

# Flood frequency analysis of Ganga river at Haridwar and Garhmukteshwar

Vikas Kamal<sup>1</sup> · Saumitra Mukherjee<sup>1</sup> · P. Singh<sup>1</sup> · R. Sen<sup>1</sup> · C. A. Vishwakarma<sup>1</sup> · P. Sajadi<sup>1</sup> · H. Asthana<sup>1</sup> · V. Rena<sup>1</sup>

Received: 12 June 2015 / Accepted: 4 January 2016 / Published online: 1 February 2016  
© The Author(s) 2016. This article is published with open access at Springerlink.com

**Abstract** The Ganga River is a major river of North India and is known for its fertile alluvium deposits formed due to floods throughout the Indo-Gangetic plains. Flood frequency analysis has been carried out through various approaches for the Ganga River by many scientists. With changes in river bed brought out by anthropogenic changes the intensity of flood has also changed in the last decade, which calls for further study. The present study is in a part of the Upper Indo-Ganga plains subzone 1(e). Statistical distributions applied on the discharge data at two stations found that for Haridwar lognormal and for Garhmukteshwar Gumbel EV1 is applicable. The importance of this study lies in its ability to predict the discharge for a return period after a suitable distribution is found for an area.

**Keywords** Discharge · Flood frequency · Generalized extreme value · Goodness of fit tests · Gumbel distribution · Lognormal 3P · Log Pearson type III

## Introduction

Agriculture, hydroelectricity and industrial sector derive their water resource indirectly from Summer Monsoon rainfall in the month of June–September. The irony is that the same Monsoon often cause flood in many parts of the country like Assam, Bihar and Uttar Pradesh. The worst drought years were 1877, 1899, 1911, 1918, 1920, 1951, 1965 and 1972 and the worst flood years were 1892, 1933,

1961 and 1983 when many subdivisions reported extremely low and excess rainfall, respectively (Parthasarathy et al. 1987). 1988, 1994, 1998, 2000, 2005, 2008, 2009, 2010, 2013 and 2014 can also be regarded as flood years in India. Assam, the state which lies in Brahmaputra River floodplains has been experiencing flood annually since 1998. Floods in Ganga River have been very common and their cause is attributed to heavy downpour in upper reaches in Uttarakhand district and in the floodplains. Ganga Flood Control Commission was established in 1972 to look into the causes and sort out the flooding problem by suggesting structural measures. The National Flood Control Program was launched in the country in 1954. Since then a good progress has been made in the flood protection measures. About one-third of the flood prone area had been afforded reasonable protection. Besides, many steps were undertaken in planning, implementation and performance of flood warning, protection and control measures (CWC 2007). On an average 32.92 million people are affected by floods every year in India (Report 2011).

Attempts of flood frequency analysis have been made for deltaic region (Jha and Bairagya 2011) and Middle subzone 1(f) of Ganga basin (Kumar et al. 2003). They have adopted the normal, lognormal, gumbel maximum value and Log Pearson type III probability distribution functions to find the flood frequency for different return periods. Now-a-days the L-moment approach is widely used for developing regional flood frequency relationships (Hosking and Wallis 1997). There is a need of data on flood magnitudes and their frequencies for designing of hydraulic structures like dams, spillways, culverts, urban drainage systems; also for road and railway bridges, flood plain zonation, etc. (Kumar et al. 2003). Singo et al. (2013) used similar approach to find that the Log Pearson type III best fits the model to find the flood intensity of different

✉ Saumitra Mukherjee  
saumitramukherjee3@gmail.com

<sup>1</sup> School of Environmental Sciences, Jawaharlal Nehru University, New Delhi, India

return period in flood prone Luvuvhu River Catchment (LRC) of South Africa.

There are many frequency models which are now used for determining hydrologic frequency of flood. Probabilistic model rely on the use of existing data to forecast future scenario and deterministic model rely on the different physical parameters to bring out the result and verify it with the existing data to develop a best fit model. Probabilistic approach is commonly practiced in hydrology (Helsel and Hirsch 2010). Within probabilistic models, the two most popular are Gumbel maximum value and Log Pearson type III distribution.

The development of model for hydrological data is driven by the pattern that one obtains through fitting of various equations into an orderly arrangement of data. Overtime the hydrological models have become more complex with the advent of new theories in mathematical sciences. But in terms of result they are more reliable than before. The most common distribution that have been explored here are: Lognormal 3P, Generalised Extreme Value, Log Pearson type III and Gumbel distribution. These distributions are finally tested to find which one gives the best results and can be utilized for modelling flood hazards in that area.

Log Pearson type III distribution has found very wide use in hydrological sciences, especially in flood frequency analysis. Bernard Bobee discussed its limitation and utilization in his paper in 1975. This method retains the original data and it gives better fit over other distribution for long return period. Similar studies have shown that GEV distribution is a more acceptable distribution over Log Pearson type III (Vogel et al. 1993). Nazemi et al. (2011) corroborated this fact by his studies in Saskatoon city of Saskatchewan in Canada. *Environment Canada* (EC) prefers to use Gumbel distribution with the method of moments (MOM) for precipitation analysis.

The lognormal, GEV, EV1 and LP3 distributions are explained here along with their advantages and disadvantages. A random variable  $x$  (variate) is said to be in lognormal distribution if the logarithmic values of  $x$  is distributed normally, as derived using central limit theorem. The mean and the standard deviation are the two parameters here and third frequency factor is derived from the exceedance probability value. GEV (Generalized Extreme Value Distribution) is a continuous probability distribution method that uses three parameters: location, scale and shape. The shift of a distribution in a particular direction is explained by location parameter, spreading out of the distribution is explained by the scale parameter similar to kurtosis and tails of each distribution is governed by the shape parameter like skewness. For shape parameter ( $k$ ) = 0, Gumbel or EV1 distribution is applicable, for  $k > 0$ , EV2 or Frechet is applicable, and for  $k < 0$  EV3 or Weibull is applicable. In general, GEV which has more

parameters will be able to model the input data more accurately than a distribution with a lesser number of parameters. GEV is also good for sample size greater than 50 (Cunnane 1989). Cunnane also found that 3–4 parameter distributions have less bias. Gumbel Distribution (EV1) uses 2 parameters, location ( $\xi$ ) and scale ( $\alpha$ ) and is used for all Precipitation Frequency Analysis in Canada. The LP3 distribution is also referred to as the Gamma distribution. The LP3 distribution is complex due to 2 interacting shape parameters (Stedinger and Griffis 2007).

The parameter estimation is done by using many ways, viz. by maximum likelihood estimators, method of moments (MOM) or by methods of L-Moments. L-Moments are based on probability-weighted moments (PWMs), for the data arranged in ascending order. The MOM technique is good for limited range of parameters, whereas L-Moments can be more widely used, and are unbiased (Rowinski et al. 2001).

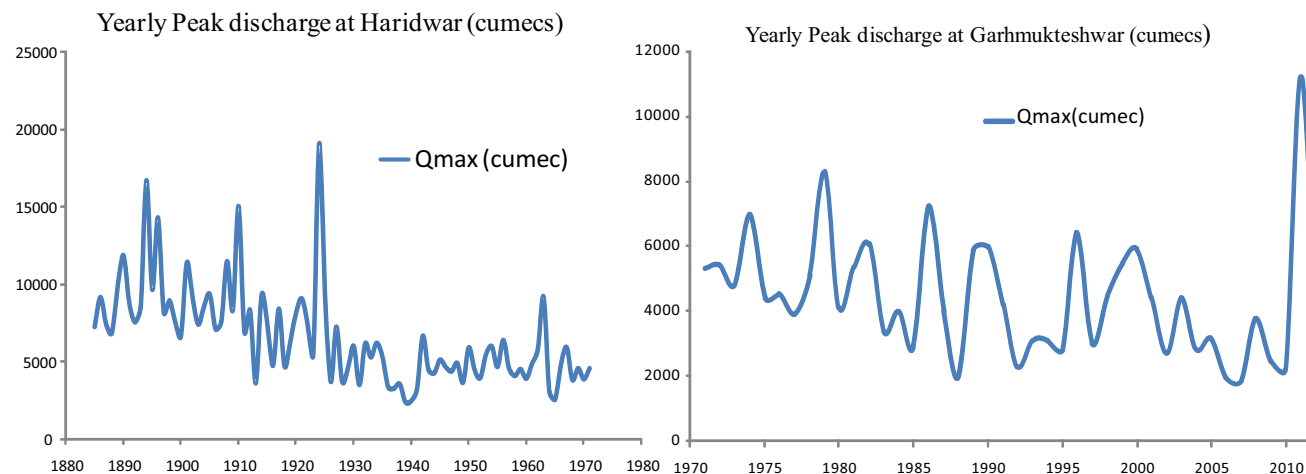
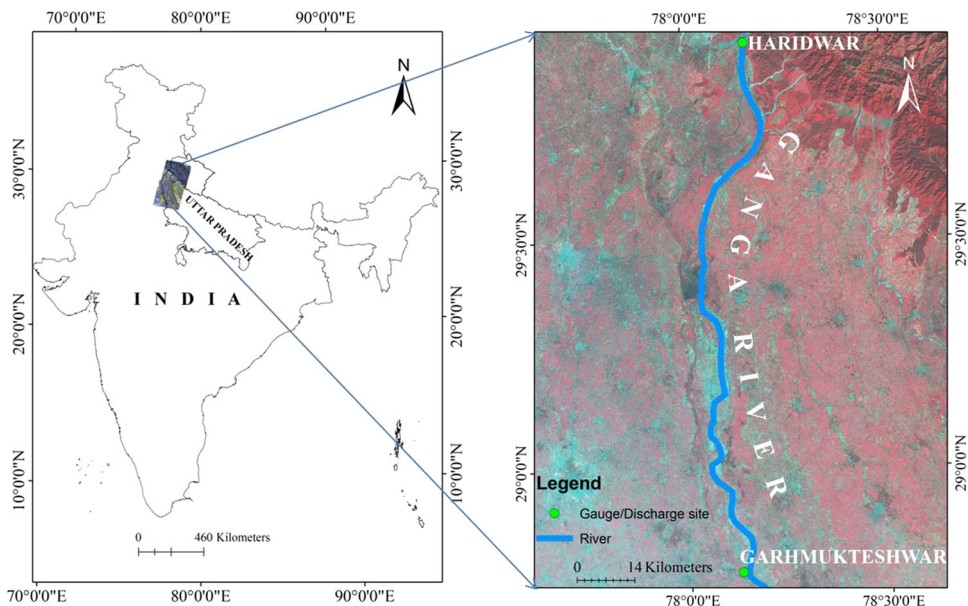
## Study area and data availability

The yearly discharge data from two locations on the Ganga River have been used here, one at Haridwar and the other located 145 km downstream at Garhmukteshwar. Haridwar site is located at 78.165°E longitude and 29.942°N latitude and Garhmukteshwar site is located at 78.148°E longitude and 28.758°N latitude (Fig. 1). The maximum yearly discharge data of Haridwar is taken from the book authored by Professor H.M. Raghunath, Hydrology Principles, Analysis and Design. This data is from 1885 to 1971. The yearly discharge data of Garhmukteshwar (1970–2010) has been obtained through proper channel from CWC (Central Water Commission), and since the data is restricted by the Indian Government due to international character of Ganga River, it has not been shown here; only the graph is shown (Fig. 2). The data of Haridwar is also shown along with the Summer Monsoon rainfall data of Eastern U.P. region which is the region where Haridwar falls, to show how well the rainfall peaks match with that of discharge (Fig. 3). The lognormal values help in synchronizing the data of rainfall and discharge which are in different units. The rainfall data is used from the work of Parthasarathy et al. 1987.

## Methodology

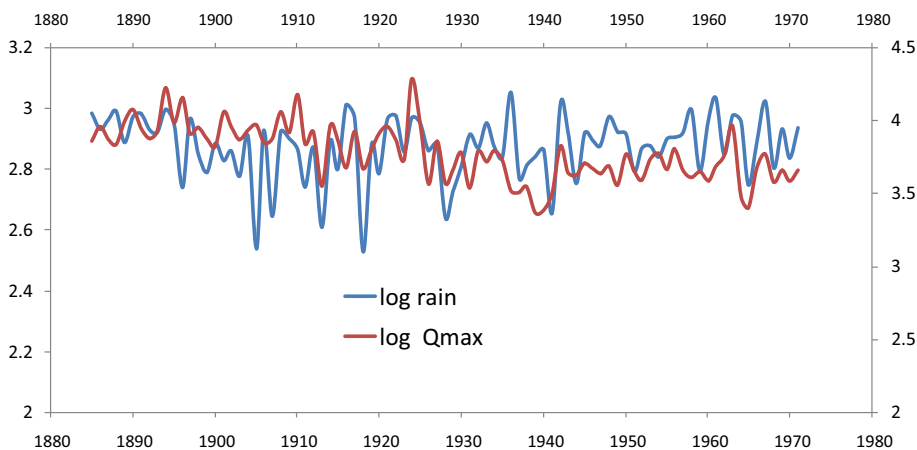
Generalized Extreme Value distribution is done on the L-Moments approach and MOM is used in LP3 and EV1. PWMs are needed to find L-Moments. The data is first arranged in ascending order, and then following equations are used to calculate PWMs: M100, M110, M120 and M130 (Cunnane 1989).

**Fig. 1** Study area with two gauging sites on Ganga River



**Fig. 2** Graphs showing yearly maximum discharge variation in both the stations

**Fig. 3** Graph showing log values of Summer Monsoon rainfall and yearly discharge



$$M_{100} = \frac{1}{N} \sum_{i=1}^N Q_i \quad (1)$$

$$M_{110} = \frac{1}{N} \sum_{i=1}^N \frac{i-1}{n-1} Q_i \quad (2)$$

$$M_{120} = \frac{1}{N} \sum_{i=1}^N \frac{(i-1)(i-2)}{(n-1)(n-2)} Q_i \quad (3)$$

$$M_{130} = \frac{1}{N} \sum_{i=1}^N \frac{(i-1)(i-2)(i-3)}{(n-1)(n-2)(n-3)} Q_i \quad (4)$$

in which  $N$  is the sample size,  $Q$  is the data value, and  $i$  is the rank of the value in ascending order. The L-Moments are then calculated as follows (Cunnane 1989):

$$\lambda_1 = L1 = M_{100} \quad (5)$$

$$\lambda_2 = L2 = 2M_{110} - M_{100} \quad (6)$$

$$\lambda_3 = L3 = 6M_{120} - 6M_{110} + M_{100} \quad (7)$$

$$\lambda_4 = L4 = 20M_{130} - 30M_{120} + 12M_{110} - M_{100} \quad (8)$$

The L-moments are further used to derive variation coefficient L-CV ( $\tau_2$ ), symmetry coefficient L-Skewness ( $\tau_3$ ) and peakedness coefficient L-Kurtosis ( $\tau_4$ ) as follows, (Hosking and Wallis 1997):

$$\tau_2 = L2/L1 \text{ (L - CV)} \quad (9)$$

$$\tau_3 = L3/L2 \text{ (L - Skewness)} \quad (10)$$

$$\tau_4 = L4/L2 \text{ (L - Kurtosis)} \quad (11)$$

Generalized Extreme Value (GEV) distribution uses three parameters:  $\xi$ , the location parameter,  $\alpha$ , the scale parameter and  $\kappa$ , the shape parameter. The parameters are defined from (Hosking and Wallis 1997) as:

$$\kappa = 7.8590c + 2.9554c^2 \quad (12)$$

$$\text{in which } c = \frac{2}{3 + \tau_3} - \frac{\ln 2}{\ln 3} \quad (13)$$

$$\xi = \lambda_1 - \frac{\{1 - \Gamma(1 + \kappa)\}}{\kappa} \quad (14)$$

in which  $\Gamma$  = the gamma function.

Finally the return period discharge is calculated using the following formulae:

$$Qt = \xi + (\alpha/k) \{1 - \log((T-1)/T)\}k \quad (15)$$

in which  $T$  is the desired return period in years.

Step by step GEV performed in excel (Millington et al. 2011) is as follows :

- Firstly, sort the data set by ordering all of the data points in ascending order (lowest to highest)
- Calculate the 4 PWM's (M100, M110, M120, M130)

- Calculate the 4 L-Moments ( $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ ) using the PWMs
- Calculate  $k$ , the shape parameter
- Calculate  $\xi$ , the location parameter and  $\alpha$ , the scale parameter
- Using the desired return period, apply all parameters to the Return Period equation to calculate the discharge value.

The US Water Resources Council (1967) adopted the Log-Pearson Type-III distribution. The procedure is to first convert the data to logarithms and calculate the following (Raghunath 2006):

$$\text{Mean } \overline{\log x} = \frac{\sum \log x}{n} \quad (16)$$

$$\text{Standard deviation : } \sigma_{\log x} = \sqrt{\frac{\sum (\log x - \overline{\log x})^2}{n-1}} \quad (17)$$

$$\text{Skew coefficient : } g = \frac{n \sum (\log x - \overline{\log x})^3}{(n-1)(n-2)(\sigma_{\log x})^3} \quad (18)$$

The values of  $x$  for various recurrence intervals are computed from,

$$\log x = \overline{\log x} + K\sigma_{\log x} \quad (19)$$

The frequency factor  $K$  is obtained from the following Table 1 for the computed value of 'g' and the desired recurrence interval.

Gumbel's method by V.T. Chow is used. The equation is

$$Qt = a + bX_t \quad (20)$$

$$\text{where, } X_t = \log \left( \log \left( \frac{T}{T-1} \right) \right) \quad (21)$$

$a, b$  = parameters estimated by the method of moments. The following equations are derived from the method of least squares.

$$\sum Q = an + b \sum X_t \quad (22)$$

$$\sum (QX_t) = a \sum X_t + b \sum X_t^2 \quad (23)$$

Now, 'a' and 'b' can be solved.

In this method, a plotting position has been assigned for each value of  $Q$  when arranged in the descending order. For example, if an annual flood peak  $Q_T$  has a rank  $m$ , its plotting position

$$T = \frac{n+1}{m} \quad (24)$$

From Eq. (21),

$$X_t = \log \left( \log \left( \frac{n+1}{n+1-m} \right) \right) \quad (25)$$

**Table 1** Table of frequency factor ‘K’ for LogPearson III distribution

Skewness coefficient (g)	Recurrence interval try in years							
	1	2	5	10	25	50	100	200
Annual probability of occurrence in % = 1 – F								
Cs	99	50	20	10	4	2	1	0.5
0	-2.326	0	0.842	1.282	1.751	2.054	2.326	2.576
-0.1	-2.4	0.017	0.846	1.27	1.716	2	2.252	2.482
-0.2	-2.472	0.033	0.85	1.258	1.68	1.945	2.178	2.388
-0.3	-2.544	0.05	0.853	1.245	1.643	1.89	2.104	2.294
-0.4	-2.615	0.066	0.855	1.231	1.606	1.834	2.029	2.201
-0.5	-2.686	0.083	0.856	1.216	1.567	1.777	1.955	2.108
-0.6	-2.755	0.099	0.857	1.2	1.528	1.72	1.88	2.016
-0.7	-2.824	0.116	0.857	1.183	1.488	1.663	1.806	1.926
-0.8	-2.891	0.132	0.856	1.166	1.448	1.606	1.733	1.837
-0.9	-2.957	0.148	0.854	1.147	1.407	1.549	1.66	1.749
-1	-3.022	0.164	0.852	1.128	1.366	1.492	1.588	1.664

We substitute the values and solve the equations for getting ‘a’ and ‘b’, finally to get *Qt*.

The three parameter lognormal (TPLN) distribution is used as the fourth method of distribution. Properties of this distribution are discussed by Aitchison and Brown (1957), and Johnson and Kotz (1970). For a random variable X, if  $Y = \ln(X - a)$  has a normal distribution then X will have a lognormal distribution whose probability density function (pdf) can be expressed as

$$f(x) = \frac{1}{(x - a)c\sqrt{2\pi}} \exp\left[-\frac{[\ln(x - a) - b]^2}{2c^2}\right] \quad (26)$$

where ‘a’ is a positive quantity defined as a lower boundary, and ‘b’ and ‘c<sup>2</sup>’ are the form and scale parameters of the distribution. ‘b’ is equal to the mean and ‘c<sup>2</sup>’ is equal to the variance of log values. The cumulative distribution function (cdf) of the TPLN is an integral function from x to a of f(x) (Singh 1998). The cdf obtained from EasyFit software is used to calculate the Annual Exceedance Probability (AEP), or the probability that the event is exceeded or equaled in any single year. This is calculated as (1 – P). Return period is calculated as inverse of AEP. Then finally the *Qt* for a return period ‘r’ is obtained using the logarithmic relation between *return period* and discharge values.

**Goodness of fit tests**

Climatic datasets are analyzed using different distribution techniques and to find which one is most reliable, we use the goodness of fit tests. These tests are:

1. The Anderson–Darling (AD) and
2. Kolmogorov–Smirnov (KS)

Solaiman 2011 described all test statistics. The goodness of fit tests was carried out using EasyFit, available at <http://www.mathwave.com/easyfit-distribution-fitting.html>.

**Anderson–Darling Test**

The Anderson–Darling test compares an observed CDF to an expected CDF. The Anderson–Darling test gives more weight to the tail of the distribution than KS test. The test hypothesis is rejected if the AD statistic is greater than a critical value of 2.5018 at a given significance level  $\alpha = 0.05$ . The AD test statistic ( $A^2$ ) is:

$$A^2 = -n - \frac{1}{n} \sum_{i=1}^n (2i - 1) \cdot [\ln F(xi) + \ln(1 - F(xn - i + 1))] \quad (27)$$

**Kolmogorov–Smirnov Test**

The Kolmogorov–Smirnov test statistic is based on the greatest vertical distance from the empirical and theoretical CDFs. Similar to the AD test statistic, a hypothesis is rejected if the KS statistic is greater than the critical value 0.1255 at a chosen significance level  $\alpha = 0.05$ .

The samples are assumed to be from a CDF  $F(x)$ . The test statistic (D) is:

$$D = \max\left(F(xi) - \frac{i - 1}{n}, \frac{i}{n} - F(xi)\right) \quad (28)$$



Log Normal, Log Pearson type III, Gumbel EV1 (Ven T Chow method) and Generalised Extreme Value (L Moments method) as discussed above were used to calculate maximum discharge for return period of 2, 5, 10, 25, 50, 100 and 200 years in Ganga river at the discharge site of Haridwar and Garhmukteshwar.

### Results and discussion

The following table (Table 2) shows the outcome of the various distributions. The entire process was executed in Microsoft Excel 2007. The graph in Figs. 4 and 5 shows the comparison of discharge calculated by different distributions. It comes out that for both Haridwar and Garhmukteshwar discharge sites; GEV gives maximum values, followed by Gumbel, Log Pearson III and Log-normal 3P at last. To find statistically which distribution best fits the discharge data and gives the best output in terms of return period, the available data was processed in Easyfit software. Easyfit software compares the three Goodness of Fit (GOF) tests. According to the theory discussed before, the statistic is calculated from Kolmogorov–Smirnov test, Anderson–Darling test and Chi-Squared test (Tables 3, 4). The Chi-Squared test determines if a sample comes from a given distribution. It is not considered a high power statistical test and is not so useful (Cunnane 1989). So, the Chi-square has not been adopted here for GOF test.

The critical value at  $\alpha = 0.05$ , i.e. 95 % confidence level for all three test is shown in the table (Table 3, 4). This value decides which distribution is to be rejected from the study. We see that all the distributions are accepted with no rejection statistically. The other fact that is brought out is the significance of the distribution. Ranking is given on the difference between statistic value and the critical value. Lognormal (3P) is given ranking 1 in case of Haridwar data and Gumbel is given ranking 1 in case of

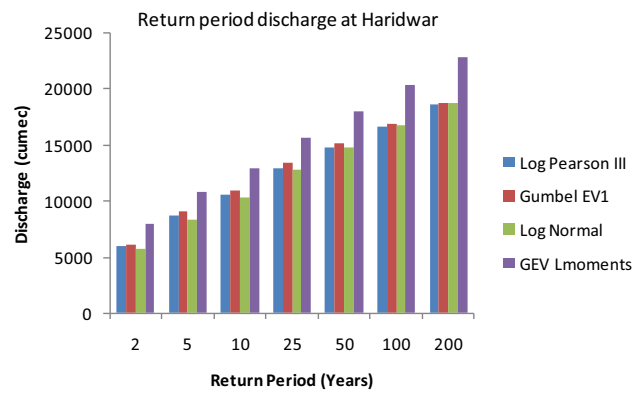


Fig. 4 Bar graph showing the comparison of outcome of four distribution for Haridwar data

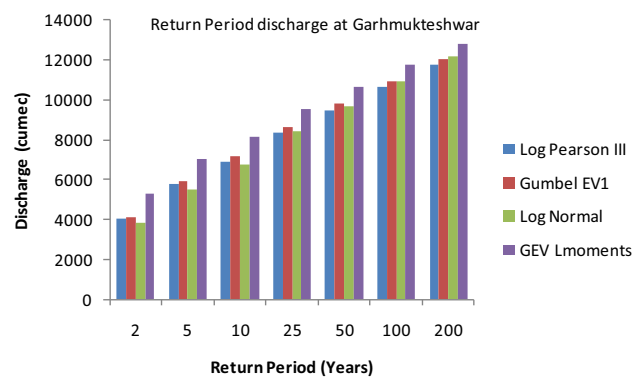


Fig. 5 Bar graph showing the comparison of outcome of four distribution for Garhmukteshwar data

Garhmukteshwar data. The sample size in terms of number of years is high for Haridwar i.e. 87 (1885–1971) and low for Garhmukteshwar, i.e. 42 (1971–2013). The present study is corroborated by the previous similar studies on latest data done by Kumar et al. (2003) where GEV (L moments method) was found to be robust for Middle Ganga subzone (1f) and Singo et al., where Gumbel

Table 2 Final result showing the return period discharge values expected through four distribution methods

Return period (years)	Garhmukteshwar				Haridwar (discharge in cumecs)			
	Log Pearson III	Gumbel EV1 Chow	Lognormal Normal	GEV L-Moments	Log Pearson III	Gumbel EV1 Chow	Lognormal Normal	GEV L-Moments
2	4056	4124	3857	5313	6023	6147	5728	7932
5	5749	5941	5509	7001	8714	9044	8314	10,820
10	6898	7143	6759	8131	10,570	10,962	10,271	12,898
25	8377	8663	8411	9574	12,985	13,385	12,858	15,727
50	9497	9791	9660	10,656	14,831	15,182	14,815	17,984
100	10630	10,910	10910	11,738	16,711	16,967	16,771	20,368
200	11790	12,025	12160	12,826	18,648	18,744	18,728	22,893

**Table 3** Goodness of fit test for Haridwar data

S. No.	Distribution	Kolmogorov–Smirnov (critical value at 0.05 = 0.14355)			Anderson–Darling (critical value at 0.05 = 2.5018)			Chi-Squared (critical value at 0.05 = 12.592)		
		Statistic	Reject	Rank	Statistic	Reject	Rank	Statistic	Reject	Rank
	Haridwar									
1	GEV	0.0728	No	3	0.39622	No	3	6.578	No	3
2	Gumbel EV1	0.07369	No	4	0.54909	No	4	9.0697	No	4
3	Log-Pearson 3	0.06812	No	2	0.37121	No	2	4.4149	No	2
4	Lognormal (3P)	0.06451	No	1	0.3707	No	1	3.4708	No	1

**Table 4** Goodness of fit test for Garhmukteshwar data

S. No.	Distribution	Kolmogorov–Smirnov (critical value at 0.05 = 0.20517)			Anderson–Darling (critical value at 0.05 = 2.5018)			Chi-Squared (critical value at 0.05 = 11.07)		
		Statistic	Reject	Rank	Statistic	Reject	Rank	Statistic	Reject	Rank
	Garh-Mukteshwar									
1	GEV	0.0661	No	2	0.20129	No	2	1.2078	No	2
2	Gumbel EV1	0.0631	No	1	0.19808	No	1	2.141	No	3
3	Log-Pearson 3	0.0734	No	3	0.22259	No	3	0.4161	No	1
4	Lognormal (3P)	0.07255	No	4	0.22832	No	4	0.6585	No	4

distribution and Log Pearson 3 gave good results for steep Luvuhu river catchment. Haridwar is analogous to Luvuhu as it lies in foothills and Garhmukteshwar is very close to Middle Ganga subzone (1f).

So we can conclude that Gumbel is good for low sample size and Lognormal (3P) gives good result for large sample size (Table 4). Log Pearson III is placed at poor ranking in Garhmukteshwar data which supports the fact that Log Pearson III is not good for small sample size, and Gumbel is better than this. Now, the question arises, why discharge is less at Garhmukteshwar, though it is downstream of Haridwar and theoretically the discharge increases downstream. The answer can be easily given from the fact that there is significant withdrawal of river water via canal at Bijnor barrage which lies in between Haridwar and Garhmukteshwar. Bijnor barrage is also known as Madhya Ganga canal project which started in 1976. Also there are no perennial tributaries which come and join in between. So, naturally the discharge level goes down at Garhmukteshwar, which has discharge data after 1970. This underlines the methodological limitations of statistical distributions which primarily rely on the fact that the flow in a river is not altered through unnatural ways and the data availability is continuous and of long duration at every station along the river. Ironically, such conditions are hard to find for any river and field data availability is also scarce for such rivers due to the legal and technical issues involved.

## Conclusion

The present study has been done on the data available for Upper Ganga region, and is important because of dearth of data availability, for the Ganga River. The floodplain of Ganga River is facing danger of encroachment by illegal construction. The future scope of the present work is that the values of return period flood can be used to construct the flood hazard zones and define the river space. This river space is to be preserved for the sake of ecology, riparian vegetation and nutrient recycling during floods. It signifies the horizontal connectivity in a fluvial system.

The statistical approaches have been used widely to fit the data and predict the values for return period by many authors. The study has shown that the recent technique of GEV distribution that uses L-Moments does not fits well with the discharge data of Ganga in Haridwar for long term data but Log normal (3P) fits and prove more reliable for flood frequency analysis. Goodness of fit tests validated that Gumbel EV1 distribution stand high in ranking for short term data of Garhmukteshwar at 145 km downstream. The comparison of return period discharge further proves that Log normal (3P) gives more practical result if we have more historical data, with values neither overshooting nor undershooting.

**Acknowledgments** The authors wish to thank Central Water Commission Upper Ganga Division for providing the data. This work was possible due to fellowship grant provided by Council of Scientific and Industrial Research, India. The authors also thank Sri Dayaram

Yadavji (lab attendant) for helping in data collection. Mr. Ritesh Sipolya and Ms. Neha Singh are also acknowledged for discussion regarding the work.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

- Aitchison J, Brown JAC (1957) The lognormal distribution with special reference to its uses in economics. Cambridge University Press, London **18**
- Cunnane C (1989) Statistical distributions for flood-frequency analysis: World Meteorological Organization, Operational Hydrology Report No. 33 Secretariat of the World Meteorological Organization–No. 718, 61 p. plus appendixes (1989)
- CWC 2007 Annual Report 2007–08 (2007) Central Water Commission, Chapter III, p 23
- Government of India Report on disaster management in India (2011) Ministry of Home Affairs Chapter 1, p 20
- Helsel DR, Hirsch RM (2010) Statistical methods in water resources. U.S. Geological Survey, Investigations Book 4, Chapter A3, pp 97–113
- Hosking JRM, Wallis JR (1997) Regional frequency analysis. Cambridge University Press, Cambridge
- Jha VC, Bairagya H (2011) Environmental impact of flood and their sustainable management in deltaic region of West Bengal, India. *Caminhos de Geografia*. 12(39)
- Johnson WL, Kotz S (1970) Distributions in statistics: continuous univariate distributions 1. Houghton-Mifflin, Boston, Massachusetts **1**
- Kumar R, Chatterjee C, Kumar S, Lohani AK, Singh RD (2003) Development of regional flood frequency relationships using L-moments for middle Ganga Plains Subzone 1(f) of India. *Water Resour Manage* 7:243–257
- Millington N, Das S and Simonovic SP (2011) The Comparison of GEV, Log-Pearson Type 3 and Gumbel Distributions in the Upper Thames River Watershed under Global Climate Models. *Water Resources Research Report No: 077*. pp 10–19
- Nazemi AR, Elshorbagy A, Pingale S (2011) Uncertainties in the estimation of future annual extreme daily rainfall for the City of Saskatoon under Climate Change Affects 20th Canadian Hydrotechnical Conference, CSCE
- Parthasarathy B, Sontakke NA, Monot A, Kothawale DR (1987) Droughts/floods in the summer monsoon season over different meteorological subdivisions of India for the period 1871–1984. *J Climatol* 7:57–70
- Raghunath HM (2006) Hydrology: principles, analysis and design Second revised edition. pp 354
- Rowinski PM, Strupczewski WG, Singh VP (2001) A note on the applicability of log-Gumbel and log-logistic probability distributions in hydrological analyses, *Hydrological Sciences ~ J Sci Hydrol*, 47:1
- Singh VP (1998) Entropy-based parameter estimation in hydrology, 30, Ch 7, pp 82–107
- Singo LR, Kundu PM, Odiyo JO, Mathivha FI, Nkuna TR (2013) Flood frequency analysis of annual maximum stream flows for Luvuvhu river catchment, Limpopo province, South Africa, University of Venda, Department of Hydrology and Water Resources, Thohoyandou, South Africa, pp 1–9
- Solaiman TA (2011) Uncertainty estimation of extreme precipitations under climatic change: a non-parametric approach, PhD Thesis, Department of Civil and Environmental Engineering, University of Western Ontario
- Stedinger JR, Griffis VW (2007) Log Pearson Type 3 distribution and its application in flood frequency analysis. I: Distribution characteristics, *J Hydrol Eng ASCE*
- Vogel RM, Thomas WO, McMahon TA (1993) Flood-flow frequency model selection in southwestern United States. *J Water Res Plan Manag* 119(3):353–366