

# Flood frequency analysis for an ungauged Himalayan river basin using different methods: a case study of Modi Khola, Parbat, Nepal

Bibek Acharya

Tribhuvan University, Department of Civil Engineering, Pulchowk Campus, 44600 Lalitpur, Nepal, e-mail:  
acharya.bibek76@gmail.com

Bisesh Joshi

University of Memphis, Department of Civil Engineering, 104 Engineering Science Building, Memphis, TN 38152, USA

## Abstract

Predicting flood discharges in the rivers of an ungauged basin is tedious because essential hydrological data is lacking. In mountainous countries like Nepal, the design of hydraulic structures in these steeply sloped rivers is of prime importance for flood control, as well as for electricity generation where hydraulic head is gained over short, steep reaches. This study illustrates a variety of approaches that can be used to perform flood frequency analysis of typical ungauged mountainous rivers, where discharge data are available from hydrologically similar catchments. The various methods are evaluated by comparing the goodness of fit of an array of hydrologic distribution functions. From each probability density function or regional empirical method, we predict the multi-year return periods for floods, information that is generally required to design the hydraulic structures. The analysis was done based on the annual maxima, peaks above threshold, and widely used regional empirical methods. This analysis was accomplished using the discharge data of Nayapul station near Jhapre Bagar collected from the Department of Hydrology and Meteorology, Government of Nepal, Kathmandu. The analysis and results of this study paved the way for the hydraulic design of water systems in the ungauged study region and demonstrated how the information acquired can be used for water resource management in catchments with similar hydrologic features.

## Keywords

Return periods, ungauged basin, design flood, hydrologic data, probability distribution functions, annual maximum discharge, peaks above threshold.

Submitted 26 June 2020, revised 24 November 2020, accepted 2 December 2020

DOI: 10.26491/mhwm/131092

## 1. Introduction

The design and construction of water systems, as well as water resource management, requires in-depth knowledge of different flood events for different return periods (Tao et al. 2002). The faulty design of engineering structures will have a serious economic impact due to structural damage. Over-designing or under-designing of a hydraulic structure may result in the waste of natural resources or may compromise the structural safety (Reich 1961, 1963). Developing such designs becomes more challenging because of the impact of greenhouse gases, which are changing the hydrological cycle, precipitation patterns, and temperature regimes. Increasing temperatures are altering the physical characteristics of catchments by melting snow and glaciers

(Singh et al. 2018). Researchers are thus challenged to devote more effort to analyzing discharges in the water sources for planning and management.

Adequate discharge data are required for the study, analysis, and quantification of various parameters, including design flood. Hydrological stations are not established in all rivers due to economic and geographical limitations, and hence hydrological analysis in such areas is complicated. The availability of discharge data in Nepal is limited. Rivers that descend from hilly areas of Nepal carry large amounts of sediments, so for the ungauged rivers, the design of hydraulic structures such as weirs, canals, sluice gates, and dams become more complex (Sapkota et al. 2016). Therefore, the main objective of this study is to estimate flood discharges at specific place in ungauged river basins for various return periods, compare estimates and determine the best fit.

The discharge used for the design of a hydraulic structure is called the “design flood”. Designing hydraulic structures for the maximum possible flood for a catchment is very costly. Engineering structures, whose failure may lead to huge loss of lives and properties, are generally designed for floods of large return periods (Izinyon et al. 2011). Design flood estimation is essential for the design of hydraulic structures, flood management and insurance studies, development, and planning. (Rahman et al. 2013).

Hydrologic events have random probability distributions for which statistical analysis can be performed, but precise predictions might not be achieved. Flood frequency analysis is used to estimate design floods for sites along a river that uses observed flow discharge data to calculate statistical information, which is utilized to construct frequency distributions. There is no specific rule for the length of data required for the frequency analysis. Parameter estimation techniques in flood frequency analysis include the graphical method, frequency factor method, method of moments, and method of probability-weighted moments and L-moments (Ojha et al. 2008). Flood frequency estimation is a challenging task for a researcher and has been associated with confusion and controversies (Bobee et al. 1993). Flood frequency analysis helps to predict future flows of different magnitudes and provides reliable predictions in regions of similar climatic conditions. A wide range of research has been conducted to predict the suitable probability distribution functions for annual maximum flood events. Some of the commonly used probability distribution functions include general extreme value, log-normal, normal, Gumbel, Weibull 3P and log Pearson methods. For analysis of the short-term annual maximum discharge, there is no strict rule for using a particular distribution function (Alam et al. 2016).

The bed slope is steep in the mountainous rivers where water flows rapidly, and it is necessary to predict floods for various return periods to design hydraulic structures. Hydropower generation is most common in these rivers. Therefore, designing hydraulic structures like levees, guide walls, dams, intakes, weirs, and barrages need estimates of 10, 20, 50, 100, 200 etc. years return period floods to reduce the risk. Return period flood predictions differ based on the hydrologic distributions selected.

## 2. Study area

Modi Khola is a major tributary of the Kali Gandaki River, which originates from the Annapurna Conservation area of Nepal. The study basin (Fig. 1) has an area of 510 square kilometers. The location selected for study is at 28.273 N and 83.744 E, in the Parbat district. Climate varies from warm temperate to alpine (Rijal 2007), and most of the precipitation occurs during monsoon season (June, July, August, and September). Quartzite, phyllitic slate, schist, and gneiss were found during a site visit. The sediment yield in the river is high because of the steep gradient, erosion of riverbanks, and fragile geological conditions in the upper part of the river basin. Modi Khola carries sediments ranging from sand to huge boulders during the monsoon, eroded from the banks and transported into it by its tributaries. Many hydropower projects are in operation or under construction; there are also new projects proposed for this river.

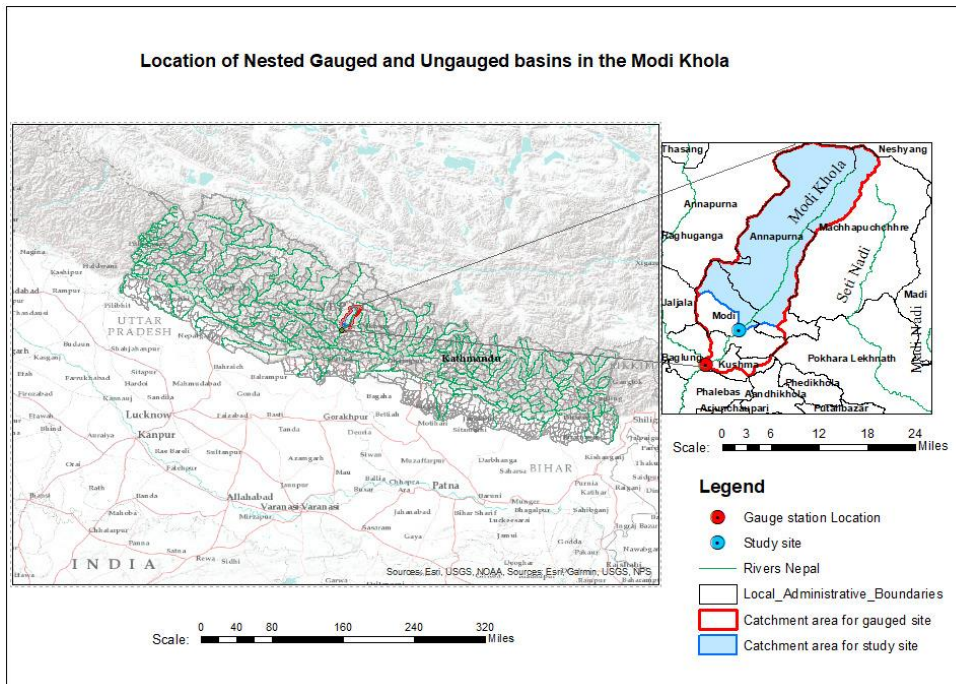


Fig. 1. Location of study basin.

## 3. Methodology

### 3.1. Data collection

Daily discharge data used in this study are collected from the Department of Hydrology and Meteorology, Government of Nepal, from 1976 to 2010, except data from 1980 to 1987 were not available. The flow data available were point discharge data measured once a day at Nayapul near Jhapre Bagar.

### 3.2. Analysis

The oldest and most common technique to estimate the daily flow of an ungauged catchment with the use of a reference catchment is the drainage area ratio method (Archfield, Vogel 2010; Gianfagna et al. 2015). Our study catchment is part of a gauged catchment therefore we used the drainage area ratio method for the nested watersheds.

$$Q_1/A_1 = Q_2/A_2 \quad (1)$$

where  $Q_1$  and  $A_1$  are discharge and area for the gauged catchment, and  $Q_2$  and  $A_2$  are discharge and area for the ungauged catchment.

Annual maximum discharge data were obtained from the dataset for each year by selecting the largest daily flood from that particular year. For hydrological analysis, these data were transferred to the selected outlet point of the study basin by the drainage area ratio method; the data are plotted in Figure 2.



Fig. 2. Transfer of flow data from Nayapul (Gauged Station) to study basin.

For flood frequency analysis using peaks above threshold, identifying such large floods from years of daily data is difficult. We simplified this job by taking the largest flood in a monthly interval from the chunk of daily discharge data. This data series cutoff value was set equal to the smallest discharge from the annual maxima series at Nayapul Station. The selected 61 discharge data points were then transferred to the study basin using the drainage area ratio method. For partial flood frequency analysis, we calculated the average return interval ( $ARI$ ) from these selected floods. The average number of occurrences of peak flood events ( $k$ ) is equal to 61 events divided by 27 years, i.e. 2.26 events per year. After adjusting  $ARI$  to  $k$  times  $ARI$ , the flood values were predicted based on the adjusted  $ARI$  values from various distribution functions.

To select the best fit probability distribution, first of all, alternative probability distribution models need to be analyzed. Continuous probability distributions used in the hydrology sector, including generalized extreme value, Gumbel maximum, log Pearson type III, log Normal (3P), normal, and Weibull 3P were fitted to the processed flood data. The regional empirical methods Hydest and Modified Hydest were used for predicting discharges for different return periods.

### 3.3. Theoretical description

#### 3.3.1. WECS/DHM method (Hydest Method)

The WECS/DHM method was developed by the Water and Energy Commission Secretariat, Department of Hydrology and Meteorology (WECS/DHM) of Nepal. This method is generally used to determine the hydrological features of an ungauged basin for the pre-feasibility study of hydro-electric projects in Nepal. For this purpose, the whole country is considered as a single hydrological region, and the method is suitable for any basin with area  $\geq 100$  km<sup>2</sup>. Hydest is available in the form of an Excel file, which requires input for total catchment area, area of catchment below 5,000 m elevation, area of catchment below 3,000 m elevation, and monsoon wetness index.

Instantaneous peak flood discharges for return periods of 2 and 100 years are:

$$Q_2 = 1.8767 \times (A_{below3000m} + 1)^{0.8783} \quad (2)$$

$$Q_{100} = 14.630 \times (A_{below3000m} + 1)^{0.7342} \quad (3)$$

Peak flood discharge for different return periods:

$$Q_T = e^{(\ln Q_2 + S\sigma)} \quad (4)$$

$$\sigma = \ln(Q_{100}/Q_2)/2.326 \quad (5)$$

where:  $Q_2$  – two-year instantaneous flood in m<sup>3</sup>/s;  $Q_{100}$  – 100-year instantaneous flood in m<sup>3</sup>/s;  $Q_T$  –  $T$ -year instantaneous flood in m<sup>3</sup>/s;  $A_{below 3000m}$  – basin area below 3000 m elevation in km<sup>2</sup>;  $\sigma$  is a parameter;  $S$  is a standard normal variate whose value depends on return periods (Table 1).

Table 1. Values of standard variates based on return periods.

Return Period ( $T$ ) in years	Standard Variate ( $S$ )
2	0
5	0.842
10	1.282
20	1.645
50	2.054
100	2.326
200	2.576

### 3.3.2. Modified Hydest

This method is the updated version of WECS/DHM method in which one more parameter, basin average elevation, is also taken into consideration.

For 2- and 100-year return periods, flood discharges are given by:

$$Q_2 = 2.29(A_{\text{area below 3000m}})^{0.86} \quad (6)$$

$$Q_{100} = 20.7(A_{\text{area below 3000m}})^{0.72} \quad (7)$$

Peak flood discharge for other return periods ( $T$ ):

$$Q_T = e^{\ln Q_2 + S\sigma} \quad (8)$$

$$\sigma = \ln(Q_{100}/Q_2)/2.32 \quad (9)$$

The relationship between  $T$  and  $S$  is shown in Table 1; where:  $Q_2$  – two-year instantaneous flood in  $\text{m}^3/\text{s}$ ;  $Q_{100}$  – 100-year instantaneous flood in  $\text{m}^3/\text{s}$ ;  $Q_T$  –  $T$  year instantaneous flood in  $\text{m}^3/\text{s}$ ;  $A_{\text{below 3000m}}$  – basin area below 3000 m elevation in  $\text{km}^2$ ;  $\sigma$  is a parameter;  $S$  is a standard normal variate whose value depends on return periods.

### 3.3.3. Goodness of fit tests

The goodness of fit technique is a method of examining how a sample of data aligns with a given distribution as its population (Wickramaarachchi 2016). The data were fitted in the EasyFit software (<https://easyfit.soft32.com/>) to check fits for distributions common in hydrology, and then floods for selected return periods were predicted.

## 4. Results and discussions

The discharge data were obtained for 27 years, and maximum annual discharge data were calculated from maximum daily discharge values. Sixty-one peaks above threshold flood over a period of 27 years were taken to predict floods for various return periods. There is no particular rule for establishing the trim level for partial frequency analysis, so we took the lowest annual maximum value as our trim level, and floods greater or equal to that value were fed into probabilistic distribution models. The data were evaluated with the probability distribution functions mentioned above to determine the flood discharges for return periods of 2, 10, 20, 50, 100, and 200 years. Comparisons of the different frequency analysis methods and empirical methods are shown in Tables 2, 3, and 4.

Estimated floods are presented in Tables 2, 3, and 4 for comparative analysis. Figures 3a, 3b, and 3c also provide information about floods of different return periods using various methods. It has been found that the estimated flood values from different methods diverge for higher return periods. The best fits of distribution

functions are shown by ranking in Tables 5 and 6. When designing hydraulic structures for a river like Modi Khola, the choice of distribution to estimate the flood for different return periods should be based on the fit of the distribution to the discharge data. From Table 4, we found that Hydest and Modified Hydest estimated smaller floods than other frequency analysis methods.

Table 2. Floods [ $\text{m}^3/\text{s}$ ] of different return periods using different methods for annual maxima (GEV = generalized extreme value, LP 3 = log Pearson, Type III, LN = log Normal, 3P = Three Parameter).

Return period $T$ (years)	Exceedance Probability ( $p$ ) = $1/T$	Annual Maxima					
		GEV	LP 3	LN (3P)	Gumbel Maximum	Normal	Weibull 3P
2	0.5	321	324	330	392	468	306
10	0.1	756	824	836	1,073	1,062	926
20	0.05	1,085	1,188	1,156	1,333	1,231	1,274
50	0.02	1,762	1,900	1,698	1,670	1,420	1,793
100	0.01	2,554	2,690	2,216	1,923	1,547	2,225
200	0.005	3,715	3,791	2,839	2,174	1,663	2,686

Table 3. Floods [ $\text{m}^3/\text{s}$ ] of different return periods using different methods for peaks above threshold.

$ARI$	$ARI \times k$	Exceedance Probability ( $p$ ) = $1/(ARI \times k)$	Peaks above threshold					
			GEV	LP 3	LN (3P)	Gumbel Maximum	Normal	Weibull 3P
2	4.519	0.221	359	378	389	603	655	446
10	22.593	0.044	801	916	833	1,108	1,009	1,049
20	45.185	0.022	1,180	1,353	1,124	1,316	1,126	1,381
50	112.963	0.009	1,996	2,247	1,615	1,581	1,258	1,866
100	225.926	0.004	3,271	3,571	2,198	1,821	1,367	2,358
200	451.852	0.002	5,032	5,316	2,819	2,025	1,452	2,814

Table 4. Floods [ $\text{m}^3/\text{s}$ ] of different return periods using empirical methods.

Return period $T$ (years)	Empirical Methods	
	Hydest	Modified Hydest
2	216	240
10	438	531
20	535	665
50	670	858
100	778	1016
200	892	1186

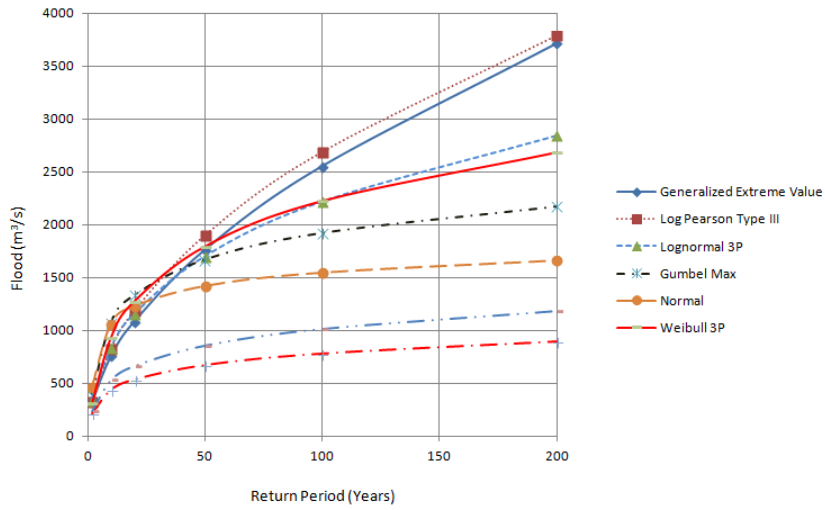


Fig. 3a. Plot of return period vs. flood for different distributions using annual maximum floods.

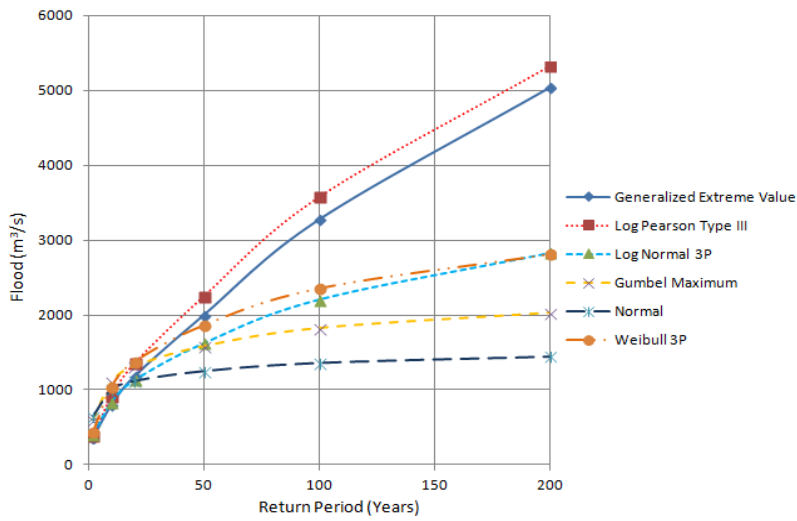


Fig. 3b. Plot of return period vs. flood for different distributions using peaks above threshold.



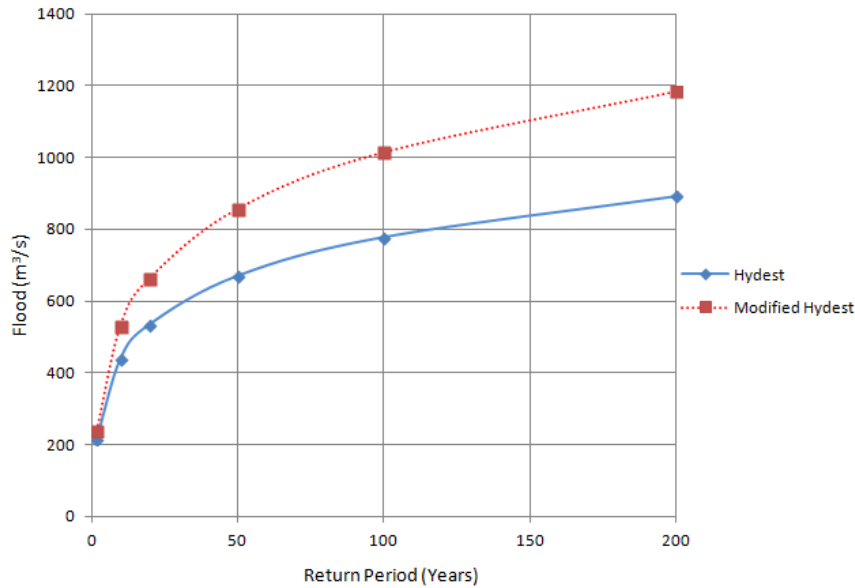


Fig. 3c. Plot of return period vs. flood using empirical methods.

To fit the probability distribution functions with the flood data at a certain significance level ( $\alpha$ )  $\times$  100%, the test statistics and critical values were analyzed. The test statistics for different kinds of tests: Kolmogorov Smirnov (K-S), Anderson Darling (A-D), and chi-squared ( $\chi^2$ ) should be less than the critical value corresponding to significance level  $\alpha$ .

The following tables give the details of the probabilistic analysis carried out for annual maximum floods and peaks above threshold.

Table 5. Fitness of hydrologic distributions for annual maxima.

Test	Distribution	GEV	LP 3	LN (3P)	Gumbel Maximum	Normal	Weibull 3P
K-S	Test Statistic	0.0893	0.10369	0.10661	0.28203	0.28361	0.20639
	Critical value at $\alpha = 0.05$	0.25438	0.25438	0.25438	0.25438	0.25438	0.25438
	Rank	1	2	3	5	6	4
	Decision at 5% significance level	Accept	Accept	Accept	Reject	Reject	Accept
A-D	Test Statistic	0.30003	0.47387	0.37112	2.9499	3.8913	4.8144
	Critical value at $\alpha = 0.05$	2.5018	2.5018	2.5018	2.5018	2.5018	2.5018
	Rank	1	3	2	4	5	6
	Decision at 5% significance level	Accept	Accept	Accept	Reject	Reject	Reject
$\chi^2$	Test Statistic	0.20178	0.66937	1.7897	6.2978	5.3965	N/A
	Critical value at $\alpha = 0.05$	5.9915	5.9915	5.9915	5.9915	5.9915	N/A
	Rank	1	2	3	5	4	N/A
	Decision at 5% significance level	Accept	Accept	Accept	Reject	Accept	Reject

Table 6. Fitness of hydrologic distributions for peaks above threshold.

Test	Distribution	GEV	LP 3	LN (3P)	Gumbel Maximum	Normal	Weibull 3P
K-S	Test Statistic	0.06286	0.08828	0.08642	0.34283	0.32045	0.19077
	Critical value at $\alpha = 0.05$	0.17091	0.17091	0.17091	0.17091	0.17091	0.17091
	Rank	1	3	2	6	5	4
	Decision at 5% significance level	Accept	Accept	Accept	Reject	Reject	Reject
A-D	Test Statistic	0.2523	15.441	0.56032	9.5651	11.419	2.8666
	Critical value at $\alpha = 0.05$	2.5018	2.5018	2.5018	2.5018	2.5018	2.5018
	Rank	1	6	2	4	5	3
	Decision at 5% significance level	Accept	Reject	Accept	Reject	Reject	Reject
$\chi^2$	Test Statistic	3.366	N/A	6.7201	27.89	21.765	14.735
	Critical value at $\alpha = 0.05$	11.07	N/A	11.07	7.8147	7.8147	9.4877
	Rank	1	N/A	2	5	4	3
	Decision at 5% significance level	Accept	Reject	Accept	Reject	Reject	Reject

Each distribution was assigned a rank, the first rank indicating the best fitting distribution, and the last indicating the worst fitting among the distributions used for comparison. The N/A value indicates that the distribution is not applicable for the given data at the 5% significance level and hence rejected. From Tables 5 and 6, generalized extreme value and log Normal (3P) functions are accepted at a 95% confidence interval.

## 5. Conclusion

The frequency analysis of annual maximum and peak discharges above threshold for identifying the best fit probability distribution was performed using normal, Gumbel maximum, log Pearson type III, log Normal (3P), generalized extreme value, Hydest, Modified Hydest and Weibull 3P distributions. Most of the research on flood flow estimations has been conducted at gauged locations, with very little research for ungauged locations. Estimating flood discharges at different locations in an ungauged catchment requires a hydrologically similar reference catchment, the choice of which is a challenge.

For our study basin, based on K-S, A-D and  $\chi^2$  tests, we found that GEV and LN (3P) are well fitted compared to other hydrological distributions. Selection of the suitable distribution also depends upon financial considerations as well as risk optimization. Designing hydraulic structures based on the design floods from the GEV distribution may not be cost-effective because of the large predicted flood values for larger return periods, so the LN (3P) can be suggested as an alternative. This study indicates that at least GEV and LN (3P) distributions are better suited for flood frequency analysis of an ungauged basin where the geographical and hydrological features are similar to that of the study basin.

Nevertheless, the limited data available for both spatial and temporal resolution for the gauged basin should be acknowledged, and the hydrological similarities between the catchments should be carefully assessed; these

characteristics can vary drastically from one place to another. A cross-check of transferred data should be done where another similar catchment is available to enhance the credibility of the data while selecting the appropriate method. Hydest and Modified Hydest are commonly used in Nepal for preliminary assessment of the hydrology of ungauged basins.

The methodology we used in this study can be adopted to study the hydrology of an ungauged site in the basin where the hydrological and meteorological stations are very sparse. Because of epistemic as well as aleatory uncertainties, we cannot exactly quantify the hydrological characteristics of even the gauged river basins. Moreover, for ungauged basins, epistemic uncertainty is significantly high. Therefore, we should be skeptical about our probabilistic analysis while selecting the appropriate distributions, and thus a comparative analysis of hydrological distributions for different tests is recommended. Moreover, choosing a probability distribution function does not depend only on its goodness of fit but also on the optimization of hydraulic structures based on safety and cost. Therefore, this study helps in estimating return floods for various return years in ungauged basins and in selecting design flood for engineering structures, in developing hydrology models, agriculture, flood management, river training works, and environmental studies.

## References

- Alam S., Khan M.S.M.K., Rahat S.H., 2016, A study on selection of probability distributions of extreme hydrologic parameters for the peripheral river system of Dhaka city, [in:] Proceedings of Academics World 15<sup>th</sup> International Conference, Bangkok, Thailand, 29-34
- Archfield S.A., Vogel R.M., 2010, Map correlation method: selection of a reference streamgage to estimate daily streamflow at ungauged catchments, *Water Resources Research*, 46, 1-15, DOI: 10.1029/2009WR008481
- Bobee B., Cavidas G., Ashkar F., Bernier J., Ramussen P., 1993, Towards a systematic approach to comparing distributions used in flood frequency analysis, *Journal of Hydrology*, 142 (1-4), 121-136, DOI: 10.1016/0022-1694(93)90008-W
- Gianfagna C.C., Johnson C.E., Chandler D.G., Hofmann C., 2015, Watershed area ratio accurately predicts daily streamflow in nested catchments in the Catskills, New York, *Journal of Hydrology: Regional Studies*, 4, 583-594, DOI: 10.1016/j.ejrh.2015.09.002
- Izinyon O.C., Ihimekpen N., Igbino G.E., 2011, Flood frequency analysis of Ikpoba River catchment at Benin City using Log Pearson Type III distribution, *Journal of Emerging Trends in Engineering and Applied Sciences*, 2 (1), 50-55
- Ojha G.S.P., Berndtsson R., Bhunya P., 2008, *Engineering hydrology*, Oxford University Press, New Delhi, India
- Rahman A.S., Rahman A., Zaman M.A., Haddad K., Ahsan A., Imteaz M., 2013, A study on selection of probability distribution for at-site flood frequency analysis in Australia, *Natural Hazards*, 69, 1803-1813, DOI: 10.1007/s11069-013-0775-y
- Reich B.M., 1961, Short duration rainfall intensity in South Africa, *South African Journal of Agricultural Science*, 4 (4), 589-614
- Reich B.M., 1963, Short-duration rainfall-intensity estimates and other design aids for regions of sparse data, *Journal of Hydrology*, 1 (1), 3-28, DOI: 10.1016/0022-1694(63)90029-5
- Rijal S.P., 2007, Land holding and livelihoods: a synthesis from Modi Khola watershed, Nepal, *The Third Pole*, 5 (7), 43-51, DOI: 10.3126/ttp.v5i0.1952
- Sapkota A., Pandit H.P., Shrestha R., 2016, Hydraulic and sediment handling performance assessment of Rani Jamara Kulariya Irrigation Project (RJKIP) by conjunctive use of 1D and 3D simulation models, *Naresuan University Engineering Journal*, 11 (1), 1-6, DOI: 10.14456/nuej.2016.1

- Singh U.K., Kumar B., 2018, Climate change impacts on hydrology and water resources of Indian river basins, *Climate World Environment*, 13 (1), 32-43
- Tao D.Q., Nguyena V.T.V., Bourque A., 2002, On selection of probability distributions for representing extreme precipitations in Southern Quebec, [in:] *Annual Conference of the Canadian Society for Civil Engineering*, Montreal, Quebec, Canada, 1-8
- Wickramaarachchi T.N., 2016, Frequency analysis of flood discharge in Gin River at Tawalama, Sri Lanka, [in:] *7<sup>th</sup> International Conference on Water Resources and Environment Research, ICWREER 2016*, Kyoto, Japan, 7, 1-6