Assessment of Extreme Rainfall Using Gumbel Distribution for Estimation of Peak Flood Discharge for Ungauged Catchments

N. Vivekanandan

Scientist-B, Central Water and Power Research Station, Pune 411024, Maharashtra, India

***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 catchments, rainfall depth becomes an important input in derivation of PFD. So, rainfall depth can be estimated through statistical analysis by fitting probability distribution to the rainfall data. In this paper, the series of annual maximum 1-day rainfall derived from the daily rainfall data observed at Dhaulakuan rain gauge station is used for estimation of 1-day extreme rainfall adopting Gumbel distribution. Maximum likelihood method is used for determination of parameters of the distribution. Anderson-Darling test is applied for checking the adequacy of fitting of the distribution to the observed rainfall data. The estimated 1-day extreme rainfall obtained from Gumbel distribution is used to compute the 1-hour maximum value of distributed rainfall that is considered as an input to estimate the PFD by rational formula adopting CWC guidelines. The study suggests the estimated PFD could be used for design of flood protection works for different ungauged catchments of river Yamuna.**

***Index Terms*—Anderson-Darling test, Gumbel, Rainfall, Peak flood discharge, Maximum likelihood method**

I. INTRODUCTION

Estimation of Peak Flood Discharge (PFD) at a desired location on a river is important for planning, design and management of hydraulic structures such as dams, bridges, barrages and design of storm water drainage systems. These include different types of flood such as standard project flood, probable maximum flood and design basis flood. In case of large river basins, the hydrological and stream flow series of a significant duration are generally available. However, for ungauged catchments, more data is not available other than rainfall. The rainfall data is also of shorter duration and may becomes an important input in derivation of PFD [1]. For arriving at such design values, statistical analysis by fitting probability distribution to the rainfall data is carried out.

Out of a number of probability distributions, the family of Extreme Value Distributions (EVDs) includes Generalized Extreme Value, Extreme Value Type-1 (Gumbel), Extreme Value Type-2 (Frechet) and Extreme Value Type-3 (Weibull) is widely adopted for Extreme Value Analysis (EVA) of rainfall [2]. EVDs arise as limiting distributions for the sample of independent, identically distributed random variables, as the sample size increases. In the group of EVDs, Gumbel distribution has no shape parameter as when compared to other distributions and this means that there is no change in the shape of Probability Distribution Function (PDF). Moreover, the Gumbel distribution has the advantage of having only positive values, since the data series of rainfall are always positive (greater than zero); and therefore Gumbel distribution is important in statistics. Lee et al. [3] applied Gumbel and Weibull probability distributions for estimation of extreme wind speed for Korea region and found that the Gumbel distribution gives better results than Weibull.

Daneshfaraz et al. [4] carried out frequency analysis of wind speed adopting 2-parameter log-normal, truncated extreme value, truncated logistic and Weibull probability distributions and found that the truncated extreme value is the most appropriate distribution for Iran. Esteves [5] applied Gumbel distribution to estimate the extreme rainfall depths at different rain-gauge stations in southeast United Kingdom. Olumide et al. [6] applied normal and Gumbel distributions for prediction of rainfall and runoff at Tagwai dam site in Minna, Nigeria. They have also expressed that the normal distribution as better suited for rainfall prediction while Log-Gumbel for runoff. Rasel and Hossain [7] applied Gumbel distribution for development of intensity duration frequency curves for seven divisions in Bangladesh. In view of the above, Gumbel distribution is used in the present study. Parameters of the Gumbel distribution are determined by Maximum Likelihood Method (MLM) and used to estimate the 1-day extreme rainfall. For quantitative assessment on rainfall data within the observed range, Anderson-Darling test (A^2) is applied. The estimated 1-day extreme rainfall from Gumbel distribution is used to estimate the PFD for ungauged catchments of river Yamuna. The methodology adopted in EVA of Annual Maximum 1-day Rainfall (AMR) using Gumbel distribution, computation of GoF test statistic and estimation of PFD using rational formula are briefly described in the ensuing sections.

II. METHODOLOGY

The study is to estimate PFD for ungauged catchments of river Yamuna. Thus, it is required to process and validate the data for various application such as (i) assess the adequacy of fitting of Gumbel distribution to the series of AMR using GoF test; (ii) estimate the 1-day extreme rainfall adopting Gumbel distribution (using MLM); (iii) compute the 1-hour maximum rainfall from the estimated 1-day extreme rainfall using CWC guidelines; (iv) compute the PFD using rational formula; and (v) analyze the results obtained thereof.

A) PDF and CDF of Gumbel Distribution

The PDF and Cumulative Distribution Function (CDF) of the Gumbel distribution are given as below:

$$\left.\begin{aligned} &\text{PDF}: f(R) = \frac{e^{-(R-\alpha)/\beta} e^{-e^{-(R-\alpha)/\beta}}}{\beta} \\ &\text{CDF}: F(R) = e^{-e^{-(R-\alpha)/\beta}}, \beta > 0, \text{ where } (R = R_1, R_2, R_3, \ldots, R_N) \end{aligned}\right\} \quad (1)$$

where, α and β are the location and scale parameters of the distribution [8]. The parameters are computed by MLM through Equations (2) and (3), and used to estimate the rainfall (R_T) for different return periods from $R_T = \alpha + Y_T\beta$. Here, $Y_T = -\ln(-\ln(1-(1/T)))$ is called as a reduced variate for a given return period T (year).

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

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

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

where R_i is the observed AMR of i^{th} sample and $\overline{R}$ is the average value of AMR. The lower and upper confidence limits (LCL and UCL) of the estimated extreme rainfall are obtained from the linear expressions viz., LCL=ER-1.96(SE) and UCL=ER+1.96(SE). Here, ER is the estimated Extreme Rainfall and SE the Standard Error.

B) Goodness-of-Fit Test

GoF test is essential for checking the adequacy of probability distribution to the observed series of AMR. Out of a number GoF tests available, the widely accepted GoF test is A^2, which is used in the study. The theoretical description of A^2 test statistic [9] is as follows:

$$A^2 = (-N) - (1/N) \sum_{i=1}^{N} \left\{ \begin{aligned} &(2i-1)\ln(Z_i) + \\ &(2N+1-2i)\ln(1-Z_i) \end{aligned} \right\} \quad (5)$$

Here, $Z_i = F(R_i)$ for i=1,2,3,…,N with $R_1 < R_2 < \ldots < R_N$, $F(R_i)$ is the CDF of i^{th} sample (R_i) and N is the sample size.

Test criteria: If the computed value of GoF test statistic given by Gumbel distribution is less than that of the theoretical value at the desired significance level then the distribution is considered to be acceptable for modelling the series of AMR.

III. APPLICATION

In this paper, efforts were made to estimate the PFD for different return periods for ungauged catchments of river Yamuna was carried out. The series of AMR was extracted from the daily rainfall data observed at Dhaulakuan rain gauge station during the period 1988 to 2015 and used for estimation of 1-day extreme rainfall. Figure 1 presents the time series plot of the observed AMR. The descriptive statistics such as average, standard deviation, coefficient of variation, coefficient of skewness and coefficient of kurtosis of the observed AMR was determined as 171.7 mm, 58.8 mm, 33.3%, 0.428 and -0.523 respectively. The estimated 1-day extreme rainfall obtained from Gumbel distribution (using MLM) was used as an input to estimate the PFD.

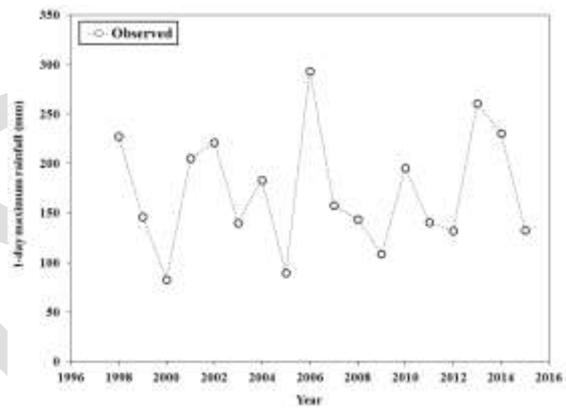

Fig. 1: Time series plot of the observed AMR

IV. RESULTS AND DISCUSSIONS

A) Statistical Analysis of Rainfall using Gumbel Distribution

By applying the procedures of Gumbel distribution, parameters were determined by MLM and used for estimation of 1-day extreme rainfall for different return periods. The parameters of Gumbel distribution were used to develop the CDF plot for the series of AMR. The plots observed and estimated CDF for the AMR series of Dhaulakuan rain gauge station are presented in Figure 2.

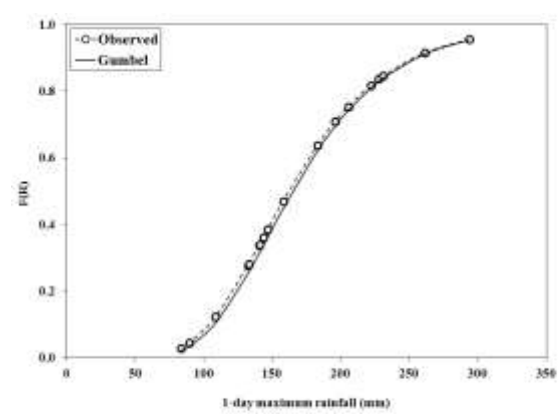

Fig. 2: Plots of observed and estimated CDF for the series of AMR

Table 1 gives the 1-day extreme rainfall estimates with confidence limits for different return periods adopting Gumbel distribution. The observed and estimated AMR are presented in Figure 3 along with confidence limits. From Figure 3, it can be seen that the observed AMR data are within the confidence limits of the estimated 1-day extreme rainfall using Gumbel distribution.

TABLE 1
ESTIMATED 1-DAY EXTREME RAINFALL WITH LOWER AND UPPER CONFIDENCE LIMITS

Return period (year)	1-day Extreme Rainfall (mm)	Standard Error (mm)	Confidence limits at 95% level	
			Lower	Upper
2	162.0	13.4	135.8	188.2
5	216.8	20.5	176.6	257.0
10	253.1	26.3	201.5	304.7
15	273.6	29.8	215.2	331.9
20	287.9	32.2	224.7	351.0
25	298.9	34.1	232.0	365.8
50	332.9	40.1	254.4	411.4
75	352.7	43.6	267.3	438.1
100	366.7	46.0	276.4	456.9

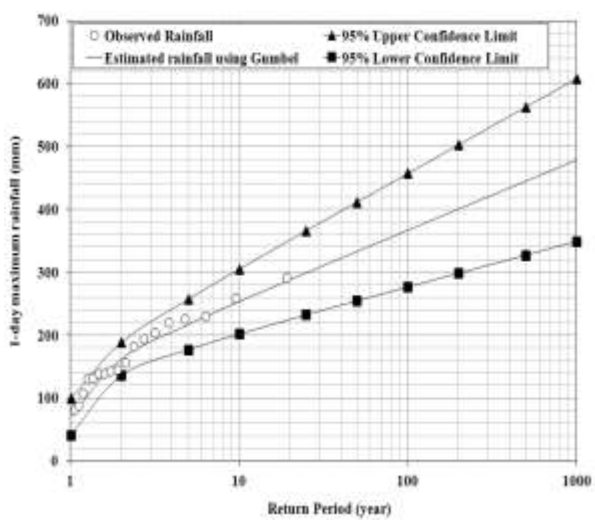

Fig. 3: Plots of observed and estimated 1-day extreme rainfall with confidence limits

B) Analysis Based on GoF Test

The adequacy of fitting of Gumbel distribution to the series of AMR was performed by adopting A^2 test, as described above. From the result, it was observed that the computed value of A^2 test statistic is 0.281, which is not greater than the theoretical value of 0.757 at 5% significance level, and at this level, the A^2 test confirmed the suitability of Gumbel distribution for modelling the series of AMR.

C) Estimation of Peak Flood Discharge

Estimation of peak flood discharge for four catchments of river Yamuna at Ponta Sahib was required for the model studies on flood protection works for river Yamuna and its streams. The size of these catchments is presented in Table 2 and the locations of the catchments are shown in Figure 4.

TABLE 2
CATCHMENT AREA OF DIFFERENT STREAMS

Sl. No.	Name of Catchment	Area (km^2)
1	Touns	365.87
2	Gojar Khala	5.43
3	Baghani Khala	8.29
4	Jambu Khala	32.87

These streams are ungauged and hence the peak flood discharges for these catchments were computed by using rational formula, which is given below:

$$q = 0.278 * C\ I\ A \quad (6)$$

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 studying topography of the catchments using maps and other available literature, the value of the C was considered as 0.55 for computing the flood discharge. The time of concentration (t_c) estimated was 1-hour. In the absence of the short duration rainfall, say, 1-hour, 2-hour, 3-hour, etc., it was computed from estimated extreme 1-day rainfall (100-year return period) by using distribution factor (Figure 5), as given in CWC [10] report entitled 'Flood estimation report for Western Himalayas-Zone 7' in which the study area falls.

Fig. 4: Locations of the sub-catchments of river Yamuna

The computed 1-hour maximum values of distributed rainfall from the 1-day maximum rainfall are presented in Table 3. In the present study, the distributed 1-hour rainfall was used as rainfall intensity in rational method for computation of peak flood discharges of different catchments. The estimated PFD for different return periods for ungauged catchments of river Yamuna presented in Table 4 could be taken as design flood for flood protection works.

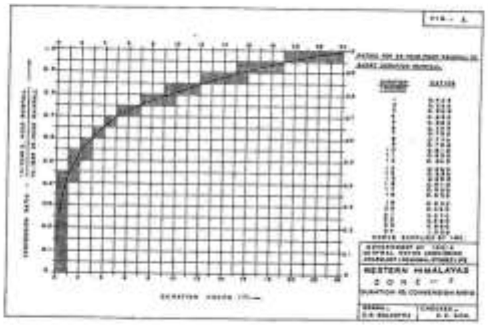

Fig. 5: Conversion factor for computation of distributed rainfall for shorter duration

TABLE 3
DISTRIBUTED RAINFALL FOR SHORT DURATIONS

Return period (year)	Estimated 1-day extreme rainfall (mm)	1-hour maximum rainfall (mm)
2	162.0	68.9
5	216.8	92.1
10	253.1	107.6
15	273.6	116.3
20	287.9	122.4
25	298.9	127.0
50	332.9	141.5
75	352.7	149.9
100	366.7	155.8

TABLE 4
ESTIMATED PFD (m^3/s) FOR DIFFERENT CATCHMENTS OF RIVER YAMUNA

Return period (year)	Peak flood discharge (m^3/s) for			
	Touns	Gojar Khala	Baghani Khala	Jambu Khala
2	3851.6	57.2	87.3	346.0
5	5154.5	76.5	116.8	463.1
10	6017.5	89.3	136.3	540.6
15	6504.9	96.5	147.4	584.4
20	6844.9	101.6	155.1	614.9
25	7106.4	105.5	161.0	638.4
50	7914.7	117.5	179.3	711.1
75	8385.5	124.5	190.0	753.4
100	8718.3	129.4	197.5	783.3

V. CONCLUSIONS

The paper described briefly the study carried out for statistical analysis of rainfall data adopting Gumbel distribution and estimation of PFD for ungauged catchments of river Yamuna. The following conclusions were drawn from the study:

i) The A^2 test result confirmed the suitability of Gumbel distribution (using MLM) for modelling the data series of AMR.
ii) From Figure 3, it was observed that the observed AMR data are within the confidence limits of the estimated 1-day extreme rainfall.
iii) The estimated 1-day extreme rainfall was used to compute 1-hour maximum value of distributed rainfall adopting CWC guidelines described in Flood estimation report for Western Himalayas-Zone 7.
iv) By using the 1-hour distributed rainfall, the PFD for ungauged catchments of river Yamuna was computed by rational formula.
v) The study suggested that the PFD, as given in Table 4, could be considered for design of flood protection works. However, considering the data length made available for the study, it was cautioned to use the PFD for return periods beyond 50-year because of uncertainty in the estimated values.

ACKNOWLEDGMENTS

The author is grateful to Dr. M.K. Sinha, 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. R.G. Patil, Scientist-D, CWPRS, for supply of rainfall data used in the study.

REFERENCES

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

[2] M.C. Casas, R. Rodriguez, M. Prohom, A. Gazquez, and A., Redano, “Estimation of the probable maximum precipitation in Barcelona (Spain)”, *Journal of Climatology*, vol. 31, no. 9, pp. 1322–1327, 2011.

[3] B.H. Lee, D.J. Ahn, H.G. Kim, and Y.C. Ha, “An estimation of the extreme wind speed using the Korea wind map”, *Renewable Energy*, vol. 42, no. 1, pp. 4–10, 2012.

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

[5] L.S. Esteves, “Consequences to flood management of using different probability distributions to estimate extreme rainfall”, *Journal of Environmental Management*, vol. 115, no. 1, pp. 98-105, 2013.

[6] B.A. Olumide, M. Saidu, and A. Oluwasesan, “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*, vo1. 3, no. 2, pp. 94-98, 2013.

[7] M.M. Rasel, and S.M. Hossain, “Development of rainfall intensity duration frequency equations and curves for seven divisions in Bangladesh”, *International Journal of Scientific and Engineering Research*, vol. 6, no. 5, pp. 96-101, 2015.

[8] E.J. Gumbel, "Statistic of Extremes", 2nd edition, Columbia Univ. Press, New York, 1960.
[9] P.E. Charles Annis, "Goodness-of-Fit tests for statistical distributions", [http://www.statistical engineering.com/ goodness. html], 2009.
[10] CWC, "Flood estimation report for Western Himalayas-Zone 7", Central Water Commission (CWC) Design Office Report No.: WH/22/1994, New Delhi, 1994.

N. Vivekanandan is presently working as Scientist in Central Water and Power Research Station, Pune. He did post graduation in Mathematics from Madurai Kamaraj University in 1991. He obtained post graduate degree in Hydrology from University of Roorkee in 2000 and Master of Philosophy degree in Mathematics from Bharathiar University in 2006. He also obtained Master of Business Management (Human Resources) degree from Manonmaniam Sundaranar University in 2013. His research interest includes hydrometeorological studies using probabilistic approach, prediction of hydrometeorological variables using soft computing techniques, optimization of hydrometric network using spatial regression approach, irrigation planning, water resources planning and management, assessment of low-flows and drought, etc.