Impact of Non-Stationarity on Flood Frequency Estimates for a Himalayan Sub-basin

Jagadish Prasad Patra*, Rakesh Kumar, Pankaj Mani, Sanjay Kumar
National Institute of Hydrology, Roorkee, India
**Corresponding author e-mail: patra.nih@gmail.com*

Abstract: Estimation of design floods of various return periods is essential for design of drainage systems, bridges, embankments, dams and flood risk assessment etc. Traditionally, frequency analysis of peak flood data is carried for estimation of floods of various return periods. The general assumption of such analysis is that parameters describing the flood frequency distribution are stationary and do not change over time. However, various researchers have raised question on such assumption as the frequency of hydrological extremes has been changing and is likely to continue to change because of various anthropogenic activities and climate change. There is need to analyse the non-stationarity aspects of discharge data and evaluate time-dependent parameters in flood frequency distributions. In this study 45 years of annual maximum flood series of a Himalayan sub-basin having catchment area of 9031 km^2 is analysed for estimation of design floods of various return periods using stationary and non-stationary approach. In first case, assuming stationarity, General Extreme Value (GEV) distribution is fitted and parameters are estimated using three methods viz. Maximum Likelihood, Bayesian and L-moments. The results shows that at low return periods up 25 years the variation in estimated design flood is within 6%, which increases to 17% for 100 year return period. Further, along with GEV, eleven additional frequency distributions viz. Extreme Value (EV1), Logistic (LOS), Generalized Logistic (GLO), Normal (NOR), Generalized Pareto (GPA), Generalized Normal(GNO), Uniform (UNF), Exponential (EXP), Pearson Type-III (PT3), Kappa (KAP) and Wakeby (WAK) have been used to identify robust distribution based on the L-moments ratio diagram and the $Z_i^{dist}-$ statistic criteria. In this case, the GLO is identified as the robust distribution. It is observed that the design flood estimated using fitted L-moments based GLO and GEV distributions are very similar and the maximum variation is within 2% up to 100 year return period. However, the estimate is 8% higher in case of 500 year return period for the GLO distribution. The Mann-Kendall test suggests significant increasing trend in the data series. In the second case, various combinations of time dependant EV1 and GEV parameters (location and scale) are estimated using the Maximum Likelihood method. The EV1 distribution with linear time dependant location parameter is select to be the best model based on lower Akaike Information Criterion (AIC). The design flood corresponding to years 2050 and 2100 are estimated using the fitted non-stationary EV1 distribution. The design flood for 25, 50, 75 and 100 year return period floods are estimated using stationary assumption would become about 9, 25, 49 and 80 year return period in the year 2050 considering non-stationary. Further, in year 2100 the corresponding return periods are estimated to be 3, 7, 12 and 20 years, respectively. However, it is to be noted that the 100 year return period flood estimated using L-moments approach (GLO as robust distribution) is 6030.4 m^3/s in comparison to 5937.7 m^3/s estimated for the year 2100 using non-stationary EV1 distribution. The results show that along with non-stationarity, various other aspects like selection of parameter estimation method, goodness of fit criteria, uncertainty of parameter estimation, covariates etc. should be studied, while estimating design flood in respect to design life of a hydraulic structure.

Keywords: Design flood; Non-Stationarity; L-moments; Trends; Akaike Information Criterion.

1. Introduction

Estimation of design floods of various return periods is essential for design of drainage systems, bridges, embankments, dams and flood risk assessment etc. Traditionally, frequency analysis of peak flood data is carried for estimation of floods of various return periods. The general assumption of such analysis is that parameters describing the flood frequency distribution are stationary and do not change over time. The traditional assumption of stationary in flood frequency analysis is being questioned by various researchers in the scenarios of climate change. Cyclical or monotonic non-stationarities pose a serious challenge to flood frequency and risk analysis and flood control design and practice. Nonstationary frequency analysis can integrate trends and climate variability by introducing covariates in the distribution parameters (Ouarda and Charron 2019). Most of the non-stationary flood frequency analysis methods proposed in the literature involve parameters of a chosen distribution that are dependent on time. The location and/or scale parameters of the probability distribution are expressed as a function of covariates and shape parameters are generally considered to be constant over time (Katz et al., 2002; Adlouni et al., 2007; Cheng et al., 2014; Thiombiano et al. 2018). Linear and nonlinear conditional quantiles are observed in various studies and further research is required on the physical mechanism behind these simple and complex interactions (Thiombiano et al. 2018). Further, the problem with time varying distribution parameters is that, it is difficult to justify why they would continue to change in the future exactly in the same way they did in the past. Agilan and Umamahesh (2016) used Global warming, ENSO cycle, Indian Ocean Dipole, local temperature changes and urbanization as covariate along with time to developed a non-stationary rainfall IDF curves for Hyderabad city in India. Covariate uncertainty is sometimes significant and in cases it is nearly equivalent to the parameter uncertainty (Agilan and Umamahesh, 2018). Akaike information criterion (AIC) and the Bayesian information criterion (BIC) are commonly used to compare various nonstationary models.

In this study, design flood for a Himalayan sub-basin in Uttarakhand is estimated using stationary and non-stationary approach. The stationary General Extreme Value (GEV) distribution is fitted to annual maximum flood series and parameters are estimated using three methods viz. Maximum Likelihood, Bayesian and L-moments. Further, along with GEV, eleven additional frequency distributions have been considered and robust distribution is selected based on the L-moments ratio diagram and the Z_i^{dist}–statistic criteria. Moreover, various combinations of time dependant EV1 and GEV parameters (location and scale) are also estimated using the Maximum Likelihood method and best model is selected best model based on lower Akaike Information Criterion (AIC).

2. Study area and Methods

The index map of the study area showing Rudraprayag gauging site in Alakhananda river in shown in Figure 1.The catchment area of Alakhananda river at the gauging site is about 9031 km^2. Annual maximum peak flood data of 45 years are used for flood frequency analysis.

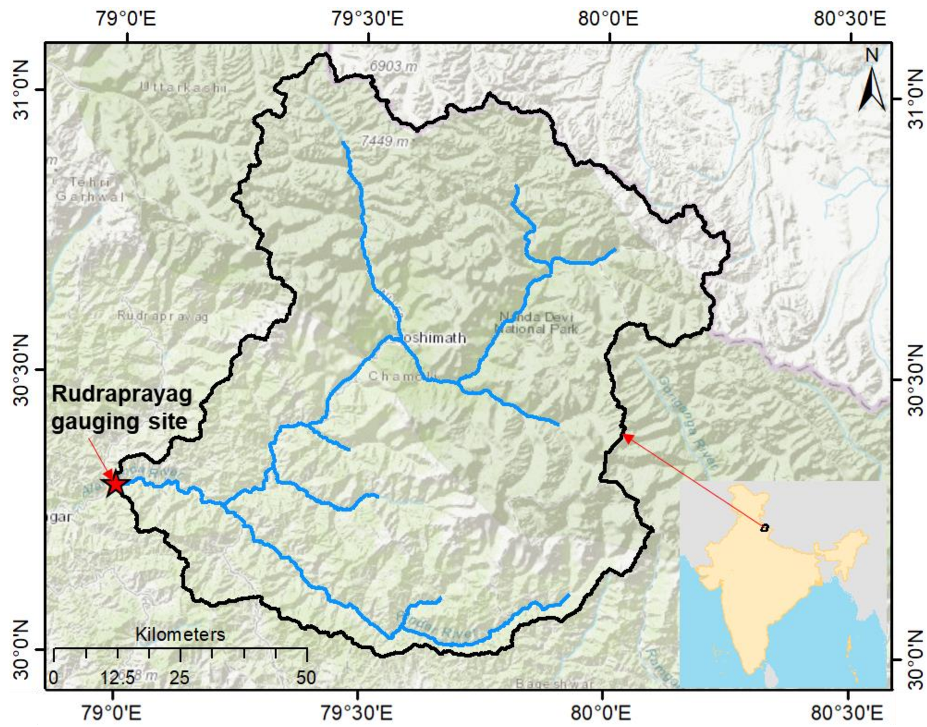

Figure 1: Index map of the study area.

L-Moments based rainfall frequency analysis

At site and regional rainfall frequency analysis of 1-day annual maximum rainfall values of seven raingauge stations lying within and near study area has been carried out using the L-moments approach as described elsewhere (Hosking &Wallis, 1997; Kumar & Chatterjee, 2005). Twelve frequency distributions viz. extreme value (EV1), general extreme value (GEV), logistic (LOS), generalized logistic (GLO), normal (NOR), generalized pareto (GPA), generalized normal(GNO), uniform (UNF), exponential (EXP), pearson Type-III (PT3), kappa (KAP) and wakeby (WAK) have been used to identify robust distribution based on the L-moment ratio diagrams and the Z_i^{dist}–statistic criteria. The details about these distributions and relationships among parameters of these distributions and L-moments are available in literature (Hosking &Wallis, 1997).

Non-Stationary frequency analysis.

The generalized extreme value (GEV) distribution has theoretical justification for fitting to block maxima data given by:

$$G(z) = \exp\left[-\left\{1+\xi\left(\frac{z-\mu}{\sigma}\right)\right\}_{+}^{-1/\xi}\right] \quad (1)$$

The Eq. 1 envelops three types of distributions depending on the sign of the shape parameter (ξ). The heavy-tailed Fréchet distribution results from $\xi > 0$, and the upper bounded Weibull distribution When $\xi < 0$. The Gumbel type is obtained by taking the limit as $\xi \rightarrow 0$ given by:

$$G(z) = \exp\left[-\exp\left\{-\left(\frac{z-\mu}{\sigma}\right)\right\}\right], \ -\infty < z < \infty \quad (2)$$

If it is believed that the extremes of the data are not stationary, then it is possible to incorporate this information into the parameters as given below (Gilleland and Katz, 2016).

$$\mu(t) = \mu_0 + \mu_1 t + \mu_2 t^2$$
$$\sigma(t) = \sigma_0 + \sigma_1 t \quad (3)$$

In the above models, each parameter varies with time; the location parameter follows a quadratic formula of time, the scale parameter a linear function. There are different methods available for performing parameter estimation including: Method of Moments Estimation (MME), Probability Weighted Moments (PWM) or equivalently L-Moments (LM), Maximum Likelihood Estimation (MLE), and Bayesian methods. MLE allows one to easily incorporate covariate information into parameter estimates. As no analytic solution to the optimization problem exists, the MLE solution can be found using numerical routines.

Various goodness–of–fit indicators (diagnostic plots) were inspected to check the overall performance of the fitted models, e.g. a quantile–quantile plot, a probability plot, etc. The likelihood ratio test can be used to compare the goodness-of-fit of two hierarchically nested models. The Akaike information criterion (AIC) and the Bayesian information criterion (BIC) are also widely applied for model selection. More details of model selection procedures can be found in Claeskens and Hjort (2008). The model associated with the smallest AIC value indicates the best model performance.

3. Result and discussions

The analyses were carried out for the selected gauging stations in a Himalayan river in India. As a first step the annual maximum (AM) flood series is assumed to be stationarity and General Extreme Value (GEV) distribution is fitted and parameters are estimated using three methods viz. Maximum Likelihood, Bayesian and L-moments. The various parameters of the GEV distribution estimated using different approach are given in Table 1. The estimated flood for various return periods using GEV distribution is shown in Figure 2. It is observed that the floods for lower return periods up to 25 year are of similar magnitude estimated using various parameter estimation method. At low return periods up 25 years the variation in estimated design flood is within 6%, which increases to 17% for 100 year return period as there is variation in the estimates for higher return periods.

Table 1: Value of different parameters of GEV distribution

Method	**Values of parameter**		
	μ	**σ**	**ξ**
MLM	1856.516	528.098	0.118
Bayesian	1841.148	559.667	0.147
L-Moment	1836.509	453.514	0.268

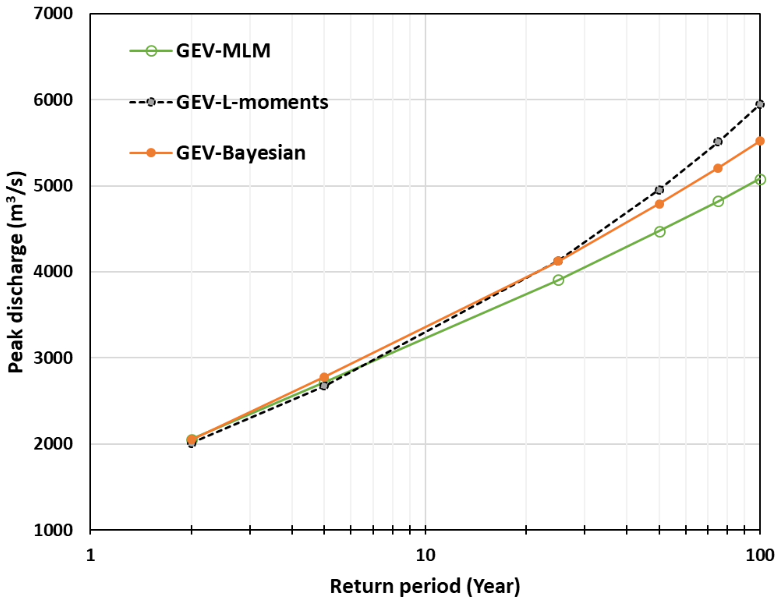

Figure 2: Estimated flood for various return periods using GEV distribution

Further flood frequency analysis has been carried out using the L-moments approach consider other distributions also. The robust frequency distribution is identified based on Z_i^{dist} Statistic and L-moment ratio diagram. The Z_i^{dist} Statistic for various distributions are given in Table 2 and the L-moment ratio diagram shown in Figure 3. The various distributions having Z^{dist} –statistic value lower than 1.64 are accepted at the 90% confidence level. In this case GLO distribution is identified as the robust frequency distribution The estimated maximum floods for 10, 25, 100 and 500 return periods are 3197.08 m^3/s, 4082.77 m^3/s, 6030.39 m^3/s, 9785.55 m^3/s respectively. It is observed that the design flood estimated using fitted L-moments based GLO and GEV distributions are very similar and the maximum variation is within 2% up to 100-year return period. However, the estimate is 8% higher in case of 500 year return period for the GLO distribution.

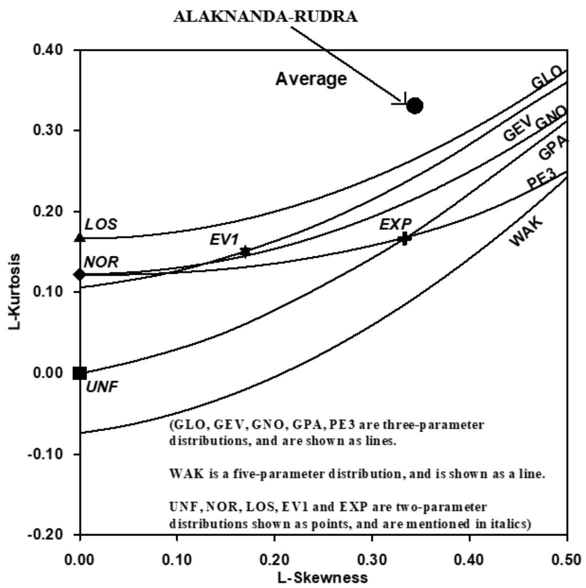

Figure 3: L-moment ratio diagram for the AM flood series of gauging site.

Table 2: Z_i^{dist} – Statistic of various distributions AM flood series

Distribution	**Z_i^{dist} –statistic**
Generalized logistic (GLO)	1.45
Generalized Extreme Value (GEV)	1.66
Generalized Normal (GNO)	1.92
Generalized Pareto (GPA)	2.31
Pearson Type III (PE3)	2.36

In the second case, various combinations of time dependant EV1 and GEV parameters (location and scale) are estimated using the Maximum Likelihood method. The EV1 distribution with linear time dependant location parameter is select to be the best model based on lower Akaike Information Criterion (AIC). The estimated parameter values are presented in Table 3. The fitted models are ranked based on AIC values. The Comparison of estimated flood quantiles for selected non-stationary (NS) and stationary (S) models are shown in Figure 4. Moreover, the flood quantiles for year 2050 and 2100 using selected NS model is also plotted. It is observed that there is wide variations in the estimated flood quantiles using stationary (S) and non-stationary (NS) models for lower return periods for future scenarios. The design flood corresponding to years 2050 and 2100 are estimated using the fitted non-stationary EV1 distribution as shown in Figure 5. It may be noticed that the design flood for 25, 50, 75 and 100 year return period floods are estimated using stationary assumption would become about 9, 25, 49 and 80 year return period in the year 2050 considering non-stationary. Further, in year 2100 the corresponding return periods are estimated to be 3, 7, 12 and 20 years, respectively (Figure 5). However, it is to be noted that the 100 year return period flood estimated using L-moments approach (GLO as robust distribution) is 6030.4 m^3/s in comparison to 5937.7 m^3/s estimated for the year 2100 using non-stationary EV1 distribution. Predicting the design flood for the life span of structure in the future using time dependant non-stationary models needs to be considered with care. It is possible that changes in the future will not be exactly the same as those in the past, especially for periods longer than 10 years into the future.

Table 3: Parameters of stationary and non-stationary models

Distribution	Model	Model parameters					AIC	Rank
		μ_0	μ_1	σ_0	σ_1	ξ		
NS GEV	μ(t) linear	1808.58	6.78	490.67		0.09	715.60	4
NS GEV	μ(t), α (t) linear	1559.07	16.15	552.54	-2.72	0.16	714.08	2
NS GEV	α (t) linear	1836.51	-	453.66	3.74	0.09	719.17	5
NS Gumble	μ(t) linear	1621.10	14.70	522.80	-		713.58	1
NS Gumble	μ(t), α (t) linear	1689.65	12.23	421.21	4.66		714.62	3
Gumble (S)	MLM	1904.18	-	556.77	-		719.57	6

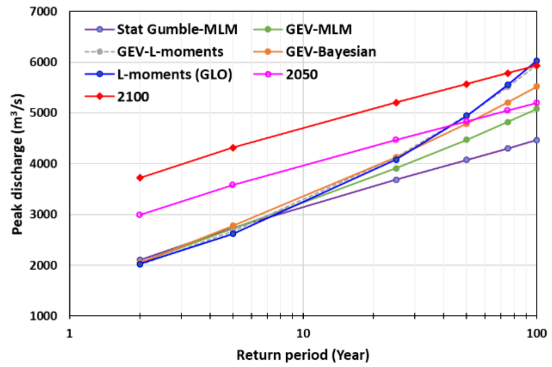

Figure 4: Comparison of estimated flood quantiles for selected non-stationary (NS) and stationary (S) models

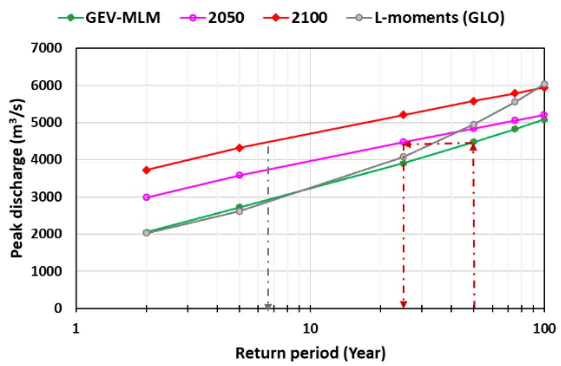

Figure 5: Effects of NS in estimated flood quantiles

5. Conclusions

Estimation of design floods is an essential step in design of drainage systems, bridges, embankments, dams and flood risk assessment etc. The traditional assumption of stationary in flood frequency analysis is being questioned by various researchers in the scenarios of climate change. Nonstationary frequency analysis can integrate trends and climate variability by introducing covariates in the distribution parameters. In this study 45 years of annual maximum flood series of a Himalayan sub-basin having catchment area of 9031 km^2 is analysed for estimation of design floods of various return periods using stationary and non-stationary approach. In first case, assuming stationarity, General Extreme Value (GEV) distribution is fitted and parameters are estimated using three methods viz. Maximum Likelihood, Bayesian and L-moments. The results shows that at low return periods up 25 years the variation in estimated design flood is within 6%, which increases to 17% for 100 year return period. Tis highlight importance of using a suitable parameter estimation method and uncertainty associated with the estimated parameters of a fitted frequency distribution. Further, along with GEV, eleven additional frequency distributions viz. Extreme Value (EV1), Logistic (LOS), Generalized Logistic (GLO), Normal (NOR), Generalized Pareto (GPA), Generalized Normal(GNO), Uniform (UNF), Exponential (EXP), Pearson Type-III (PT3), Kappa (KAP) and Wakeby (WAK) have been used to identify robust distribution based on the L-moments ratio diagram and the Z_i^{dist}–statistic criteria. Based on the lowest Z_i^{dist}–statistic and L-moments ratio diagram , the GLO is identified as the robust frequency distribution. It is observed that the design flood estimated using fitted L-moments based GLO and GEV distributions are very similar and the maximum variation is within 2% up to 100 year return period. However, the estimate is 8% higher in case of 500 year return period for the GLO distribution.

In second case, various combinations of time dependant EV1 and GEV parameters (location and scale) are estimated using the Maximum Likelihood method. The EV1 distribution with linear time dependant location parameter is select to be the best model based on lower Akaike Information Criterion (AIC). The design flood corresponding to years 2050 and 2100 are estimated using the fitted non-stationary EV1 distribution. It is observed that the design flood for 25, 50, 75 and 100 year return periods estimated using stationary assumption would become about 9, 25, 49 and 80 year return period in the year 2050 if non-stationary is considered. Further, in year 2100 the corresponding return periods would be 3, 7, 12 and 20 years, respectively. However, it is to be noted that the 100 year return period flood estimated using L-moments approach (GLO as robust distribution) is 6030.4 m^3/s in comparison to 5937.7 m^3/s estimated for the year 2100 using non-stationary EV1 distribution. The results show that along with non-stationarity, various other aspects like selection of parameter estimation method, goodness of fit criteria, uncertainty of parameter estimation, covariates etc. should be studied, while estimating design flood in respect to design life of a hydraulic structure. Predicting the design flood for the life span of structure in the future using time dependant non-stationary models needs to be considered with care. It is possible that changes in the future will not be exactly the same as those in the past, especially for periods longer than 10 years into the future.

References

Adlouni, S. E., Ouarda, T. B. M. J., Zhang, X., Roy, R., and Bobée, B. (2007). Generalized maximum likelihood estimators for the nonstationary generalized extreme value model. *Water Resources Research*, *43*(3). https://doi.org/10.1029/2005WR004545

Agilan, V., and Umamahesh, N. V. (2016). Is the covariate based non-stationary rainfall IDF curve capable of encompassing future rainfall changes? *Journal of Hydrology*, *541*, 1441–1455. https://doi.org/10.1016/j.jhydrol.2016.08.052

Agilan, V., and Umamahesh, N. V. (2018). Covariate and parameter uncertainty in non-stationary rainfall IDF curve. *International Journal of Climatology*, *38*(1), 365–383. https://doi.org/10.1002/joc.5181

Cheng, L., AghaKouchak, A., Gilleland, E., and Katz, R. W. (2014). Non-stationary extreme value analysis in a changing climate. *Climatic Change*, *127*(2), 353–369. https://doi.org/10.1007/s10584-014-1254-5

Claeskens, G. and Hjort, N. (2008) Model Selection and Model Averaging. Cambridge University Press, Cambridge. http://dx.doi.org/10.1017/CBO9780511790485

Hosking, J. R. M., and Wallis, J. R. (1993). Some statistics useful in regional frequency analysis. *Water Resources Research*, *29*(2), 271–281. https://doi.org/10.1029/92WR01980

Hosking, J.R.M. abd Wallis, J.R. (1997). Regional frequency analysis: an approach based on L-moments. Cambridge University Press, Cambridge.

Katz, R. W., Parlange, M. B., and Naveau, P. (2002). Statistics of extremes in hydrology. *Advances in Water Resources*, *25*(8), 1287–1304. https://doi.org/10.1016/S0309-1708(02)00056-8

Ouarda, T. B. M. J., and Charron, C. (2019). Changes in the distribution of hydro-climatic extremes in a non-stationary framework. *Scientific Reports*, *9*(1), 8104. https://doi.org/10.1038/s41598-019-44603-7

Thiombiano, A. N., St- Hilaire, A., Adlouni, S.-E. E., and Ouarda, T. B. M. J. (2018). Nonlinear response of precipitation to climate indices using a non-stationary Poisson-generalized Pareto model: case study of southeastern Canada. *International Journal of Climatology*, *38*(S1), e875–e888. https://doi.org/10.1002/joc.5415