Research article

Bangladesh Journal of Ecology

ISSN (Print): 2708-6356, ISSN (Online): 2708-6364, URL: bsmrau.edu.bd/bjoe
Bangladesh J. Ecol. (2021) 3 (1) : 31-37

STATISTICAL MODELING AND RETURN LEVELS OF EXTREME RAINFALL EVENTS IN FLOOD PRONE AREA OF BANGLADESH

Nasrin Sultana* **and Mohammad Nazmol Hasan**

Department of Statistics, Bangabandhu Sheikh Mujibur Rahman Agricultural University, Gazipur-1706, Bangladesh

***Correspondence e-mail:** nasrin@bsrmau.edu.bd

Received: 16 March 2021, Revised: 17 May 2021, Accepted: 22 May 2021

ABSTRACT

The connection between lives and water is well-known where rain is the important source of water for the plants and animals on earth. However, excessive rainfall raises water level and sometimes results serious floods, flash floods, river/sea flooding etc. In this regard, efficient planning considering characteristics of heavy rainfall is essential to protect lives. This article intended to determine the best-fit probability distributions of annual maximum rainfall from eight flood prone areas in Bangladesh. Consequently, the Lognormal, Generalized extreme value distribution and Inverse-Gaussian distribution were considered and maximum likelihood methods was utilizing to estimate model parameters. The power of the fitted model assessed through some goodness of fit test such as Kolmogorov-Smirnov statistic, Cramér–von Mises criterion and Anderson-Darling statistic. Furthermore, some of the graphical measures such as QQ plot, PP plot and density plots examined the competence of fitted model. The lognormal and Inverse-Gaussian distribution had similar inference, resulted best fitted distribution for almost all station. The return levels of 10-year, 20-year, 50-year and 100-year return period estimated from the best fitted models of this study were essential measures in applied point of view. Hence, protective measures for higher return levels areas may save from the adverse loss of agriculture, ecosystems, bio diversities and livelihood.

Keywords: Return level, probability distribution, GEV, trend, flood

Introduction

Lives of living being including plants and animals on earth depends on safe water. Plants need water to grow and stay alive. They use water directly when they capture the light energy from the sun and transform it into their useable energy. Animals need fresh water for their bodies to function like regulation of temperature, nutrient uptake, removing wastes, body weight, and health. Though 71 percent of the Earth's surface is covered with water, 97 percent of this surface water is salty sea water. Therefore, plants and animals have to rely on the rain water directly or indirectly. However, excessive rainfall is threat to lives specially plants and animals in some areas of the world including Bangladesh. These excessive rainfall rise water level and results serious floods, flash flood, river/sea flooding and rainwater are three major types occurs almost every year in Bangladesh (Brammer 1989) and experienced severe flood recorded in 1988, 1998, 2004, and 2007, 2014, 2015.

The intensity of heavy rainfall events in northeast India as well as in Bangladesh are increasing (May 2004). This rising trend of rainfall with north and northwest part with that might increase of having flash flood in that region and experienced flood almost every year due to excessive rainfall during early monsoon. Heavy pre-monsoon and monsoon rainfall ponded low land which increase the water levels in the attached rivers also trigger rain related flood (Mirza and Ahmed 2005) and hillside area in Chittagong causes landslides in the recent years. Excessive rainfall pollutes surface water due to organic and inorganic components create deficiency of safe drinking water. Outbreak of several water borne diseases including diarrhea, cholera, and other intestine diseases increase terribly during and after floods (Brouwer *et al*. 2007, Hashizume *et al*. 2008, McMichael *et al*. 2006). The alteration on extreme events is systematically connected to ecosystems and diversity on productivity, reproduction, phenology, nutrient cycling and community resistance to invasion are remarked as the ecological effects due to extreme weather phenomenon (Jentsch and Beierkuhnlein 2008). Determining appropriate probability distribution of extreme rainfall is truly an important task for risk assessment to protect life and property and environment.

Most commonly distribution that are applied to model rainfall around the world are LN2, Log-Pearson type 3 (LP3), generalized extreme value (GEV) Pearson type 3 (P3) Gumbel (GUM) found in literature.

The annual extreme precipitation in Bangladesh follows the generalized extreme value and generalized gamma four parameter distributions (Khudri *et al.* 2013), monthly maximum data followed GEV, P3 and LP3 distributions (Alam *et al.* 2018) and GEV for the rainfall annual maximum rainfall based on a daily sample of coastal areas (Sultana and Roni 2015). Whether the probabilistic behavior of extreme/heavy rainfall of flood prone area is not studied so far. The people living in the flood prone areas suffer tremendously, not having defensive measures as well as have very limited capability to grip the destruction of property and income. Statistical inference regarding extreme rainfall of flood prone areas help policy maker to take appropriate plans and policies for reducing risk and damage due to heavy rainfall. Therefore, the aim of this study is to determine the best fitted probability distribution of the annual maximum rainfall based on a daily sample data of flood prone regions of Bangladesh. Finally, estimation of the return level of the rainfall related flood risk of next 10, 20, 50 and 100 years would help to measure the extent of heavy return levels of flood prone area of Bangladesh.

Materials and Methods

This study considered lognormal, generalized extreme value distribution and Inverse-Gaussian distribution to evaluate the best fit probability distribution of annual maximum rainfall (Table 1). At first stationary and trend of the annual maximum rainfall were assessed. The parameters of the distribution were estimated by maximum likelihood approach and both numerical and graphical approach of goodness of fit tests for model selection were analyzed. The return levels were estimated from best fitted model for 10, 20, 50 and 100 years return periods. In extreme value analysis, the return level or quantiles of the fitted parametric distribution is more concern than parameter. The estimate of return level of rainfall for a given probability p associated with $\frac{1}{p}$ return period is

Table 1: Probability density function and parameter used in this study

Distribution	Probability density function	Parameters
Log-normal	$f(x,\mu,\sigma)=\frac{1}{x\sigma_y\sqrt{2\pi}}exp\left[-\frac{1}{2\sigma_y^2}(\ln(x)-\mu_y)^2\right]$ $y<x<\infty$	μ = location parameter σ =scale parameter
GEV	$f(x;\gamma,\mu,\sigma)=exp\left\{-\left(1+\gamma\left(\frac{x-\mu}{\sigma}\right)\right)_+^{-\frac{1}{\gamma}}\right\}$ $\{1+\gamma\frac{x-\mu}{\sigma}>0 \text{ for } \gamma\neq 0\}$ $-\infty<x<\infty \; for \; \gamma=0$	μ = location parameter σ =scale parameter γ= shape parameter $\gamma>0$ type I (*Fréchet distribution*) $\gamma\to 0$ type II (*Weibull distribution*) $\gamma<0$ type III *(Gumbel)*
Inverse-Gaussian	$f(x,\mu,\lambda)=\sqrt{\frac{\lambda}{2\pi x^3}}e^{-\lambda(X-\mu)^2/2X\mu^2}$ $0<x<\infty$	μ = location parameter λ = scale parameter

contented with the following relation $G(z_p) = 1 - p$ For a random variable *z*, the probability of occurring an event $(z \geq z_p)$ in a given p year is

$$P(Z \geq Z_p) = \frac{1}{p} or\ p = \frac{1}{1 - p(z < z_p)}$$

The 50- or 100-year return level can be estimated 98th and 99th percentiles of the fitted distribution, respectively (Wilks 1993).

Extreme rainfall model adequacy checking: The Augmented –Dickey-Fuller (ADF) test and Kwiatkowski–Phillips–Schmidt–Shin (KPSS) unit root test (Gujrati 2003) were applied to check stationary of the maximum rainfall. Non parametric Mann–Kendall test applied to check the trend in the dataset or the data are randomly ordered in time. (Drapela *et al.* 2011)

Goodness-of-Fit test: The validity of model is assessed by goodness-of-fit test statistics and goodness of fit criteria. Several graphical methods such as histograms, and Quantile-Quantile plots (Q-Q plots) and density plots and statistical significance test was conducted. The null hypothesis regarding this test was

H_0 : The annual maximum rainfall data follow the specified distribution

H_1 : The annual maximum rainfall data do not follow the specified distribution

Kolmogorov-Smirnov (K-S) test and Anderson-Darling (Anderson and Darling,1954) goodness of fit test were applied to determine power of the fitted model. Furthermore, Akaike Information Criteria (AIC) and Bayesian Information Criteria (BIC) also observed to find the model's goodness-of-fit.

Kolmogorov-Smirnov test: The Kolmogorov-Smirnov test statistic (D) is concerned the maximum vertical distance between theoretical and hypothesized or empirical distribution functions. The K-S statistics estimates the maximum difference between hypothesized distribution $Z_{(i)} = (F(x_{(i)}\hat{\theta}))$ and empirical cumulative distribution function, $F_n(x_{(i)})$ where $x_{(i)}$ is the order statistics of observation x and make decision whether the sample comes from a hypothesized continuous distribution (Chakravarti *et al*. 1967)

$$D^+ = \max(i/n - Zi);\ D^- = \max[-(i-1)/n];\ D = \max(D^+, D^-)$$

Anderson-Darling test: The A-D test, was first introduced by Anderson & Darling and suitable for the sample having comparatively large extremes values. It estimates the weighted squared difference between hypothesized distribution $Z_{(i)} = (F(x_{(i)}\hat{\theta}))$ and empirical cumulative distribution function, $F_{(n)} = (x_{(i)})$.

The test statistics

$$A^2 = -n - \frac{1}{n}\sum_{i=1}^{n}(2i-1)\{\ln F(X_i) + ln[X_{n-i+1}]\}$$

Where $x_{(i)}$ is the order statistics of observation and the test used to compare the fit of an observed cumulative distribution function to an expected cumulative distribution function.

Cramér–von Mises criterion: The Cramér–von Mises criterion is another goodness of fit test of a hypothesized distribution function $Z_{(i)} = (F(X_{(i)}\hat{\theta}))$ compared to a given $F_n(x_{(i)})$ empirical distribution function considering minimum distance.

$$T = \frac{U}{NM(N+M)} + \frac{4MN-1}{6(M+N)} where$$

$$U = N\sum_{i=1}^{N}(r_i - i)^2 + m\sum_{j=1}^{M}(s_j - j)^2$$ Anderson 1962)

Data: Daily rainfall data recorded over 1948–2014 from eight metrological stations were collected from Bangladesh Meteorological Department (BMD). Srimangal, Chittagong, Cox's Bazar were selected as a flash flood prone area and Barisal, Comilla, Dhaka, Faridpur and Rangpur were taken as river flood area for this analysis. The daily rainfall data was recorded in millimeter and missing values are remove from the data set before analysis. Model parameter and return level estimation and further empirical analysis were conducted using R.

Results and Discussion

The visual review of the annual maximum rainfall appeared in Figure 1 illustrated that the maximum amount of rainfall varies across the study period in Barisal, Comilla, Chittagong, Cox's bazar, and Dhaka districts. Meanwhile, the less variation observed in Faridpur, Srimangal, and Rangpur district. Additionally, no trend observed graphically in the time series of annual maximum rainfall.

The maximum mean annual rainfall occurred in Cox's bazar with high variability and maximum 146 mm rainfall was recorded in 2005 (Table 2). The highest amount 98% quantile of daily rainfall of measured 141.

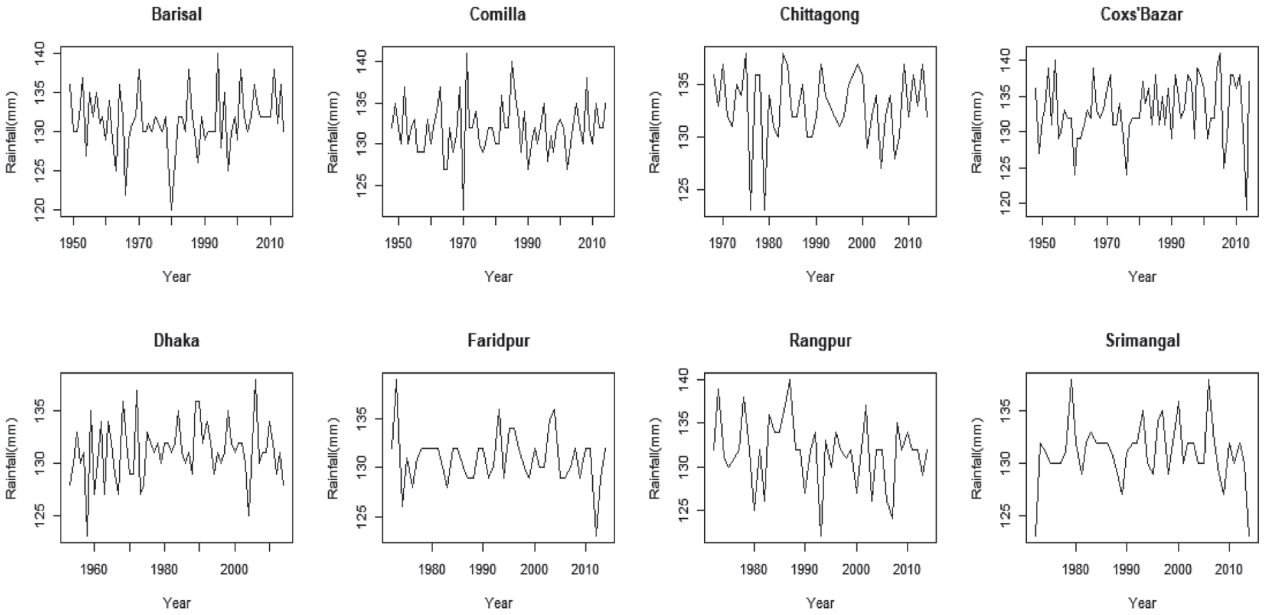


Figure 1. Time series plot of annual maximum rainfall

Total number of extreme events more than 98% quartile was calculated from daily recorded rainfall data, found the extremes cases were more frequent at Barisal and Comilla. The accumulated data in this study exposed the average maximum annual rainfall was 131mm to 135 mm approximately with variation extensively in all the stations. Maximum 146 mm rainfall was also recorded at Comilla and Rangpur on 1972 and 1988 respectively. Descriptive characteristics found almost similar for stations Barisal and Comilla in respect of average rainfall, and number of extreme cases above 98% quantile (Table 2). The lowest maximum rainfall was recorded at Srimangal and lowest variability of annual rainfall observed in Dhaka and Srimangal. The station Srimangal is known as rainier area and Dhaka city almost every year facing flash flooding during rainy seasons due to excessive rainfall.

The quantiles and total extreme events over 98% quantiles provide preliminary idea about distribution pattern of the daily rainfall. It may conclude that the daily rainfall has long tail of extremes cases. Before fitting distribution, the stationarity was determined through ADF test and KPSS test at 5% level of significance and found independent. The Mann-Kendall test concluded as no significant trend exists among the rainfall at 5% level of significance. Finally, annual maximum rainfall the data were fitted according to Lognormal, Weibull and Inverse-Gaussian distribution. The QQ plots, density plots and PP plots were examined for all three distributions and all the plots seems good enough to make conclusion regarding the goodness of fit

Table 2. Rainfall characteristics (mm)

Stations	Mean (Sd)	Max rainfall (Year)	98% quantile	Total rainfall Events > 98% quantile
Barisal	133.34 (3.86)	145 (1996)	138.82	16
Comilla	133.8 (3.12)	146 (1972)	138.5	16
Chittagong	135 (3.78)	143 (2013)	140	6
Cox’s bazar	135.07 (4.27)	146 (2005)	141	15
Dhaka	133.94 (2.82)	143 (1972)	137	13
Faridpur	133.56 (3.64)	144 (1973, 1993)	137	12
Rangpur	131.66 (3.96)	146 (1988)	134	7
Srimangal	132.22 (2.99)	140 (1979)	134	8

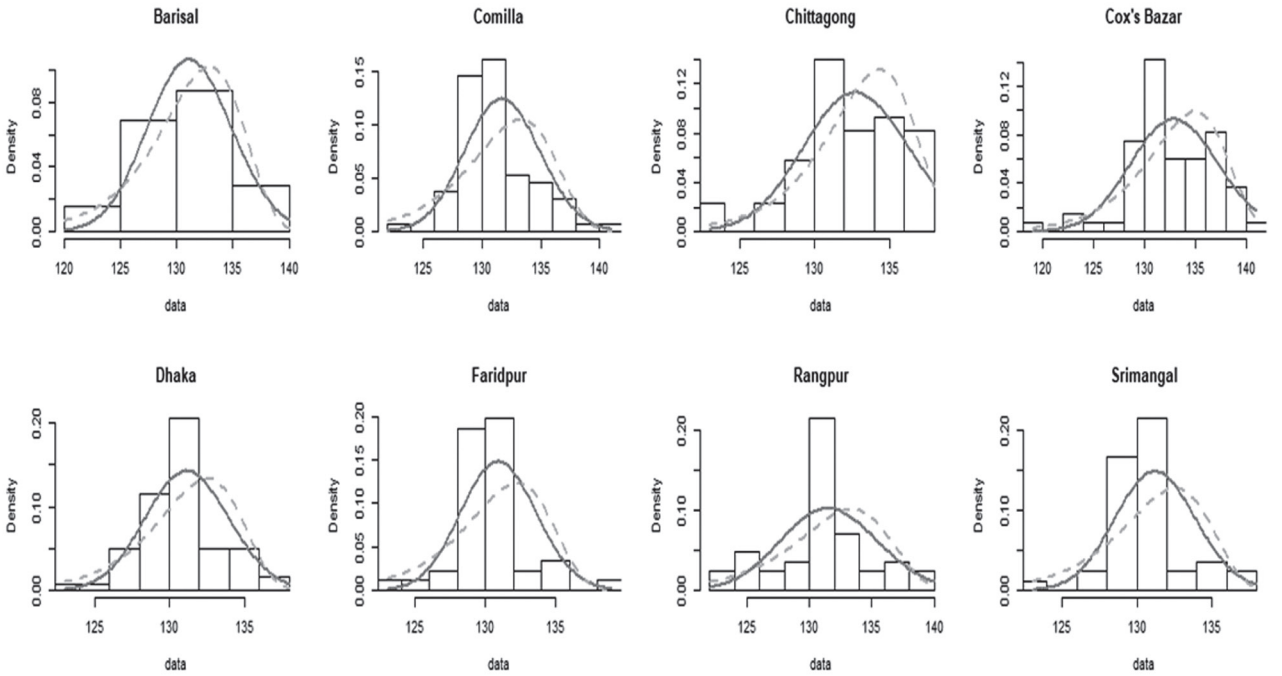


Figure 2. Density plots of Lognormal, Inverse-Gaussian and GEV distribution

Table 3: Goodness of fit statistics

Station	Goodness-of-fit statistics									Goodness-of-fit criteria					
	K-S test			C-M			A-D test			AIC			BIC		
	LN	WB	ING	LN	WB	ING	LN	WB	ING	LN	GEV	ING	LN	GEV	ING
Barisal	0.19	0.25	0.19	0.34	0.51	0.34	1.68	2.47	1.68	355.04	360.61	355.05	359.36	364.93	359.37
Comilla	0.18	0.22	0.18	0.26	0.56	0.25	1.30	2.92	1.30	339.37	356.66	339.37	343.72	361.01	343.72
Chittagong (E)	0.13	0.14	0.13	0.11	0.10	0.11	0.84	0.58	0.84	234.68	227.61	234.68	238.20	231.14	238.20
Cox's Bazar	0.10	0.16	0.10	0.15	0.25	0.15	0.98	1.28	0.98	389.59	386.40	389.59	394.0	390.83	394.0
Dhaka	0.16	0.22	0.16	0.17	0.33	0.17	0.85	1.73	0.85	302.75	310.41	302.75	306.97	314.63	306.98
Faridpur	0.21	0.27	0.21	0.25	0.44	0.25	1.35	2.43	1.36	211.28	222.92	211.29	214.81	226.45	214.80
Rangpur	0.21	0.21	0.21	0.29	0.29	0.29	1.34	1.38	1.34	237.44	239.40	237.44	240.91	242.88	240.91
Srimangal	0.22	0.28	0.22	0.26	0.47	0.26	1.42	2.57	1.42	206.34	216.20	206.34	209.81	219.68	209.81

of the data to the specific distribution. To find the best fitted distribution among three, the graphical comparison of density plots was given in Figure 2. The density plot of Lognormal distribution and Inverse-Gaussian distribution were overlapped whereas the density plots (green curve) of GEV appeared different for all stations.

Table 3 presented the results of Goodness-of-fit statistics obtained K-S, Cramér–von Mises criterion and A-D Goodness-of-fit criteria of probability distributions. The smallest goodness-of-fit among all probability distributions for each stations were considered as best fitted model and majority (9 out of 10) stations followed Lognormal and Inverse-Gaussian among three fitted distributions. Only Chittagong followed GEV of type II that was Weibull class. Moreover, the goodness of fit statistics and goodness of fit criteria obtained from Inverse-Gaussian were very close to Lognormal distribution for all stations.

Table 4. Return level of rainfall

Stations		Return level			
		10 year	20 years	50 year	100 year
Barisal	lognormal	136.07	137.49	139.11	140.20
	Inv-Gaussian	136.07	137.49	139.11	140.21
Comilla	lognormal	135.92	137.12	138.48	139.40
	Inv-Gaussian	135.92	137.12	138.48	139.40
Chittagong	GEV	136.79	137.54	138.31	138.78
Cox's Bazar	lognormal	138.62	140.26	142.12	143.38
	Inv-Gaussian	138.62	140.25	142.12	143.38
Dhaka	lognormal	134.84	135.89	137.08	137.88
	Inv-Gaussian	134.84	135.89	137.08	137.88
Faridpur	lognormal	134.45	135.46	136.60	137.37
	Inv-Gaussian	134.45	135.46	136.61	137.37
Rangpur	lognormal	136.70	138.18	139.86	141
	Inv-Gaussian	136.70	138.18	139.86	141
Srimangal	lognormal	134.73	135.74	136.88	137.65
	Inv-Gaussian	134.73	135.74	136.88	137.64

The return levels were estimated for 10-year, 20-years, 50-year and 100-year return periods were presented in the following Table 4, calculated from best fitted model. The maximum 10 years return levels of is approximately 138.62 mm at Cox's bazar district. Therefore, the average rainfall 138.38 mm is to be exceeded on average once every 10 years with probability 1/10. However, Cox's bazar might experience maximum 142.05 mm average rainfall at 100 years return level. In practical point of view, this research is providing strong assessment of the range of projected return periods from a good fitted model. The return levels of 10-year, 20-year, 50-year and 100-year return periods of rainfall are affording to provide risk of heavy rainfall for the upcoming days.

Conclusion

Rainfall is an important sources of water for both plants and animals on earth to survive. However, excessive of rainwater causes loss of lives and properties in some parts of the world including Bangladesh. Therefore, to reduce these losses, proper planning and precaution against extreme rainfall is necessary. This study has been carried out to characterize the extreme rainfall events in flood prone area of Bangladesh. The maximum rainfall per annum was recorded over 1948-2014 of flood prone area was applied in modelling and found the Lognormal and Inverse-Gaussian distribution were best-fit distribution for all studied stations except Chittagong. The 10 years return level of rainfall are almost similar across all stations but Cox's bazar is the maximum among them and so as for 20 years, 50 years and 100 years return periods. Therefore, Cox's bazar is the most exposed to heavy rainfall and may vulnerable due to flood and cyclone and tidal surges. The high return levels of other stations for all return periods tells the regions may experience excessive rainfall from the near future and vulnerable to flood in upcoming years. Hence, this study provides a clear view of future rainfall of flood prone areas which may be effective to take necessary plan or action for flood protection system. Consequently, risk reduction policy should be taken for the severely flood affected area to save the lives and properties from substantial damage. Therefore, the findings of this study would be helpful to policy makers for taking time demanding steps against the hazards of extreme rainfall.

References

Anderson TW and Darling DA.1954. A Test of Goodness-of-fit. Journal of American Statistical Association, *49*: 765-769

Anderson TW.1962. On the Distribution of the Two-Sample Cramer-von Mises Criterion. Annals of Mathematics and Statistics, *33*: 1148-1159.

Ahmed R. 1989. Probabilistic estimates of rainfall extremes in Bangladesh during the pre-monsoon season. Indian Geographical Journal, *64*: 39-53.

Alam MA, Emura K, Farnham C and Yuan J. 2018. Best-Fit Probability Distributions and Return Periods for Maximum Monthly Rainfall in Bangladesh, Climate, *6*: 9

Brammer H. 1990. Floods in Bangladesh: geographical background to the 1987 and 1988 floods. The Geographical Journal, *156*: 12-22.

Brouwer R, Akter S, Brander L and Haque E. 2007. Socioeconomic vulnerability and adaptation to environmental risk: a case study of climate change and flooding in Bangladesh. Risk Analysis: An International Journal, *27*: 313-326.

Bangladesh Metrological Department (BMD). www.bmd.gov.bd

Chakravarti L and Roy.1967. Handbook of Methods of Applied Statistics. John Wiley and Sons. Pp. 392-394.

Coles S, Bawa J, Trenner L and Dorazio p. 2001. An introduction to statistical modeling of extreme values. Springer, London.

Drápela K and Drápelová I. 2011. Application of Mann-Kendall test and the Sen's slope estimates for trend detection in deposition data from Bílý Kříž (Beskydy Mts., the Czech Republic) 1997-2010. Beskydy, *4*: 133-146.

Gujarati D N. 2003. Basic Econometrics, Fourth Edition, MeGraw-Hill Inc., New York. Pp: 23-26.

Hashizume M, Armstrong B, Wagatsuma Y, Faruque ASG, Hayashi T and Sack DA. 2008. Rotavirus infections and climate variability in Dhaka, Bangladesh: a time-series analysis. Epidemiol Infect, *136*: 1281–1289.

Jentsch A and Beierkuhnlein C. 2008. Research frontiers in climate change: Effects of extreme meteorological events on ecosystems. Comptes rendus -Géoscience, *340*: 621-628.

Khudri MM and Sadia F. 2013. Determination of the Best Fit Probability Distribution for Annual Extreme Precipitation in Bangladesh. European Journal of Scientific Research, *103*: 391-404.

May W. 2004. Simulation of the variability and extremes of daily rainfall during the Indian summer monsoon for present and future times in a global time-slice experiment. Climate Dynamics, *22*: 183-204.

Mirza MMQ and Ahmed QK. 2005. Climate change and water resources in South Asia. Taylor & Francis, London.

McMichael AJ, Woodruff RE and Hales S. 2006. Climate change and human health: present and future risks. Lancet, 367: 859-869.

Smith RL, Tawn JA and Coles SG. 1997. Markov chain models for threshold exceedances. Biometrika, *84*: 249-268.

Sultana N and Roni R. 2015. Rainfall modelling of coastal areas in Bangladesh: extreme-value approach. Journal of Science, Technology & Environment Informatics, *02*: 42-50.

Wilks DS. 1993. Comparison of three-parameter probability distributions for representing annual extreme and partial duration precipitation series. Water Resource Research, 29:3543-3549.