International Journal of Engineering & Technology, 3 (2) (2014) 93-98
©Science Publishing Corporation
www.sciencepubco.com/index.php/IJET
doi: 10.14419/ijet.v3i2.1803
Research Paper

Frequency analysis of annual one day to five consecutive days maximum rainfall for Gandak river basin

Shivika Singla [1*], Raktim Haldar [2], Rakesh Khosa [2], Rumani Singla [1], Rudraksh Rajeev [1]

[1] *PEC university of technology, Chandigarh, Punjab, India*
[2] *IIT Delhi, new Delhi, India*
**Corresponding author E-mail: singlashivika@gmail.com*

Abstract

The present study has been conducted for rainfall intensity and frequency estimation for the Gandak basin, a region prone to high floods with an unrealized and unexplored hydro-potential. The two popular gridded precipitation datasets i.e.: (1) APHRODITE, and (2) IMD, for the years 1969-2005, has been used to calculate the mean basin precipitation through the Thiessen polygon method on the ARC-GIS interface. This computed data was used to find out the 1-day, 2-day to 5-day consecutive maximum precipitation series and hence fitted into various well-known probability distribution functions viz., Normal, Gamma, Exponential, etc. According to the best fit data in these functions, the quantiles were determined corresponding to a return period of 2, 10, 20, 25, 50 and 100 years. The two widely used tests: Chi-square Test and Kolmogorov-Smirnov Test were employed to further check the goodness of fit of the series in the distributions. The results reveal that the best fit for 1-day was achieved with the normal distribution, for 2-day with GEV and with GPAR for the remaining maximum consecutive days rainfall. Such studies have thus proven to be substantially facilitative in planning for the safe and economic design of various engineered structures such as bridges, culverts, levees, canals, irrigation and drainage works and effective reservoir management.

Keywords: *Floods, Frequency, Hydrology, Probability Distribution, Rainfall.*

1. Introduction

The uncertainty in the nature of the rainfall patterns has made realize the need for its study and prediction of its occurrence in order to successfully tackle the extreme flood and drought events. For design of any hydrologic engineering structure like reservoirs, bridges, culverts, etc. across natural drainage systems, such estimates are explicitly required. It is hence important to find future probabilities of occurrence of events for various return periods, considered as the life span of the structure, based on past data records. Both flood and drought occurrence probabilities can be interpreted by using various statistical methods. Frequency analysis is used in hydrology as one such tool. In frequency analysis, the magnitude of extreme events is related to their frequency of occurrence through the use of probability distributions. More the magnitude of an extreme event, lesser is its frequency of occurrence. For such purposes, annual maximum flood peaks are used as input data series. In absence of a long record of discharge gauge data for any watershed, rainfall data series is used to interpret extreme storm event corresponding to certain return period and then find the discharge through various possible rainfall-runoff models.
A major flood event may be a result of either a one day storm or consecutive days of rainfall. Therefore, analysis of successive days of maximum rainfall has become a tool for safe and economic planning for design of hydraulic structures. Such work has been done earlier by [1] [7] [8] [9] [10] [11]. [3] Performed a regional frequency analysis of 1-day, 2-day, 5-day and 10-day annual maxima for the period from 1961 to 2000 for 204 sites across the UK. In the present paper, a case study of the Gandak river basin, a region prone to excessive flooding has been carried out. The river though perennial in nature, its flow is dominantly seasonal and non-uniform. The river experiences high flows of water during the monsoon period and low flows during summer and winter season. During monsoons the river carries huge discharges inundating the flood plains and leaving patches in the adjoining areas under waterlogged conditions for a considerable length of time, this being further supported by the shallow water table in the plains. During lean period

the river width narrows down to a great extent of about 100 m at many places. This has made a detailed study of the rainfall intensity and frequency estimation of its basin extremely important.

2. Study area

The Gandak basin lies between latitude 25.6° N and 29.4° N and longitude 82.8° E and 85.82° E, with parts of the basin lying in Tibet, Nepal and India. The Gandak River is said to originate near the Nepal – Tibet border at an altitude of 7620 m to the north-east of Dhaulagiri [5]. After that it enters the plains of Champaran district of Bihar at Triveni. For the present study the catchment outlet has been taken a little north of its confluence with the Ganga, with the catchment area amounting to 40,713 km^2. The elevation ranges from 18 m above mean sea level to 7745 m as shown in Fig. 1.

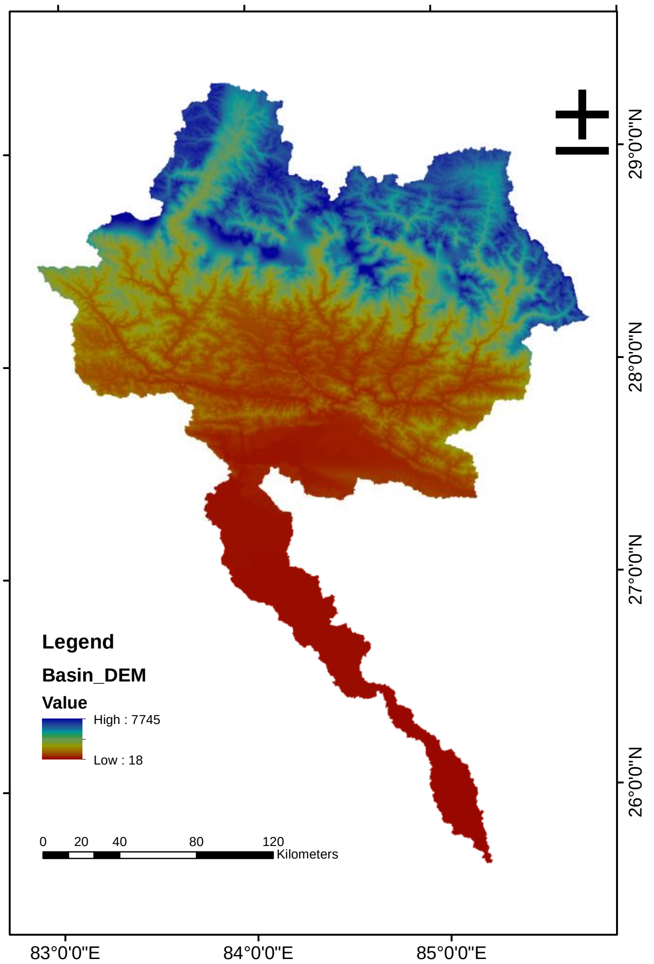

Fig. 1: Elevation map of Gandak basin form SRTM based DEM

3. Methodology

Watershed delineation of the Gandak basin for its study on the Arc GIS interface was done using terrain processing tools. The topographic data required was taken from the SRTM based DEM of 90m resolution [6]. The obtained watershed shape was overlaid on the Gandak's basin boundary available from the Google Earth 3-D interface and checked for any possible errors where it displayed a close match to the ridge lines.

Gridded precipitation data (1969 to 2005) was used to calculate mean basin precipitation for the Gandak basin. For the portion of Gandak basin lying in India, India Meteorological Department (IMD) gridded data was used while for the rest of the basin, for which IMD rainfall data is absent; Asian Precipitation Highly Resolved Observational Data Integration towards the Evaluation of Water Resources (APHRODITE) precipitation data was used. Thiessen polygons were created for the composite gridded locations on the Gandak watershed as shown in Fig. 2 which were used to calculate the mean basin precipitation.

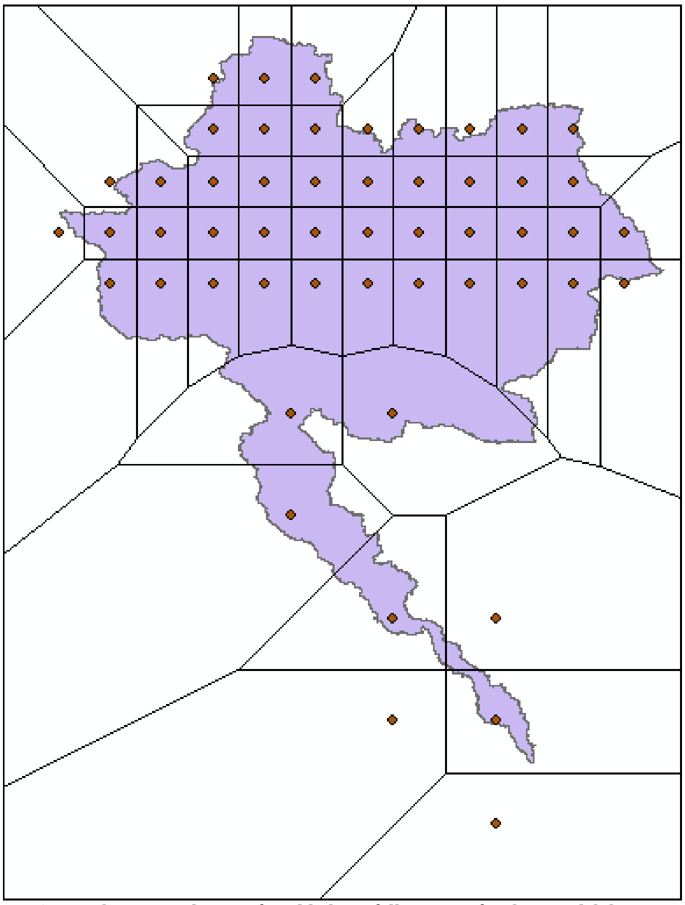

Fig. 2: Thiessen polygon of gridded rainfall stations for the Gandak basin

One day annual maximum series was derived from the mean basin precipitation series for all the 37 years. Likewise consecutive 2-day maximum to 5-day maximum series were prepared for the years 1969-2005. These formulated annual maximum series were then used as an input for the frequency analysis. The 1-day to 5-day annual maximum data series were fitted into various distributions using a frequency stand-alone computer package developed by [4]. The fitted distributions used in the present study are Normal (Normal), General extreme value (GEV), Pearson type 3 (PT-III), Three parameter log normal (LN3), Generalized logistic (GLOG), Extreme value (EV1), Generalized Pareto (GPAR), Weibul (Weibul), Exponential (EXP), Logistic (LOG), Two parameter lognormal (LN2) and Gamma (GAM) [4].

3.1. Testing the goodness of fit of probability distributions:

In this paper, we have tested the goodness of fit using two tests- Chi-square [2] and Kolmogorov-Smirnov (K-S) [10] tests. These tests calculate test-statistics, which are used to analyze how well the data fits the given distribution. These tests describe the differences between the observed and the expected data values from the distribution being tested. They are briefly discussed as below.

3.1.1. Chi-square Test

Chi-square test is the commonly used test for testing the goodness of fit of empirical data. In applying the Chi-square goodness of fit test, the data is grouped into suitable frequency classes. The test compares the actual number of observations and the expected number of observations. The expected number of observations is obtained by considering the above discussed probability distributions one by one falling in the respective class intervals. The expected numbers are calculated by multiplying the expected relative frequency by total number of observations. The chi-square test statistic is computed from the equation:

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

Where, O_i = observed rainfall, and
E_i = expected rainfall

The chi-square distribution has (N - k -1) degree of freedom. The best probability distribution function was determined by comparing Chi-square values obtained from each distribution and selecting the function that gives smallest Chi-square value.

3.1.2. Kolmogorov-Smirnov test

The K –S test statistic is based on the greatest vertical distance from the empirical and theoretical Probability distribution functions. In this test, a hypothesis is rejected if the test statistic is greater than the critical value at a chosen significance level. Similarly, the expected values are fit in the test and function is selected that gives smallest statistic value.

4. Results and discussion

The results are pure simulations. Table 1 shows the various computed statistical parameters for the 1-day to 5-day annual maximum series.

Table 1: Statistical parameters corresponding to mean daily precipitation data

Sl. No.	Parameters	1 day	2 day	3 day	4 day	5 day
1.	Minimum precipitation (mm)	13.00	21.26	26.64	32.50	37.68
2.	Maximum precipitation (mm)	43.48	63.61	80.34	87.70	94.63
3.	Mean	25.07	38.39	48.24	56.05	63.30
4.	Standard Deviation	7.31	11.01	14.34	16.20	17.37
5.	Coefficient of Variation	0.29	0.29	0.30	0.29	0.27
6.	Coefficient of Skewness	0.56	0.47	0.39	0.25	0.08
7.	Kurtosis	2.97	2.78	2.43	2.12	2.06

Table 2 and 3 respectively show the Chi-square and K –S statistic values and p-values as calculated for the considered probability distributions for 1-day, 2-day to 5-day maximum series. A p-value less than 0.05 signify that the test rejects the distribution at 95% confidence level.

Table 2: Chi-square test statistic and p-value

Distributions	1-day		2-day		3-day		4-day		5-day	
	Statistic	p-value	Statistic	p-value	Statistic	p-value	Statistic	p-value	Statistic	p-value
General Extreme Value	4.44	0.727	0.44	1	3.56	0.829	6.22	0.514	8	0.333
Pearson type 3	5.78	0.566	0.89	0.996	4	0.78	7.56	0.373	8	0.333
Three parameter log normal	5.78	0.566	0.89	0.996	3.56	0.829	6.67	0.464	10.22	0.176
Generalized logistic	4.89	0.674	1.33	0.988	7.11	0.417	6.22	0.514	8	0.333
Extreme Value type	5.78	0.566	0.89	0.996	3.56	0.829	7.56	0.373	10.67	0.154
Generalized Pareto	5.33	0.619	5.78	0.566	2.22	0.947	4.89	0.674	5.33	0.619
Weibul	-	-	16.89	0.018	13.33	0.064	20	0.006	11.56	0.116
Normal	4	0.78	2.67	0.914	4.89	0.674	7.56	0.373	8	0.333
Exponential	14.22	0.047	10.22	0.176	10.67	0.154	12	0.101	9.33	0.23
Logistic	6.22	0.514	4	0.78	4.89	0.674	7.11	0.417	8	0.333
Two parameter lognormal	4.44	0.727	0.89	0.996	3.56	0.829	6.67	0.464	10.67	0.154
Gamma	4.44	0.727	1.78	0.971	2.22	0.947	6.22	0.514	10.67	0.154

Analyzing the above values from table 2, we find that according to Chi-square test, the fitted distributions can be ranked as per the least statistic value for the various cases as below:

a) For 1-day maximum series: Normal > GEV = LN2 = GAM > GLOG > GPAR > PT - III = LN3 > EV1 > LOG > EXP > Weibul.

b) For 2-day maximum series: GEV > PT - III = LN3 = EV -1 = LN2 > GLOG > GAM > Normal > LOG > GPAR > EXP > Weibul.

c) For 3 –day maximum series: GPAR = GAM > GEV = LN3 = EV1 = LN2 > PT -III > GLOG = Normal = LOG > EXP > Weibul.

d) For 4 –day maximum series: GPAR > GEV = GLOG = GAM > LN3 > LN2 > LOG > PT –III = EV1 = Normal > EXP > Weibul.

e) For 5 –day maximum series: GPAR > GEV > PT – III = GLOG =Normal =LOG > EXP > LN3 > EV1 = LN2 = GAM > Weibul.

Table 3: K-S test statistic and p-value

Distributions	1-day		2-day		3-day		4-day		5-day	
	Statistic	p-value	Statistic	p-value	Statistic	p-value	Statistic	p-value	Statistic	p-value
General Extreme Value	0.1	0.819	0.08	0.783	0.08	0.872	0.12	0.641	0.15	0.39
Pearson type 3	0.09	0.913	0.07	0.842	0.08	0.755	0.13	0.552	0.15	0.378
Three parameter log normal	0.1	0.868	0.08	0.805	0.08	0.79	0.11	0.447	0.15	0.401
Generalized logistic	0.11	0.793	0.07	0.822	0.08	0.797	0.11	0.481	0.15	0.385
Extreme Value type	0.1	0.869	0.08	0.814	0.08	0.775	0.11	0.432	0.14	0.258
Generalized Pareto	0.09	0.709	0.09	0.905	0.06	0.999	0.1	0.866	0.15	0.384
Weibul	-	-	0.12	0.59	0.14	0.418	0.13	0.526	0.13	0.514
Normal	0.11	0.753	0.09	0.908	0.12	0.634	0.13	0.584	0.14	0.407
Exponential	0.18	0.072	0.16	0.127	0.13	0.341	0.13	0.304	0.17	0.11
Logistic	0.11	0.789	0.09	0.703	0.11	0.75	0.12	0.625	0.15	0.387
Two parameter lognormal	0.1	0.837	0.08	0.804	0.07	0.838	0.11	0.505	0.13	0.286
Gamma	0.11	0.784	0.08	0.821	0.09	0.924	0.12	0.636	0.14	0.371

Similarly, from table 3 we find the best fit probability distributions as per the least statistic value for all the consecutive days maximum rainfall series which showed slightly different results as that obtained from the previous test. For 1-day maximum, PT3 and GPAR were the best fitted distributions; for 2-day PT3 and GLOG; for 3-day, LN2; for 4-day, GPAR and for 5-day, Weibul and LN2 distributions were the corresponding best fits.

The estimates of 1-day to 5-day consecutive maximum rainfall obtained for different return periods, corresponding to the best-fit probability distribution as achieved from the Chi Square test are shown in table 4. It can be observed that for 3-day case, that Gamma and GPAR distributions have the same rank as per the Chi-square statistic value. However, the quantile estimates for various return periods are different. In such a case the distribution giving higher estimate of quantile (here, rainfall) can be chosen. Equation below gives the relationship between return period (T) and probability level (P) expressed in percentage.

$$P = 100/T \quad (2)$$

Table 4: 1-day to 5-day consecutive maximum rainfall estimated for various return periods (T)

T (YEARS)	P (%)	NORMAL 1-day	GEV 2-day	GPAR 3-day	GAM 3-day	GPAR 4-day	GPAR 5-day
2	50	25.07	37.09	46.53	46.87	55.8	65.51
10	10	34.45	52.69	69.47	66.99	79.14	88.53
20	5	37.1	57.94	74.04	73.58	83.04	91.54
25	4	37.88	59.52	75.13	75.57	83.92	92.15
50	2	40.09	64.2	77.68	81.46	85.84	93.4
100	1	42.09	68.53	79.31	87.01	86.96	94.03

As can be seen from the results a maximum rainfall of 25.07 mm in 1-day, 37.09 mm in 2-day, 46.53 mm in 3-day, 55.8 mm in 4-day and 65.51 mm in 5 consecutive days for a return period of two years is likely to occur. Similarly, a precipitation maximum of 42.09, 65.83, 79.31, 86.96 and 94.03 mm in 1, 2, and 3, 4 and 5 consecutive days respectively is expected to occur for a return period of 100 years. Such estimates of maximum precipitation for different return periods are substantially facilitative in planning the safe and economic design of various engineered structures such as bridges, culverts, levees, canals and other irrigation and drainage works. Generally, design life of a hydraulic structure decides the period to be considered for design calculations of a structure. For example, the design life of a bridge can vary from 25 years to 100 years depending on the importance and size of the structure. Similar assumptions have to be made in the case of other hydraulic structures.

References

[1] Bhakar, S., Bansal, A.K., Chhajed, N. & Purohit, R. "Frequency analysis of consecutive days maximum rainfall at Banswara, Rajasthan, India". *ARPN Journal of Engineering and Applied Sciences* 1, 64-67, 2006.

[2] Chakravarti, Indra Mohan, Radha Govira Laha, and Jogabrata Roy. *Handbook of methods of applied statistics. 1. Techniques of computation, descriptive methods, and statistical inference*. Wiley, 1967.

[3] Fowler, H. & Kilsby, C. "A regional frequency analysis of United Kingdom extreme rainfall from 1961 to 2000". *International Journal of Climatology* 23, 1313-1334, 2003.

[4] Hamed, K. & Rao, A.R. "Flood frequency analysis", *CRC press,* 2010.

[5] Jain, S.K. "Impact of retreat of Gangotri glacier on the flow of Ganga River". *Curr. Sci 95,* 1012-1014, 2008.

[6] Jarvis, A., Reuter, H., Nelson, A., & Guevara, E. Hole-filled SRTM for the globe Version 4. *Available from the CGIAR-CSI SRTM 90m Database (http: //srtm. csi. cgiar. org)*, 2008.

[7] Kwaku, X.S. & Duke, O. "Characterization and frequency analysis of one day annual maximum and two five consecutive days' maximum rainfall of Accra, Ghana". *ARPN J. Eng. Appl. Sci* 2, 27-31, 2007.

[8] Rizvi, R., Singh, R., Yadav, R., Tewari, R., Dadhwal, K., & Solanki, K. "Probability analysis of annual maximum daily rainfall for bundelkhand region of Uttar Pradesh". *Indian J. Soil Conser* 29, 259-262, 2001.

[9] Sharda, V. & Bhushan, L. "Probability analysis of annual maximum daily rainfall for Agra". *Indian Journal of Soil Conservation* 13, 16-20, 1985.

[10] Snedecor, George W., and W. G. Cochran. "Statistical methods, 8thEdn."*Ames: Iowa State Univ. Press Iowa* (1989).

[11] Upadhaya, A. & Singh, S. "Estimation of consecutive days Maximum Rainfall by Various Methods and their Comparison". *Indian J. Soil Conser* 26, 193-2001, 1998.