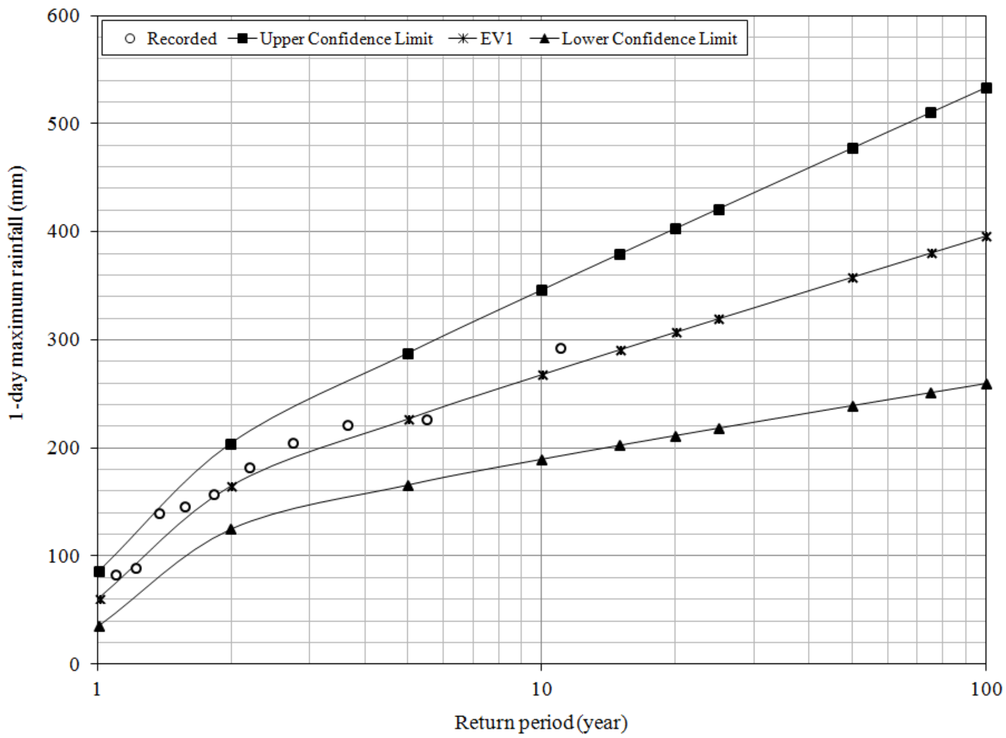

Figure 2: Plots of recorded and estimated 1-day maximum rainfall using EV1 distribution with 95% confidence limits for Dhaulakuan station

Electronic International Interdisciplinary Research Journal (EIIRJ)
Bi-monthly
Vol IV Issues V
Reviewed Journal
Impact Factor: 2.085
ISSN 2277-8721
Sep-Oct 2015

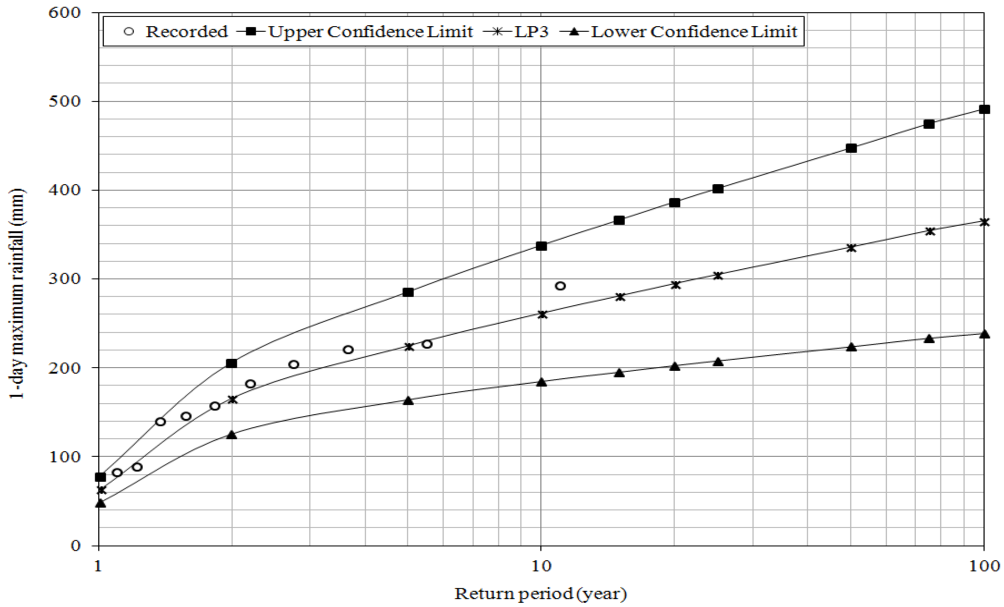

Figure 3: Plots of recorded and estimated 1-day maximum rainfall using LP3 distribution with 95% confidence limits for Dhaulakuan station

From Figures 2 and 3, it can be seen that the recorded AMR data are falling between the confidence limits of the estimated 1-day maximum rainfall obtained from EV1 and LP3 distributions.

Analysis Based on GoF Tests

The adequacy of fitting of EV1 and LP3 distributions for EVA of rainfall was performed by adopting GoF tests (A^2 and χ^2) and the results are presented in Table 2.

Table 2: Computed and theoretical values of GoF tests

Distribution	Computed value		Theoretical value	
	A^2	χ^2	A^2	χ^2
EV1	0.269	0.800	0.757	5.991
LP3	0.265	0.200	0.757	5.991

From Table 2, it may be noted that the computed values are not greater than the theoretical values at 5% significance level, and at this level, the GoF tests results supported the EV1 and LP3 distributions for modelling the series of AMR.

Analysis Based on Diagnostic Index

The selection of most suitable probability distribution amongst EV1 and LP3 distributions was made through RMSE. The RMSE value was computed as 14.1 mm for EV1 and 15.5 mm for LP3. On the basis of diagnostic index, EV1 distribution was found to be better suited probability distribution for estimation of rainfall.

Computation of PFD for Ungauged Catchments

It was required to estimate PFD for six ungauged catchments of YRB. The size of the catchment area is presented in Table 3. From an observation of catchment size and at the Google Earth of the region of these catchments, it was estimated that these are small catchments that respond quickly to rainfall, t_c (time of concentration) ≅ 1-hour.

Table 3: Catchment area of different streams

Sl. No.	Name of catchment	Area (km^2)
1	Kansar Khala	32.0
2	Chandini Khala	61.0
3	Mandi Khala	35.0
4	Maswa Khala	62.0
5	Satuan Khala	65.0
6	Khorowala Khala	45.0

In the absence of the short duration rainfall, say, 1-hour, 2-hour, 3-hour, etc., the same was computed from the estimated 1-day maximum rainfall by using conversion factors, as given in Central Water Commission (CWC, 1984) report. The 1-hour maximum value of distributed rainfall is computed by multiplying the 1-day maximum rainfall with factor of 0.340 and presented in Table 4. The distributed 1-hour rainfall was used as input for computation of PFD as the catchment areas of different tributaries of YRB are in the range of 20 to 65 km^2.

These streams are ungauged and hence the PFD for ungauged catchments is computed by using rational formula, which is given below:

$$q = 0.278 * C I A \quad \ldots (11)$$

where, q is peak discharge (m^3/s), C is runoff coefficient, I is rainfall intensity (mm/hour) and A is catchment area (km^2). By considering topography of the river basin, the value of C is considered as 0.60 while computing the PFD. The computed PFD for six ungauged catchments of YRB are presented in Table 5, which could be taken as design flood at the catchments. The PFD estimates obtained from EV1 distribution for different catchments of YRB are used to develop the rainfall frequency curves, as presented in Figure 4.

Table 4: Distributed rainfall for 1-hour duration

Return period (year)	Estimated 1-day maximum rainfall (mm)	1-hour rainfall (mm)
2	164.6	56.0
5	226.6	77.0
10	267.7	91.0
15	290.9	98.9
20	307.1	104.4
25	319.6	108.7
50	358.2	121.8
75	380.5	129.4
100	396.4	134.8

Table 5: Peak Flood Discharge (m^3/s) for six catchments of YRB

Name of the catchment	PFD (m^3/s) for different return periods of								
	2-yr	5-yr	10-yr	15-yr	20-yr	25-yr	50-yr	75-yr	100-yr
Kansar Khala	298.6	411.3	485.8	527.9	557.4	580.1	650.0	690.6	719.4
Chandini Khala	569.3	784.0	926.1	1006.3	1062.5	1105.8	1239.0	1316.4	1371.3
Mandi Khala	326.6	449.8	531.4	577.4	609.6	634.4	710.9	755.3	786.8
Maswa Khala	578.6	796.8	941.3	1022.8	1079.9	1123.9	1259.3	1338.0	1393.7
Satuan Khala	606.6	835.4	986.9	1072.3	1132.2	1178.3	1320.2	1402.8	1461.2
Khorowala Khala	420.0	578.3	683.2	742.4	783.8	815.7	914.0	971.2	1011.6

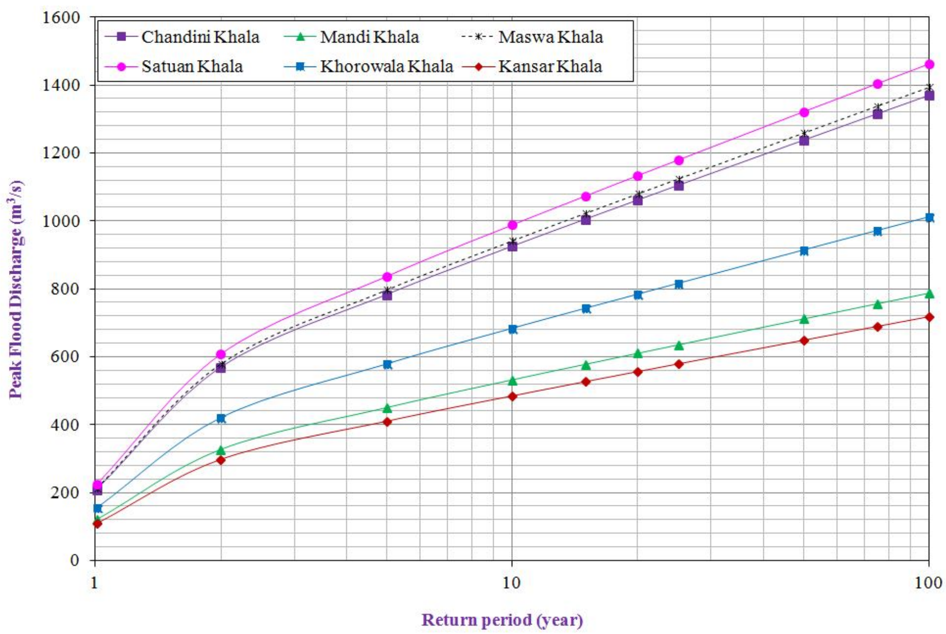

Figure 4: Plots of estimated PFD using EV1 distribution for ungauged catchments

CONCLUSIONS

The paper presents the study carried out for EVA of rainfall for Dhaulakuan station adopting EV1 and LP3 distributions (using MLM). GoF tests such as A^2 and χ^2 are applied for checking the adequacy of fitting of probability distributions to the recorded rainfall data. The selection of most suitable probability distribution is made through RMSE and used for estimation of PFD for six ungauged catchments of YRB. The following conclusions are drawn from the study:

i) From the rainfall frequency curves, it is found that the recorded AMR data are falling between the confidence limits of the estimated 1-day maximum rainfall from EV1 and LP3 distributions.
ii) GoF test results support the EV1 and LP3 distributions for modelling the series of AMR.
iii) Based on EVA results and RMSE values, the EV1 distribution is identified as better suited distribution for estimation of rainfall.
iv) The estimated 1-day maximum rainfall is used to compute 1-hour maximum value of distributed rainfall adopting CWC guidelines, as described in Flood estimation report of Indo Ganga plains.
v) By using the 1-hour distributed rainfall, the PFD for six ungauged catchments of YRB is computed from rational method.
vi) The study suggests the estimated PFD at six catchments could be used for design of flood in river protection works of YRB.

ACKNOWLEDGEMENTS

The author is grateful to the Director, Central Water and Power Research Station (CWPRS), Pune, for providing the research facilities to carry out the study. The author is thankful to Dr. C. Ramesh, and Dr. R.G. Patil, CWPRS, for their valuable suggestions and continuous encouragement during the course of study.

REFERENCES

Casas, M.C., Rodriguez, R., Prohom, M., Gazquez, A. and Redano, A., Estimation of the probable maximum precipitation in Barcelona (Spain), *Journal of Climatology*, 2011, Vol. 31, No. 9, pp. 1322–1327.

Central Water Commission (CWC), Flood estimation report for Upper Indo Ganga Plains - Sub Zone 1(e), New Delhi, 1984.

Charles Annis, P.E., Goodness-of-Fit tests for statistical distributions, [http://www.statistical engineering.com/goodness.html], 2009.

Das, L.M. and Qureshi, Z.H., Flood frequency analysis for Jiya Dhol river of Brahmaputra valley, *Journal of Sciences: Basic and Applied Research*, 2014, Vol. 14, No. 2, pp. 14-24.

Daneshfaraz, R., Nemati, S., Asadi, H. and Menazadeh, M., Comparison of four distributions for frequency analysis of wind speed: A case study, *Journal of Civil Engineering and Urbanism*, 2013, Vol. 3, No. 1, pp. 6-11.

Gumbel, E.J., Statistic of Extremes, 2nd edition, Columbia University Press, New York, 1960.

Izinyon, O.C. and Ajumuka, H.N., Probability distribution models for flood prediction in Upper Benue River Basin, *Journal of Civil and Environmental Research*, 2013, Vo1. 3, No. 2, pp. 62-74.

Jalal, S., Özgur, K., Oleg, M., Abbas-Ali, S. and Bagher, N., Fore-casting daily stream flows using artificial intelligence approaches, *ISH Journal of Hydraulic Engineering*, 2012, Vol. 18, No. 3, pp. 204-214.

Lee, B.H., Ahn, D.J., Kim, H.G. and Ha, Y.C., An estimation of the extreme wind speed using the Korea wind map, *Renewable Energy*, 2012, Vol. 42, No. 1, pp. 4–10.

Rasel, M. and Hossain, S.M., Development of rainfall intensity duration frequency equations and curves for seven divisions in Bangladesh, *International Journal of Scientific & Engineering Research*, 2015, Vol. 6, No. 5, pp. 96-101.

Rao, A.R. and Hameed, K.H., Flood Frequency Analysis, CRC Publications, New York, 2000.

National Institute of Hydrology (NIH), Technical note on hydrological process in an ungauged catchment, 2011, pp. 1-163.

Olumide, B.A., Saidu, M. and Oluwasesan, A., Evaluation of best fit probability distribution models for the prediction of rainfall and runoff volume (Case Study: Tagwai Dam, Minna-Nigeria), *Journal of Engineering and Technology*, 2013, Vo1. 3, No. 2, pp. 94-98.

Singh, R.D., Mishra, S.K. and Chowdhary, H., Regional flow duration models for 1200 ungauged Himalayan watersheds for planning micro-hydro projects, *ASCE Journal of. Hydrologic Engineering*, 2001, Vol. 6, No. 4, pp. 310-316.

Zhang, J, 'Powerful Goodness-of-Fit Tests Based on the Likelihood Ratio', *Journal of Royal Statistical Society: Series B*, 2002, Vol. 64, Part 2, pp. 281-294.