IJCRT.ORG ISSN : 2320-2882
INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)
An International Open Access, Peer-reviewed, Refereed Journal

Flood Frequency Analysis of Jamsholaghat Sub-basin using Three Methods and its Comparison

Sanjay Ku. Ray [1] and Venkappayya R. Desai [2]

[1]Civil Engineering Department, Indian Institute of Technology Kharagpur, West Bengal, India -721302,
[2]Civil Engineering Department, Indian Institute of Technology Kharagpur, West Bengal, India -721302,

Abstract: Flood frequency analysis play most important role in prediction in calculation of Peak Flood Discharge for a required return period which is essential for management and design of hydrological structures like bridges, weirs, spillways, dams etc. Also to find out the most suitable method that could anticipate extreme events of certain natural phenomena like rainfall, flood etc. This study focuses on flood frequency analysis of Jamsholaghat sub-basin using three methods (viz., Log Pearson Type-III, Partial Duration Series and Gumbel's Method). Peak discharge has been estimated for different return period (viz., 100y, 200y and 500y) by analyzing 13 years of GD data followed by goodness of fit test. The result shows the variation of peak discharge among the different methods. Goodness of fit test (Kolmogorov Smirnov and Chi Square test) were applied in which Partial duration series shows the best-fit result. In future to maintain hazardous level of flood events of the river Subarnarekha, significance steps must be taken in order to save humanity from the ravages of flood events.

Keywords: Flood Frequency, Peak Discharge, Gumbel's Method, Partial Duration Series, Log Pearson Type-III, Chi Square Test, Kolmogorov Smirnov Test.

1. Introduction

Flood recurrence investigation assumes a crucial job in many water asset ventures. During design and planning of water resource projects, Engineers often require to estimate the size and recurrence of floods that will affect the required study area. Various researchers from different fields adopted the applications of frequency analysis method in their research works (Bhagat, 2017; Salinas et al., 2014; Zaman et al., 2012). There are number of flood frequency analysis distributions that have been used to hydrological data in order to estimate the flood magnitude and frequency. Some of the flood frequency analysis methods used in hydrology are Log Pearson Type-III distribution, Log-Normal distribution, Partial Duration Series, Gumbel's Distribution etc (Bezak et al., 2014; Keast and Ellison, 2015). Flood recurrence examination gives the likelihood of event of a specific occasion. So as to appraise the likelihood of event of the floods, detail examination of stream information is required. The procedure engaged with flood recurrence examination is to fit the likelihood model to the example of yearly flood peaks that has been recorded over a period of perceptions for a given watershed. The parameters got from this model at that point can be utilized to anticipate the extreme events of various return periods. Flood recurrence examination assumes an essential job in the executives of floodplains, to limit the harm because of flood and for arranging and planning of water driven structures, for example, dam, bridge, barrage etc (Baratti et al., 2018; Izinyon et al., 2011; Luo et al., 2015). The use of statistic has a broad access of significant applications in research area of hydrology. With the help of these application we can measure complex data which can be helpful in prediction of future flood events, also to examine the dynamics of hydrological events. To ascertain that we are dealing with most significance, accurate and reliable methods, researches must be done on new thoughts about more reasonable emerging distributions (Mandal and Choudhury, 2015; Mukherjee, 2008; Suhartanto et al., 2018). This paper will show us about the accuracy and reliability of distributions.

The main objective of the investigation was to compare the three distributions i.e., Gumbel's Method, Log Pearson Type III and Partial Duration Series. For Goodness of fit test, Kolmogorov Smirnov (KS) and Chi Square tests were used in this paper.

2. Literature Review

Griffis and Stedinger (2007) presented a study on flood frequency analysis using Log Pearson Type-III (LP3) distribution. They considered three major sections for the analysis. First one is to investigate the characteristics of LP3 in both real space and log space. Second one is identify the parameter ranges for which LP3 showed the comparable result to that of U.S. Flood records. Third one is to develop the L-moments for LP3 distribution. The authors concluded that the LP3 distribution showed the reasonable result as compared to U.S. flood data. LP3 density function was able to assume many shapes with two shape parameters. Polynomial expressions for the L-kurtosis were developed, so that they can be used in L-moment frequency analyses.

Mujere (2011) presented a study on analyzing the flood frequency of Nyanyadzi river, Zimbabwe using Gumbel's distribution. The author considered 30 water years (from 1969 to 1999) for the analysis. To check the fitness of the Gumbel distribution, the author used Chi- square test. The author concluded that, Gumbel distribution shows the good result in Chi-square test and it was considered for flood frequency analysis of Nyanyadzi river, Zimbabwe. For 100yr and 200yr return period, the magnitude of floods were estimated as 276 m^3/s and 310 m^3/s respectively.

Karim et al. (2017) presented a study on frequency analysis of small magnitude floods using Annual Maximum and Partial Duration Series. The authors considered five frequency distribution models for their study and they collected data from 24 gauging stations in Great Barrier Reef (GBR) lagoon catchments in north-eastern Australia. The author concluded that the partial series results shows better result than that of annual series in case of small and medium flood. For large floods, both partial series and annual series showed similar results. In case of partial series, if flood threshold value varies largely as compared to bankfull discharge then it can underestimate or overestimate the frequency.

Alam et al. (2018) presented a study to determine the best fit probability distribution among different statistical distribution using monthly rainfall data of 30 years (from 1984 to 2013) for 35 locations in Bangladesh. Commonly used flood frequency distribution methods were used and their parameters were estimated by L-moments estimator and method of moments. The authors concluded that the Generalized Extreme value, Pearson Type-III and Log Pearson Type-III distribution showed the best fit results for large number of stations. They also estimate the peak flood for 10y, 25y, 50y and 100y return periods.

3. Methodology

This chapter deals with the methodological steps involved in estimating the peak flood using three methods (viz., Gumbel's Method, Log Pearson Type-III and Partial Duration Series) for 13 years duration followed by goodness of fit test using Chi Square Test and Kolmogorov Smirnov Test.

3.1 Study Area

Study area involves the Jamsholaghat sub-catchment of Subarnarekha Basin. The present study area extends over states of Odisha and West Bengal having total area of 552 km^2. It lies between 86°30' to 86°50' E longitudes and 22°04' to 22°32' N latitudes. The Subarnarekha and the Burhabalang form the major river systems in the sub-catchment. Length of the Subarnarekha River and Burhabalang River is 22.74 km and 44.32 km respectively. Delineated DEM of study area in Subarnarekha Basin has been shown in Fig. 1.

3.2 Estimation of Peak Discharge & Design Flood by Gumbel's Method

Gumbel's equation for estimating the peak discharge x_T is given by (Mukherjee, 2013):

$$x_T = \bar{x} + K\sigma_{n-1} \quad (3.1)$$

$$\sigma_{n-1} = \text{Standard deviation of the sample of size N} = \sqrt{\frac{\Sigma(x-\bar{x})^2}{N-1}} \quad (3.2)$$

$$K = \text{Frequency Factor} = \frac{y_T - Y_n}{S_n} \quad (3.3)$$

$$y_T = \text{Reduced variate} = -\left[\ln.\ln\frac{T}{T-1}\right] \quad (3.4)$$

Where, $\sigma_{(n-1)}$ is standard deviation, K is frequency factor. In this method, first annual maximum flood for each year need to be calculated followed by finding the $\bar{x}$ and σ_{n-1} of annual maximum flood. Then, a suitable return period for the given data has to be assumed followed by estimating Y_T and K. After substituting all the values in equation 3.1, peak discharge has been estimated.

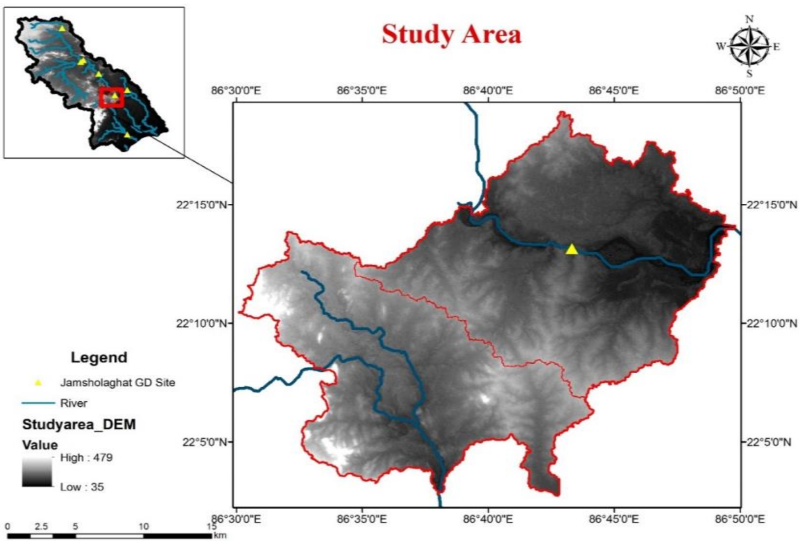

Fig. 1. Delineated DEM of study area in Subarnarekha Basin

3.3 Estimation of Peak Discharge & Design Flood by Partial Duration Series

Partial Duration Series method for estimating Peak discharge (Q_T) is given by (Subramanya, 2012):

$$Q_T = Q_b + X_T \qquad (3.5)$$

$$\text{Exceedance } X_T \text{ for an interval of T years } = \frac{1}{\beta}[\ln(\lambda T) - \ln\{\ln(\lambda T)\}] \qquad (3.6)$$

Where, Q_b is Bankfull Discharge, β and λ are Constants of Partial Duration Series as given below:

$$\beta = \frac{-\ln\{1 - H(x)\}}{x} \qquad (3.7)$$

First step is to find out the bankfull discharge followed by finding the monthly peak discharge (x_i) for each year. Then, estimating flood peak exceedance (i.e., $x_i - Q_b$), λ and β using the given equations. After assuming suitable return period, and substituting the value of λ and β, X_T has been calculated. After substituting all the values in equation 3.5, peak discharge has been estimated.

3.4 Estimation of Peak Discharge & Design Flood by Log Pearson Type III distribution

Log Pearson Type III for estimating Peak discharge (Q_T) is given by (Subramanya, 2012):

$$\log Q_T = \overline{\log Q} + K\sigma_{\log Q} \qquad (3.7)$$

$$\text{Mean } = \overline{\log Q} = \frac{\sum \log Q}{n} \qquad (3.8)$$

$$\text{Standard Deviation } = \sigma_{\log Q} = \sqrt{\frac{\sum(\log Q - \overline{\log Q})^2}{n-1}} \qquad (3.9)$$

$$\text{Skew Coefficient } = \frac{n\sum(\log Q - \overline{\log Q})^3}{(n-1)(n-2)(\sigma_{\log Q})^3} \qquad (3.10)$$

Where, K is frequency factor depends on skew coefficient, T is return Period and n is sample Size (i.e., no. of years of flood record). In this method, first annual maximum flood for each year need to be calculated followed by finding the $\overline{\log Q}$ and $\sigma_{\log Q}$ of annual maximum flood. Then, a suitable return period for the given data has to be assumed followed by calculating skew coefficient and using the skew coefficient, value of K has been taken from the Log Pearson Type III distribution table. After substituting all the values in equation 3.7, peak discharge has been estimated.

3.5 Chi Square Test

To find out goodness of fit between the observed and expected frequencies the Chi Square (χ^2) test was evaluated. The formula of Chi Square (χ^2) is given by (Chow et al., 1988):

$$\chi^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i} \qquad (3.11)$$

Where, O_i is observed frequency, E_i is expected frequency. The significance level α = 5% & 1% produced critical values which used in this paper. If calculated Chi Square value is less than critical value then hypothesis is accepted at a chosen significance level. The Chi Square test is not considered a high power statistical test and is not very useful (Cunnane, 1989).

3.6 Kolmogorov Smirnov (KS) Test

The Kolmogorov-Smirnov test statistic is based on the greatest deviation which is evaluated by comparing every variant between the theoretical and empirical probability. In this test we give rank (From small to big) to the data to find out empirical probability(Chow et al., 1988).

$$D_{max} = |1 - P_e - P_t| \qquad (3.12)$$

$$P_e = \frac{n}{N+1} \qquad (3.13)$$

$$P_t = \frac{1}{T} \qquad (3.14)$$

$$P_t = 1 - e^{-e^{Y_T}} \qquad (3.15)$$

Where, P_t is theoretical probability, P_e is empirical probability, n is rank, N is sample size. If D_{max} is less than critical value then hypothesis is accepted at a significance level.

4 Results and Discussions

Observed Peak discharge data starting from 2004 to 2016 for Jamsholaghat GD site has been shown in Table 4.1.

4.1 Gumbel's Method

Design flood estimation using Gumbl's method for Jamsholaghat GD site has been shown in Table 4.2.

4.2 Partial Duration Series

Partial duration series parameter estimation for Jamsholaghat GD site has been shown in Table 4.3. Design flood estimation using Partial Duration Series for Jamsholaghat GD site has been shown in Table 4.4.

4.3 Log Pearson Type-III Distribution

Design flood estimation using Log Pearson Type-III Distribution for Jamsholaghat GD site has been shown in Table 4.5.

4.4 Chi Square Test

Chi Square Test of all three methods for Jamsholaghat GD site have been shown in Tables 4.6 to 4.8. All the three methods qualify the criteria for Chi Square test having 5% significance level.

4.5 Kolmogorov Smirnov (KS) Test

Kolmogorov Smirnov (KS) Test of all three methods for Jamsholaghat Jamsholaghat GD site have been shown in Tables 4.9 to 4.11. For 5% significance level, ΔD_{cric} is 0.41. From the result it is clear that ΔD_{max} for Partial Duration Series and Log Pearson Type III is less than ΔD_{cric} and ΔD_{max} for Gumbel's method is higer than the ΔD_{cric}.

Table 4.1. Peak discharge of Jamsholaghat Stn.

Year	Peak Discharge (m^3/s)
2004	3369
2005	1706
2006	3242
2007	3500
2008	1859
2009	994
2010	864
2011	3750
2012	2557
2013	2267
2014	3663
2015	2561
2016	2754

Table 4.2. Gumbel's Method of Design Flood Estimation for Jamsholaghat Stn.

$$[\bar{Q} = 2545\ {m^3}/{s}, \sigma_Q = 971\ {m^3}/{s}]$$

Return Period (T) (year)	1.01	2	3.5	5	10	50	100	200	500
Reduced Variate (Yt)	-1.529	0.367	1.089	1.500	2.250	3.902	4.600	5.296	6.214
Frequency Factor (K)	-2.042	-0.141	0.584	0.996	1.748	3.405	4.105	4.803	5.723
Peak Discharge (m^3/s)	562	2408	3112	3512	4242	5850	6530	7207	8101

Table 4.3. Estimation of average value of β for Jamsholaghat Stn.

Flood Peak Exceedance x_i			Cumulative Frequency (CF)	H(x) = (CF/137)	1 – H(x)	β
Class Interval (CI) (m^3/s)	Variable (m^3/s)	Observed Frequency				
0-1000	1000	87	87	0.64	0.36	0.0010
1001-2000	2000	33	120	0.88	0.12	0.0010
2001-3000	3000	12	132	0.96	0.04	0.0011
3001-4000	4000	5	137	1.00	0.00	
		137			Average	0.0011

Table 4.4. Partial Duration Series of Design Flood Estimation for Jamsholaghat Stn.
[Q_b (m^3/s) = 259, λ [y^{-1}] = 10.54, β =0.0011]

Return Period (T) (year)	1.01	2	3.5	5	10	50	100	200	500
X_T (m^3/s)	1203	1547	1860	2070	2495	3545	4016	4495	5137
Peak Discharge (m^3/s)	1462	1806	2119	2329	2754	3804	4275	4754	5396

Table 4.5. Log Pearson Type III Design Flood Estimation for Jamsholaghat Stn.
[$\overline{\log Q} = 3.367\ m^3/s$, $\sigma_{\log Q} = 0.208\ m^3/s$]

Return Period (T) (year)	1.01	2	3.5	5	10	50	100	200	500
Frequency Factor (K)	-3.087	0.180	0.514	0.848	1.107	1.44	1.52	1.55	1.63
Log Q	2.725	3.404	3.474	3.543	3.597	3.665	3.683	3.690	3.706
Peak Discharge (m3/s)	531	2538	2978	3494	3956	4629	4817	4898	5077

Table 4.6. Chi Square Test of Gumbel's Method for Jamsholaghat Stn.

Interval of Class	Observed Frequency (OF)	Expected Frequency (EF)	$(EF-OF)^2/EF$
0 - 562	0	2.6	2.60
563 - 2408	5	2.6	2.22
2409 - 3112	3	2.6	0.06
3113 - 3512	3	2.6	0.06
> 3512	2	2.6	0.14
Total	13		5.08
		α (%)	5%
		Degree of freedom (k-h-1)	2
		Chi-square critic	5.99 (Accepted)

Table 4.7. Chi Square Test of Partial Duration Series for Jamsholaghat Stn.

Interval of Class	Observed Frequency (OF)	Expected Frequency (EF)	$(EF-OF)^2/EF$
0 - 1627	2	2.6	0.14
1628 - 2017	2	2.6	0.14
2018 - 2372	1	2.6	0.98
2373 - 2611	2	2.6	0.14
> 2611	6	2.6	4.45
Total	13		5.85
		α (%)	5%
		Degree of freedom (k-h-1)	2
		Chi-square critic	5.99 (Accepted)

Table 4.8. Chi Square Test of Log Pearson Type III for Jamsholaghat Stn.

Interval of Class	Observed Frequency (OF)	Expected Frequency (EF)	$(EF-OF)^2/EF$
0 - 531	0	2.6	2.60
532 - 2538	5	2.6	2.22
2539 - 2978	3	2.6	0.06
2979 - 3494	2	2.6	0.14
> 3494	3	2.6	0.06
Total	13		5.08
		α (%)	5%
		Degree of freedom (k-h-1)	2
		Chi-square critic	5.99 (Accepted)

Table 4.9. Kolmogorov Smirnov (KS) Test of Gumbel's Method for Jamsholaghat Stn.

Rank	Year	Peak Discharge (m^3/s)	Empirical Probability (P_e)	Frequency Factor (K)	Reduced Variate (Y_t)	Theoretical Probability (P_t)	Deviation (Δ)
1	2005	1706	0.083	-1.582	-1.032	0.940	0.856
2	2008	1859	0.167	-1.369	-0.825	0.898	0.731
3	2013	2267	0.250	-0.799	-0.273	0.731	0.481
4	2012	2557	0.333	-0.394	0.119	0.589	0.255
5	2015	2561	0.417	-0.388	0.124	0.587	0.170
6	2016	2754	0.500	-0.119	0.385	0.494	0.006
7	2006	3242	0.583	0.563	1.044	0.297	0.287
8	2004	3369	0.667	0.740	1.216	0.257	0.410
9	2007	3500	0.750	0.923	1.393	0.220	0.530
10	2014	3663	0.833	1.151	1.613	0.181	0.653
11	2011	3750	0.917	1.272	1.731	0.162	0.754

Table 4.10. Kolmogorov Smirnov (KS) Test of Partial Duration Series for Jamsholaghat Stn.

Rank	Year	Peak Discharge (m^3/s)	Empirical Probability (P_e)	Bank full discharge (m^3/s)	X_t (m^3/s)	Return Period (T) (yr)	Theoretical Probability (P_t)	Deviation (Δ)
1	2005	1706	0.083	395	1311	0.737	(Rejected as T<1)	-
2	2008	1859	0.167	395	1464	1.101	0.092	0.075
3	2013	2267	0.250	395	1872	1.356	0.263	0.013
4	2012	2557	0.333	395	2162	1.76	0.432	0.098
5	2015	2561	0.417	395	2166	2.025	0.506	0.090
6	2016	2754	0.500	395	2359	2.65	0.623	0.123
7	2006	3242	0.583	395	2847	3.74	0.733	0.149
8	2004	3369	0.667	395	2974	5.23	0.809	0.142
9	2007	3500	0.750	395	3105	9.81	0.898	0.148
10	2014	3663	0.833	395	3268	11.18	0.911	0.077
11	2011	3750	0.917	395	3355	13.54	0.926	0.009

Table 4.11. Kolmogorov Smirnov (KS) Test of Log Pearson Type III for Jamsholaghat Stn.

Rank	Year	Peak Discharge (m^3/s)	Empirical Probability (Pe)	Log Q	Frequency Factor (K)	Return Period (T) (yr)	Theoretical Probability (Pt)	Deviation (Δ)
1	2005	1706	0.083	3.232	-1.754	1.183	0.155	0.071
2	2008	1859	0.167	3.269	-1.438	1.311	0.237	0.070
3	2013	2267	0.250	3.355	-0.708	1.606	0.377	0.127
4	2012	2557	0.333	3.408	-0.265	1.786	0.440	0.107
5	2015	2561	0.417	3.408	-0.259	1.789	0.441	0.024
6	2016	2754	0.500	3.440	0.008	1.897	0.473	0.027
7	2006	3242	0.583	3.511	0.609	3.780	0.735	0.152
8	2004	3369	0.667	3.528	0.750	4.255	0.765	0.098
9	2007	3500	0.750	3.544	0.890	4.902	0.796	0.046
10	2014	3663	0.833	3.564	1.058	7.488	0.866	0.033
11	2011	3750	0.917	3.574	1.144	8.965	0.888	0.028

5 Conclusions

In the work done so far, detailed analysis flood frequency using three methods for Jamsholaghat GD site have been carried out followed by goodness of fit test. These are the following conclusions drawn from the study:

- Gumbel's method overestimated the peak discharge whereas Partial Duration Series and Log Pearson Type III methods shows almost similar result.
- All the three methods qualify the criteria for Chi Square test at 5% significance level.
- Partial Duration series and Log Pearson Type III method qualifies the criteria for Kolmogorov Smirnov (KS) test as ΔD_{max} is less than ΔD_{cric} at 5% significance level.
- Gumbel's method did not qualify the criteria for Kolmogorov Smirnov (KS) test as ΔD_{max} is higher than ΔD_{cric} at 5% significance level.
- Out of all three methods, ΔD_{max} for Partial Duration Series and Log Pearson Type III method is 0.149 and 0.152 respectively, it means these two methods are suitable methods for estimating the peak discharge in case of Jamsholaghat GD Site.

References

[1] Alam Md. A., Emura K., Farnham C. and Yuan J. (2018). Best-Fit Probability Distributions and Return Periods for Maximum Monthly Rainfall in Bangladesh, Climate, 6, 9; doi:10.3390.

[2] Baratti E., Montanari A., Castellarin A., Salinas J.L., Viglione A. and Bezzi A. (2012). Estimating the flood frequency distribution at seasonal and annual time scales, Hydrol. Earth Syst. Sci., 16, 4651–4660.

[3] Bezak N., Brilly M. and Sraj M. 2014. Comparison between the peaks-over-threshold method and the annual maximum method for flood frequency analysis. Hydrol. Sci. J., 59, 959–977.

[4] Bhagat N. (2017). Flood Frequency Analysis Using Gumbel's Distribution Method: A Case Study of Lower Mahi Basin, India. Journal of Water Resources and Ocean Science, 6(4), 51-54.

[5] Chow. V. T., David R. and Mays. Larry W. (1988). Applied hydrology, McGraw-Hill, New York, USA.

[6] Griffis V. W. and Stedinger J.R. (2007). Log-Pearson Type 3 Distribution and its application in flood frequency analysis. I: distribution characteristics, J. Hydrol. Eng., 12(5): 482-491.

[7] Izinyon O.C., Ihimekpen N. and IgbinobaG.E. (2011). Flood Frequency Analysis of Ikpoba River Catchment at Benin City Using Log Pearson Type III Distribution, J. of Emerging Trends in Engineering and Applied Sciences, 2 (1), 50-55.

[8] Karim F., Hasan M. and Marvanek S. (2017). Evaluating annual maximum and Partial Duration Series for estimating frequency of small magnitude floods, Water, 9, 481; doi: 10. 3390.

[9] Keast D. and Ellison J. 2015. Magnitude frequency analysis of small floods using the annual and partial series. Water, 5, 1816–1829.

[10] Luo P., He B., Takara K., Xiong Y.E., Nover D., Duan W. and Fukushi K. 2015. Historical assessment of Chinese and Japanese flood management policies and implications for managing future floods. Environ. Sci. Policy, 48, 265–277.

[11] Mandal S. and Choudhury B.U. 2015. Estimation and prediction of maximum daily rainfall at Sagar Island using best fit probability models. Theor. Appl. Clim , 121, 87–97.

[12] Mujere N. (2011). Flood frequency analysis using the Gumbel Distribution, International Journal on Computer Science and Engineering, 3(7), ISSN: 0975-3397.

[13] Mukherjee M. (2013). Flood frequency analysis of River Subernarekha, India, using Gumbel’s extreme value distribution, International Journal of Computational Engineering Research, 3(7), Issn 2250-3005.

[14] Mukherjee. M. K. (2008). Studies on Empirical model between Discharge (Qp) and Return period (T) of Major Himalayan Rivers in North Bengal, National Conference on Integrated Water & Wastewater Management at School of Water Resources Engineering. Jadavpur University. Kolkata, India.

[15] Salinas J.L., Castellarin A., Kohnová S. and Kjeldsen T.R. 2014. Regional parent flood frequency distributions in Europe—Part 2: Climate and scale controls. Hydrol. Earth Syst. Sci., 18, 4391–4401.

[16] Subramanya K. (2012). Engineering Hydrology; Tata McGraw-Hill 3rdEd. , New Delhi, India.

[17] Suhartanto E., Limantara L.M., Noviadriana D., Harta F.I. and Aryani D.K. (2018). Estimation of design flood with four frequency analysis distributions, Asian Journal of Applied Science and Technology, 2(1), 13-27.

[18] Zaman M.A., Rahman A. and Haddad K. 2012. Regional flood frequency analysis in arid regions: A case study for Australia. J. Hydrol., 475, 74–83.