Extreme event characterization for the river basins of Eastern Indian Gangetic Plains

Aradhana Yaduvanshi and Anand Kr Sinha

ABSTRACT

The Eastern Indian Gangetic Plains are characterized by a primarily nature-dependent region blessed with an enormous supply of mineral resources. The region is witnessing a rapid transition in its demographic structure because of rapid industrialization. The region provides a classic example of an area which shows a trend reverse to that observed globally, as far as the frequency of extreme precipitation events is concerned. This paper provides a risk characterization of the entire region, based on the empirical behavior shown by data available so far, in addition to predictions based on theory of extreme values. The long-term behavior prediction is made with an aim to provide policy makers ample time and direction to develop suitable disaster prevention measures. The focus is primarily on extremely high rainfall events, their frequency, trend and estimated long-term behavior. The study corroborates the stability assumption behind the Indian monsoon, and also provides an indication of the expected long-term as well as short-term threats. This study provides a unique application of the extreme value theory to help in developing a threat map for a region whose population is known to be highly impacted by any significant deviations from a normal monsoon.

Key words | extreme value theory, maximum likelihood estimation, prediction, rainfall

Aradhana Yaduvanshi (corresponding author)
Centre of Excellence in Climatology, Department of Physics,
Birla Institute of Technology,
Mesra,
Ranchi,
Jharkhand,
India
E-mail: *aradhanayaduvanshi10@gmail.com*

Anand Kr Sinha
Birla Institute of Technology,
Mesra,
Ranchi,
Jharkhand,
India

INTRODUCTION

The visible impact of anthropogenic forcing of climate change has become substantially evident during the last three decades. Some of the fundamental changes reported by the Intergovernmental Panel on Climate Change (2014) include impacts on natural as well as biological systems over every continent and water body, a robust increasing trend in land and ocean surface temperature, changes in ocean surface salinity and acid content, changes in mass of snow cover, shrinking glacier volume, and changing trends in the global precipitation and water cycle. One of the most perceivable changes has been in the number and intensity of extreme precipitation events. Modeling and forecasting such events remains a challenge, because of the behavioral difference across different regions of the globe. Regions in Australia, New Zealand, the USA, South Pacific islands, Mediterranean basin, South Africa and India have shown an increasing trend for extreme precipitation events (Haylock & Nicholls 2000; Groisman *et al.* 2001; Alpert *et al.* 2002; Fauchereau *et al.* 2003; Roy & Balling 2004; Tabari & Talaee 2011; Pingale *et al.* 2014; Rajendran *et al.* 2013). On the contrary, there are a few regions where a decreasing trend with statistical significance has been observed. Some of these places include some regions in the west of Australia, southeast Asia, the eastern part of New Zealand and also the eastern part of the Gangetic Plain and parts of Uttaranchal in India (Haylock & Nicholls 2000; Manton *et al.* 2001; Salinger & Griffiths 2001; Roy & Balling 2004; Guhathakurta & Rajeevan 2006; Vinnarasi & Dhanya 2015; Sabarish *et al.* 2017).

Risk distribution of extreme events related to climate change is not uniform and the impact is more severe on socially and economically disadvantaged communities. The Indian sub-continent provides a mix of situations complicated by a transition from a nature-based rural society to a rapidly growing urban society, with huge pressures of

doi: 10.2166/nh.2017.211

over-population and industrialization. India receives about 60% to 90% of its total rainfall during the southwest monsoon season, which transcends most part of the months of June to September. The economy and social life has for generations depended on the regularity of the monsoon, both regarding its timing and volume. Extreme precipitation events provide a direct indication of climate extremes, as an increase in their number indicates a serious cause of concern for the overall society. Spatial and temporal characterization based on the Indian monsoon have over the years resulted in a number of interesting studies (Roy & Balling 2004; Goswami *et al.* 2006; Rajeevan *et al.* 2006). Goswami *et al.* (2006), in their effort to analyze the stability of the Indian monsoon, found a significantly increasing trend in the frequency and magnitude of extreme rain events and a decreasing trend for moderate precipitation events. However, it was interestingly observed that the mean rainfall did not show any significant trend, thereby suggesting a stable behavior. Rajeevan *et al.* (2006) also explored the increasing trend of extreme rainfall events in the Indian context and associated it with the increasing trend of sea surface temperatures and surface latent heat flux over the tropical Indian Ocean. Pal & Al-Tabbaa (2010) analyzed the extreme rainfall excess as well as deficiency for the Indian monsoon. In this paper, we focus on the Eastern Indian Gangetic Plains (EIGP) which in particular can be visualized as a traditionally nature-dependent region abundant in mineral resources and that has witnessed a rapidly increasing population pressure. The area remains disadvantaged economically compared to the rest of the region, and hence the impact of extreme events is often severe. It also shows a trend direction for the frequency and magnitude of extreme rainfall events, which is reverse, when compared to the rest of the country, and hence provides a unique situation to explore (Roy & Balling 2004). Seven river basins of the region are explored and a characterization of the extreme event scenario is attempted. Extreme value theory along with other related ideas provide the major foundation on which the modeling and risk forecasting measures are based. The remaining sections of this paper include an outline of the study area and the data used, the methodology, and the results and discussion. The final section concludes the paper with a discussion of the major findings.

STUDY AREA AND DATA USED

Geographically, the EIGP extend from 19 °20′N to 25 °35′N latitude and 79 °30′E to 89 °02′E longitude. The river area is spread over the four states of central and east India, viz. Jharkhand, Orissa, Chattisgarh and some parts of Maharashtra (Figure 1). The study region comprises seven river basins, viz. Subarnarekha, Brahmani, Kasai, Damodar, Son, Tons and Mahanadi. Subarnarekha is the largest east flowing river, with a drainage area of 32,647 km^2. It starts near the Chotanagpur plateau and drains out into the Bay of Bengal covering a total distance of nearly 395 km. Brahmani is the next largest east flowing river, covering around 800 km. Administratively, both Subarnarekha and Brahmani are inter-state rivers. Brahmani has a drainage area of 50,581 km^2. Kasai is situated in the state of Bengal and has the smallest basin in the EIGP. Tons is located in the north-western corner of the EIGP and is 264 km long. Damodar originates from the Palamu hills of Chotanagpur and has a drainage area of around 64,753 km^2. Son is the largest tributary of the Ganges towards the southern region. Its drainage area is 784 km long and it is one of the largest river basins of India with a drainage area of 111,300 km^2. The basin of Mahanadi is the largest and drains an area over 145,040 km^2. The rainfall of the EIGP is dominated by short term fluctuations. The south-west monsoon is the major source of rainfall in this region.

In the present study, $1^\circ \times 1^\circ$ lat-long National Center for Medium Range Weather Forecasting merged satellite-gauge (NMSG) gridded data sets of daily rainfall product were used. Gridded data sets of rainfall have been used in many climatological studies for hydro-climatic forecasting, weather attribution studies (e.g., Sahai *et al.* 2013; Prakash *et al.* 2014). Therefore, in the present study $1^\circ \times 1^\circ$ gridded data sets of daily rainfall from the period 1998 to 2012 were used for the spatio-temporal variability and trend analyses of the extreme events. Large scale variations can be easily identified from the data (Mitra *et al.* 2009). The average number of gauge stations used in the study was 383 on a daily basis, however sometimes up to 400 IMD (India Meteorological Department) gauge stations were used. In the present work 32 grid points were used to cover the whole study area. The data sets were prepared using Tropical Rainfall Measuring Mission multi satellite precipitation

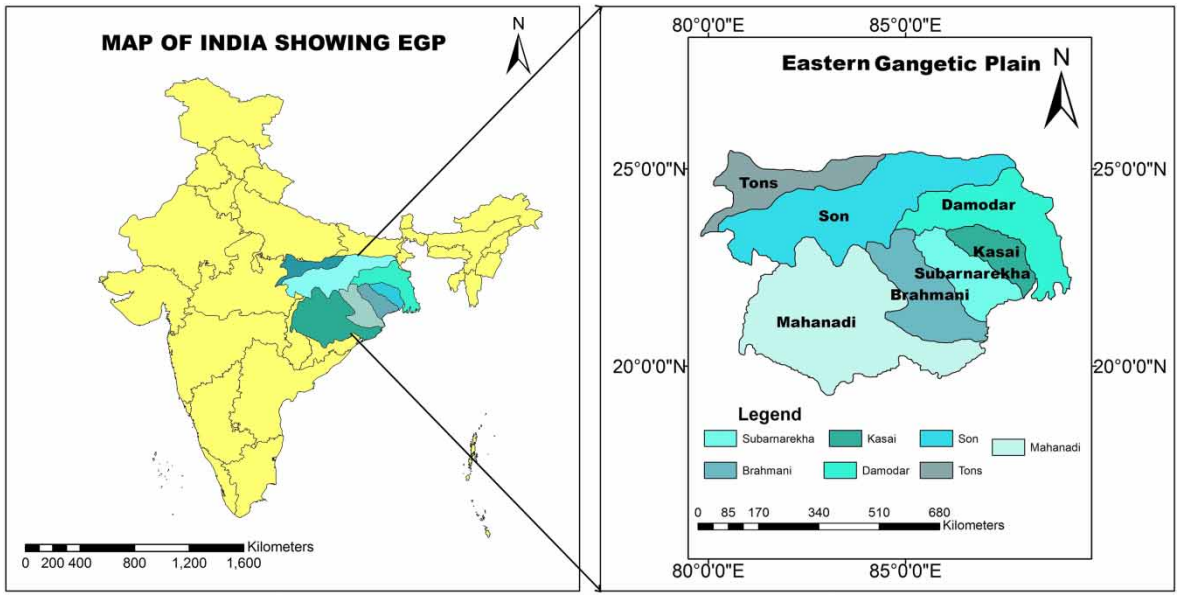

Figure 1 | Geographical location of the EIGP.

analysis (TMPA) at 0.25 degree resolution for every three hours. These data sets have turned out to be effective for a variety of regional applications such as validation of numerical and climate model simulations, extreme temperature studies, and numerous environmental applications such as trend analysis of mean and extreme events of rainfall and temperature (Rajeevan & Bhate 2009). Before the data analysis was executed, quality checks in the form of location of station, missing data, typing errors, etc., were undertaken for the data.

METHODOLOGY

This section describes the various methods and techniques used in the paper.

Basic analysis for deviation from symmetric normal behavior

The normal distribution generally provides a good theoretical approximation to most natural and physical phenomenon. However, closer observation shows that the distributions of most hydrologic variables are heavy tailed in nature (Katz *et al.* 2002). Normally distributed data are typically characterized by a reasonably smooth decrease of probability of occurrence on both sides of the mean. In particular, the distances measured in terms of standard deviations provide important boundaries, with approximately 68% of the data enclosed within one standard deviation from the mean, 95.5% within two standard deviations and around 99.7% of the values lying within three standard deviations (Hogg *et al.* 2013). The three-sigma (3σ) bound therefore accounts for nearly the whole distribution, and very rarely are values expected to lie outside this bound. In this paper the initial analysis consists of evaluating the amount of deviation from this behavior. The traditional threshold approaches, using a high tail value of the distribution or using a high percentile value, are also known to be efficient measures in establishing the deviation from normally distributed behavior. They are also used in the study as they provide a direct cut-off point to identify values in the higher tail.

Central peak and asymmetry in spread of the distribution

Once the deviation from a symmetric normal behavior is established, it becomes imperative to further explore the shape of the empirical distribution. Skewness gives an

indication of the shape of the distribution and its asymmetry with respect to the mean (Hogg *et al.* 2013). Typically, a *right skewed distribution* has a longer right tail, i.e. it is characterized by a dominance of the data towards the left of the mean. In contrast, a *left skewed distribution* has a longer left tail, with most high frequency values lying to the right of the mean. The coefficient of skewness is given as

$$\gamma_1 = \frac{\mu_3}{\mu_2^{3/2}} \tag{1}$$

where μ_3 and μ_2 are respectively the third and second moments about the mean. For a symmetrical distribution $\gamma_1 = 0$ whereas a distribution is generally considered to be highly skewed if the magnitude of γ_1 exceeds 1. The sign of γ_1 tells us whether the distribution is right skewed ($\gamma_1 > 0$) or it is left skewed ($\gamma_1 < 0$). Kurtosis on the other hand gives an idea of the central peak and the effect of tail values on the shape of the distribution. The coefficient of Kurtosis is given as follows

$$\beta_2 = \frac{\mu_4}{\mu_2^2} \tag{2}$$

It provides a better indication of the effect of values farther away from the mean because it involves only even moments and hence the tail values do not offset the contribution of one another. The standard value of β_2 corresponds to that of a normal distribution where $\beta_2 = 3$. Thinner tails lead to lower peaks and then the distribution is said to be *platykurtic* ($\beta_2 < 3$) whereas fatter tails lead to higher peaks where the distribution is called *leptokurtic* ($\beta_2 > 3$).

Mann-Kendall test

The Mann-Kendall test for trend analysis is a useful non-parametric method that does not assume any underlying distribution in the data (Mann 1945; Kendall 1975). The Mann-Kendall test statistic is given as

$$S = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \operatorname{sgn}(S_j - S_i) \tag{3}$$

where sgn represents the signum function. Hence a large positive value indicates that the later observations are larger than the previous ones, thereby indicating a positive trend, and vice versa. We use this test to explore for possible trends in tail event frequencies for the different basins.

Extreme value statistics

The basic statistical theory for extreme values in a distribution has existed for quite some time now (Leadbetter *et al.* 1983). However, the statistical modeling of real life extreme events continues to receive significant attention from researchers (Katz *et al.* 2002). One of the earliest applications of the theory of extremes was presented by Gumbel (1941) where he tried to statistically estimate the return period for floods (Gumbel 1941). The theoretical basis of the theories underlying the statistics of extremes can be traced to the Fisher-Tippet-Gnedenko and Pickands-Balkema-de Hann theorems.

The Fisher-Tippet-Gnedenko approach

The Central Limit theorem forms a basis as far as understanding the asymptotic behavior of the mean of a distribution is concerned. The Fisher-Tippet-Gnedenko theorem (Fisher & Tippett 1928; Gnedenko 1943) similarly forms the basis on which the asymptotic behavior of the extremes in a distribution depends. The theorem suggests that for a sequence of n independent and identically distributed random variables $x_1, x_2, \ldots, x_n$ with an unknown distribution function $F(x)$ and M_i as the maxima of the first $i < n$ observations, the standardized maxima converges to one of three distributions – Gumbel, Frechet and Weibull, which can be expressed in a generalized form by the so-called Generalized Extreme Value (GEV) distribution whose cumulative distribution function is defined as

$$G_{\beta,\xi}(x) = \begin{cases} e^{-\left(1+\frac{\xi x}{\beta}\right)^{1/\xi}} & \text{where } \xi \neq 0 \\ e^{-e^{-x/\beta}} & \text{where } \xi = 0 \end{cases} \tag{4}$$

where β and ξ are the two parameters of the distribution, which are the scaling and shape parameters respectively.

For $\xi < 0$ the limiting distribution corresponds to the Weibull distribution which is thin finite tailed in nature, for $\xi = 0$ it corresponds to the Gumbel distribution which is characterized by exponentially decreasing tails on both sides, and for $\xi > 0$ it corresponds to the Frechet distribution which is thick tailed decreasing by a power law.

The Pickands-Balkema-de Hann approach

The Pickands-Balkema-de Hann theorem (Balkema & de Haan 1974; Pickands 1975) provides another aspect to the analysis of extreme observations. It is often found to be more applicable, as it considers suitable threshold quantities in place of standardized maximal observations. It suggests that for a random variable x with an unknown distribution function $F(x)$, the distribution of x above a threshold τ is defined as

$$F_\tau(y) = P(x - \tau \le y | x > \tau) = \frac{F(y+\tau) - F(\tau)}{1 - F(\tau)} \tag{5}$$

which asymptotically converges to the Generalized Pareto Distribution (GPD) whose cumulative distribution function is given as

$$H_{\beta,\xi}(x) = \begin{cases} 1 - \left(1 + \dfrac{\xi x}{\beta}\right)^{-1/\xi} & \text{where } \xi \neq 0 \\ 1 - e^{-x/\beta} & \text{where } \xi = 0 \end{cases} \tag{6}$$

where the behavior of the scaling parameter β and the shape parameter ξ is similar as in the case of the GEV distribution.

Model fitting and parameter estimation

Corresponding to the two approaches the data can be dealt with in two ways. In the first approach it is divided into equal sized blocks, fitting the GEV distribution to the maxima of each block. In the second, all values above a selected threshold are considered and then the GPD distribution is fitted. A minor demerit in the first approach is that many sufficiently large values get neglected in the process of considering the block maxima. This is not the case in the second approach since all values above the specified threshold are considered. In this paper, both the methods are used. The GEV approach is used only to show the heavy tailed nature of the distribution, while the GPD approach is used to show the results for the remaining analysis. Here we have taken the 90th percentile as the threshold and for GEV we have used the maxima for the June to September period of rainfall in the region. The maximum likelihood estimation (MLE) method is used to fit the extreme value functional models. The performance of the MLE method has been questioned by many researchers but it provides straightforward methods to find standard errors for the estimated parameters and values (Gilleland & Katz 2011). The parameters of the GPD can be estimated by maximizing the following likelihood function

$$L(\beta\xi | m, x_1, x_2, \ldots, x_n) = -m \log_e \beta - \left(1 + \frac{1}{\xi}\right) \times \sum_{i=1}^{n} \log_e \left(1 + \frac{\xi(x_i - \tau)}{\beta}\right) \tag{7}$$

where τ is the threshold and m is the number of values exceeding this threshold value.

RESULTS AND DISCUSSION

Fat tail behavior

Even though the normal distribution is used as an approximation for most naturally occurring phenomenon, several climatological variables show a fat tailed behavior (Katz *et al.* 2002). Table 1 contains the frequency of empirical exceeding of the n-sigma values for n varying from 1 to 6. The values are given for each of the seven basins and the overall average. With the same values of means and standard deviations, the same number of values are generated according to a normal distribution, and their average is given in the last column of Table 1.

In all of the basins the percentage of values exceeding the n-sigma bounds for n varying from 1 to 6 are approximately seen to monotonically decrease from 10–11%, 5%, 2%, 1%, 0.5% to around 0.3%. This is significantly different behavior when compared to the randomly generated normal distribution data, whose corresponding values are 31.81%,

Table 1 | Frequency and percentage of values exceeding n-sigma bounds

Rainfall data (frequencies with percentage within parentheses)									
Outside	**Basin 1**	**Basin 2**	**Basin 3**	**Basin 4**	**Basin 5**	**Basin 6**	**Basin 7**	**Basin All**	**Theoretical data normal distribution**
1 sigma	247 (9.59)	328 (12.73)	322 (12.50)	271 (10.52)	301 (11.68)	259 (10.05)	248 (9.63)	336 (13.04)	819.5 (31.81)
2 sigma	106 (4.11)	156 (6.06)	139 (5.39)	107 (4.15)	142 (5.51)	110 (4.27)	111 (4.31)	147 (5.71)	116.75 (4.53)
3 sigma	54 (2.09)	65 (2.52)	62 (2.41)	48 (1.86)	55 (2.14)	55 (2.14)	48 (1.86)	57 (2.21)	6.9 (0.27)
4 sigma	25 (0.97)	27 (1.05)	31 (1.20)	24 (0.93)	21 (0.82)	30 (1.16)	24 (0.93)	20 (0.78)	0
5 sigma	14 (0.54)	8 (0.31)	15 (0.58)	15 (0.58)	11 (0.43)	14 (0.54)	14 (0.54)	9 (0.35)	0
6 sigma	9 (0.35)	0	4 (0.16)	11 (0.43)	6 (0.23)	12 (0.47)	9 (0.35)	2 (0.08)	0

4.53%, 0.26% and 0% for all sigma higher than 3. Each of the basins show a substantial number of days with rainfall values exceeding the 3-sigma mark and even the 4-sigma mark. This shows that the tails of the empirical distribution in each of the basins show a different behavior in comparison to the theoretical normal distribution model. A closer observation shows that the frequency of values exceeding 6-sigma is relatively higher for the Brahmani (Basin 1), Mahanadi (Basin 4), Subarnarekha (Basin 6) and Tons (Basin 7) river basins, which indicates very high single day rainfall events in these basins. Similar observations are made when we consider a threshold approach, instead of considering the sigma-bounds. Table 2 gives the values of the 90th percentiles for the distributions and also the frequencies exceeding a high threshold of 50 mm daily precipitation. It can be seen that the 90th percentile mostly lies between the first and second sigma values, and still as shown in Table 1, there are values even outside the 6-sigma bound. For the high threshold value of 50 mm daily rainfall, it is seen that the percentage of days exceeding it is not even 3% for any of the basins whereas the contribution of these days to the overall total rainfall observed during the period reaches up to 20%.

Table 3 gives the values of skewness and kurtosis coefficients for the rainfall distributions of all the considered basins.

From Table 3 it can be observed that all the empirical distributions show a high degree of right skewness and all of them are highly leptokurtic in nature. In particular, the distributions for the Brahmani (Basin 1), Mahanadi (Basin 4), Subarnarekha (Basin 6) and Tons (Basin 7) river basins have comparatively much higher values for both γ_1 and β_2. Interestingly, this corresponds exactly to the observations made earlier.

Figure 2 shows the frequency histograms for the daily rainfall in the cases of the different basins. It can be clearly observed that empirically the distributions show characteristics similar to a Frechet distribution.

The distributions for all of the basins are strongly skewed and have significantly persisting right tails. This

Table 2 | Number of days and their contribution for a threshold of $\tau = 50$ mm and position of the 90th percentile between 1-sigma and 2-sigma

	90th percentile	**σ**	**2σ**	**% of days with above 50 mm rain**	**% contribution to total rain**
Basin 1 (Brahmani)	15.60	10.33	20.66	0.82	10.84
Basin 2 (Damodar)	23.01	11.94	23.88	1.44	12.11
Basin 3 (Kasai)	25.61	14.40	28.80	2.49	21.75
Basin 4 (Mahanadi)	14.78	8.46	16.92	0.54	6.31
Basin 5 (Son)	16.16	8.66	17.31	0.39	4.59
Basin 6 (Subarnarekha)	20.38	13.37	16.74	1.98	20.78
Basin 7 (Tons)	14.03	9.39	18.78	0.66	9.95
Basin All (Average EIGP)	17.45	8.44	16.88	0.31	2.80

Table 3 | Coefficients of skewness and kurtosis

Basins	γ_1	β_2
Basin 1	4.55	36.39
Basin 2	2.43	9.54
Basin 3	2.89	13.49
Basin 4	3.55	22.97
Basin 5	3.03	17.04
Basin 6	4.49	38.94
Basin 7	4.92	44.95
Basin All	2.30	10.00

indicates that, even though small rainfall events occur most frequently, very large magnitude rainfall events are also not uncommon.

Before proceeding with trend analysis and extreme value analysis of our empirical distributions, the empirical quantiles are plotted against those of the exponential distribution. The QQ-plots in Figure 3 clearly reveal a behavior which is deviant from the assumed distribution, as is expected. In QQ-plots linearity suggests complete fit and concavity suggests heavy tailed behavior. In the cases of only Basins 2 and 8, the plot is somewhat linear, while in all other cases the concave plots dominated by the extreme values point towards a fat tail behavior.

Trends in k-sigma events

The Mann-Kendall test allows us to explore if there is any temporal trend in the number of daily rainfall events exceeding k-sigma values, where k ranges from 3 to 6. Table 4 shows the values of S given by Equation (3) computed for the four k-sigma event frequencies for the different basins.

It is observed that only for the Tons basin (Basin 7), are all of the four values negative, while most of the other values are positive with a few exceptions. On further significance testing of these values it is found that only the frequency of 3-sigma events for the Damodar basin (Basin 2) show a significant positive trend.

Fitting and verification of extreme functional models

Table 5 gives the values of the estimated parameters and their standard errors, when the GPD is fitted to the eight empirical distributions.

One peculiar observation arising out of the parameter values is that the two basins Damodar (Basin 2) and Kasai (Basin 3) and the overall average are represented by the Weibull distribution whereas all the other basins are represented by the Frechet distribution.

A plot of the distribution of excesses over the threshold and the fitted GPD is given for all the basins in Figure 4. The

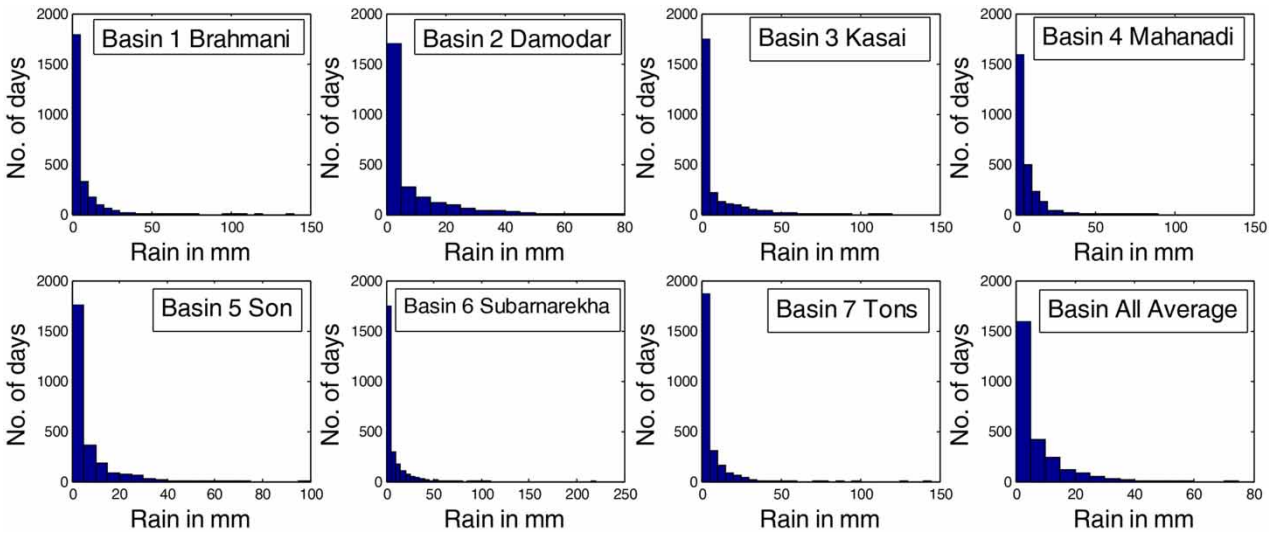

Figure 2 | Histogram of distributions for the basin rainfall values suggesting strongly skewed right tails.

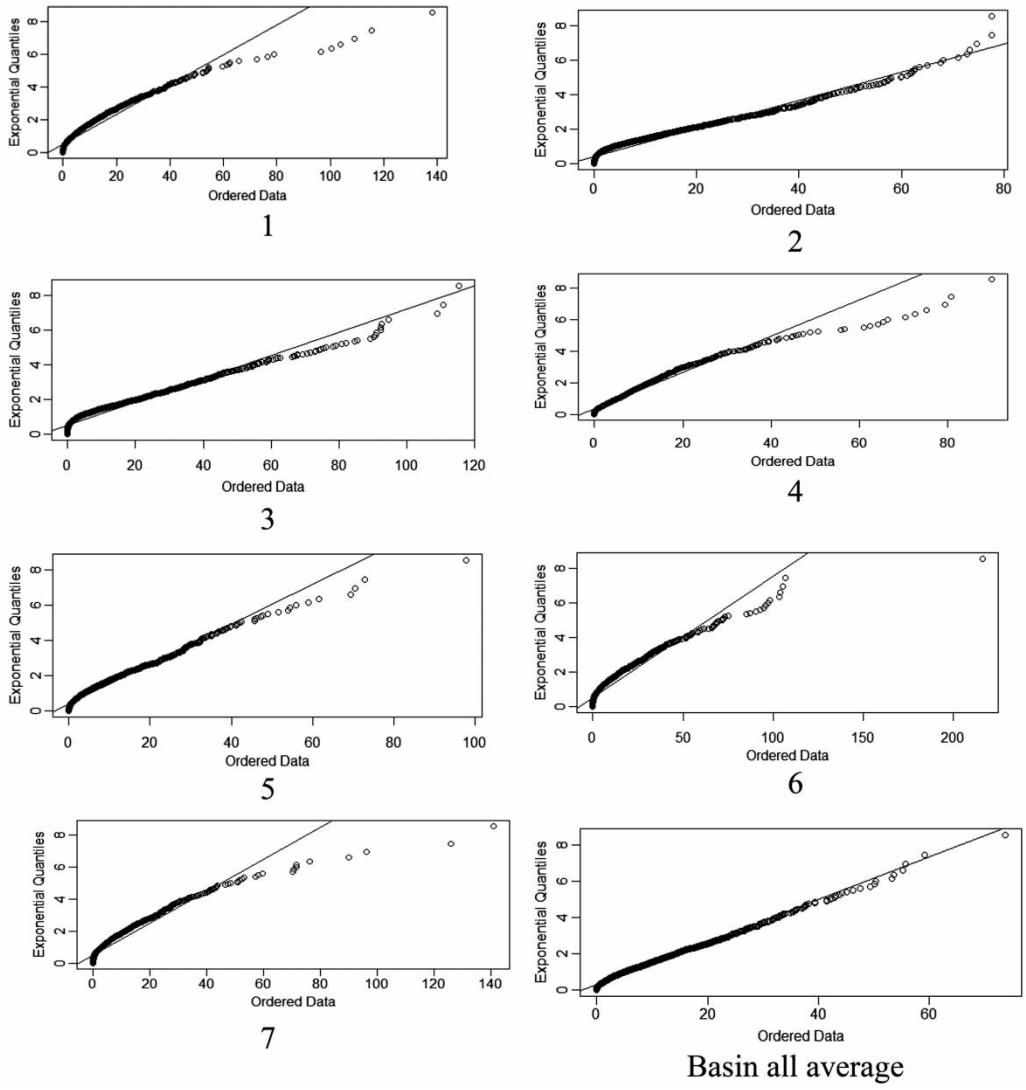

Figure 3 | QQ-plots revealing the presence of fat tails in the empirical distributions.

Table 4 | Mann-Kendall test statistic and the z-statistic values for k-sigma events

	3-sigma	4-sigma	5-sigma	6-sigma
Basin 1	24 (1.25)	14 (0.71)	5 (0.21)	−2 (− 0.05)
Basin 2	53 (2.84)	22 (1.14)	12 (0.60)	0 (− 0.05)
Basin 3	19 (0.98)	1 (0)	17 (0.87)	−14 (− 0.71)
Basin 4	14 (0.71)	5 (0.21)	17 (0.87)	6 (0.27)
Basin 5	10 (0.49)	12 (0.60)	−9 (− 0.43)	7 (0.32)
Basin 6	9 (0.43)	10 (0.49)	10 (0.49)	11 (0.54)
Basin 7	−2 (− 0.05)	−26 (− 1.36)	−22 (− 1.14)	−17 (− 0.87)
Basin All	17 (0.87)	18 (0.93)	15 (0.76)	4 (0.16)

Table 5 | Estimated values and errors in value of scale parameter and shape parameter

Basin	Estimated ξ	Estimated β	Std error ξ	Std error β
1	0.22	9.40	0.05	0.70
2	−0.11	16.49	0.04	0.97
3	−0.04	19.14	0.04	1.09
4	0.23	6.92	0.05	0.49
5	0.007	10.51	0.04	0.67
6	0.16	13.56	0.05	0.90
7	0.20	9.02	0.05	0.69
All	−0.02	9.52	0.04	0.56

QQ-plots also show that the fitted model has a good fit to the empirical distribution.

It can be observed that in all the cases the assumed models fit well into the empirical data, thereby providing us with a distribution that we can now use for predicting risk measures in the next section.

Heterogeneity within homogeneity – the risk map of EIGP

One important risk measure that can be obtained from the models fitted so far is the n-year return level. It can be

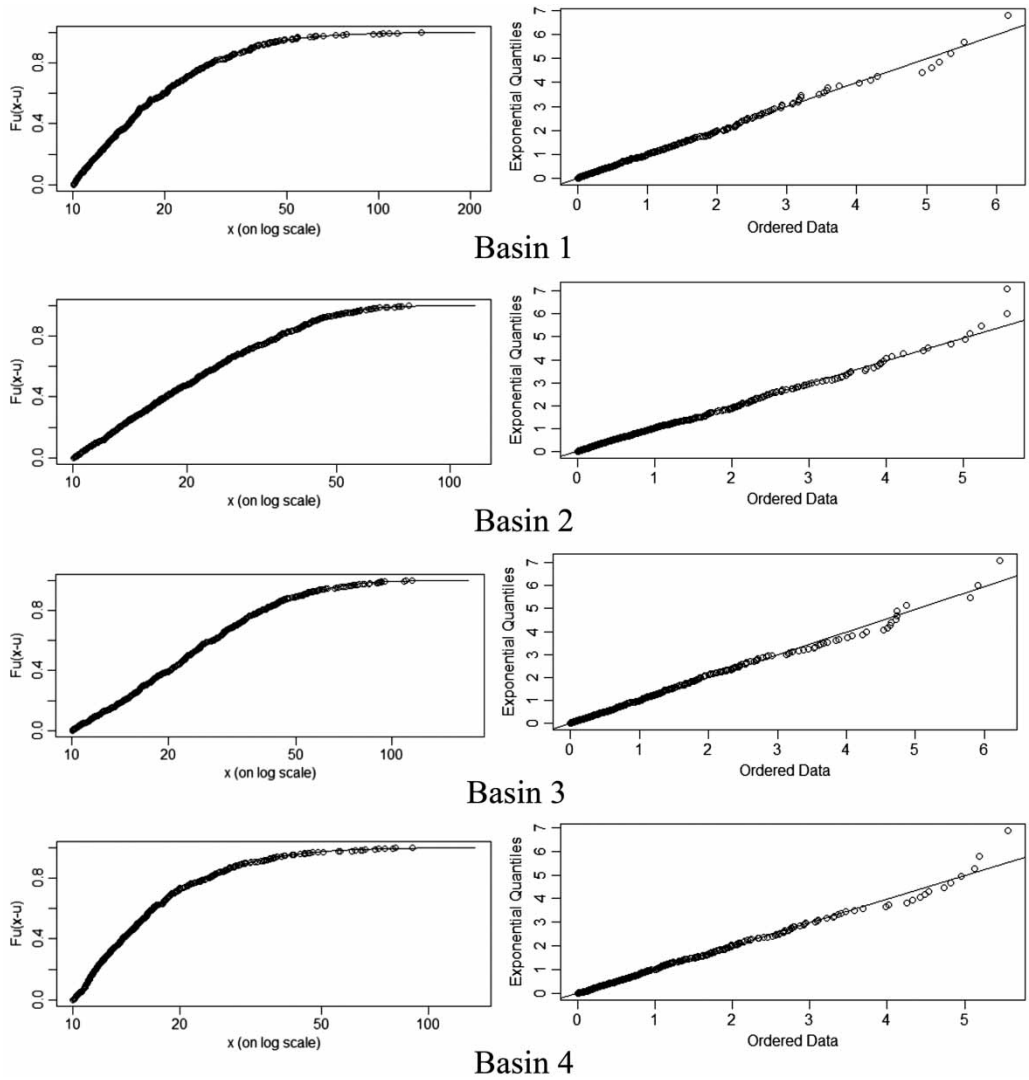

Figure 4 | Verification of fitted GPD to empirical data. (*Continued.*)

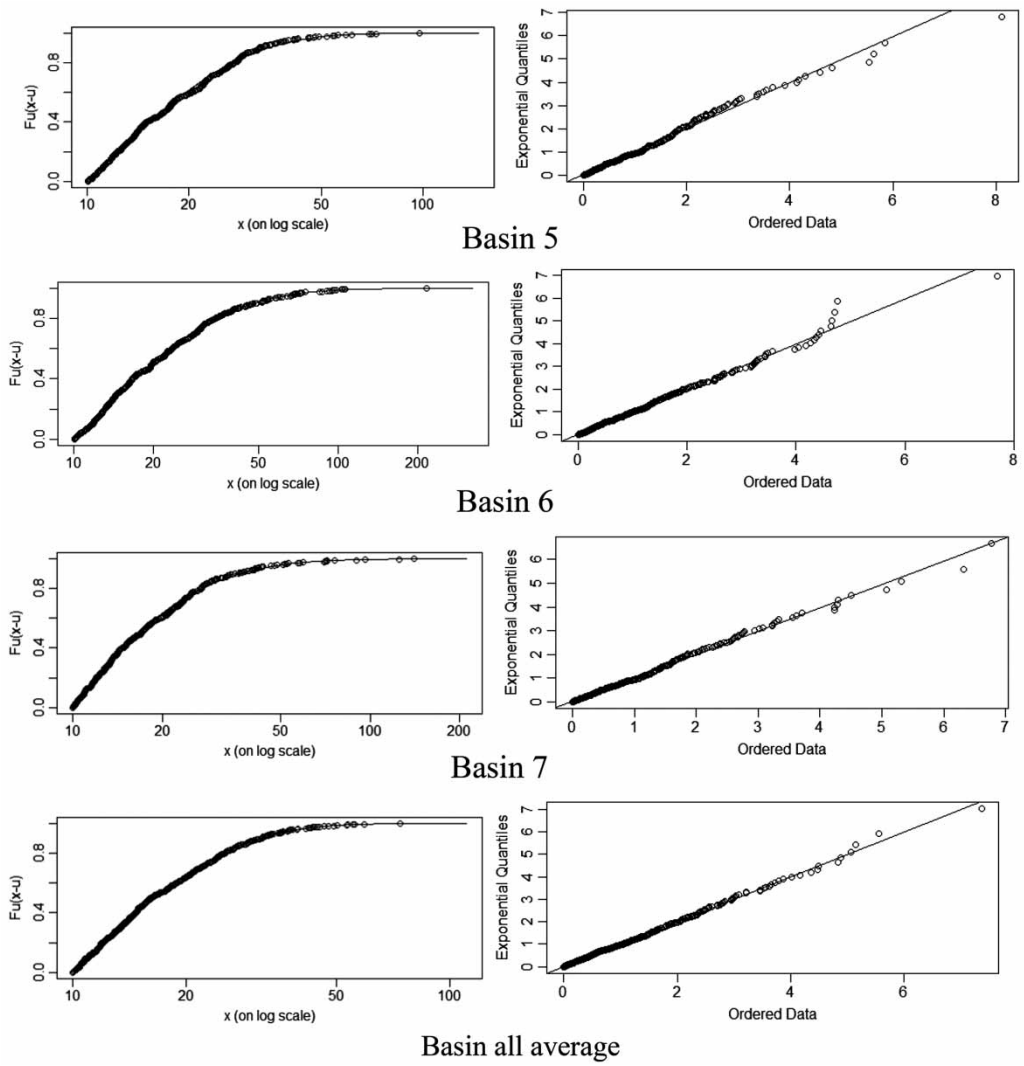

Figure 4 | Continued.

interpreted as the value that is expected to be obtained once in n years and thus provides a measure of the maximum level of risk that we need to be prepared for while designing policies. These values can be obtained by substituting the estimated parameter values into the quantile function for the fitted distribution. Figures 5–7 show a vulnerability map of the EIGP region based on three return periods, viz. extremely short (2 years), medium (20 years) and large (100 years).

When a 2-year return level is considered then the Kasai basin shows a substantially high value, followed by the Subarnarekha and Damodar basins which also have a relatively higher value of the expected return. In the earlier section 'Extreme value statistics', also it was seen that the Damodar and Kasai basins were the only basins to have negative shape parameters in the fitted model. It is difficult from the small range estimate to make any major policy decisions, because the values estimated are generally seen during the peak monsoon very frequently. A medium range prediction shows that the Subarnarekha, Tons and Brahmani river basins have far higher estimates compared to the other basins. This estimation is clearly in line with the elementary observations made while dealing with the fat tail behavior and the trends in the k-sigma events. For the long range

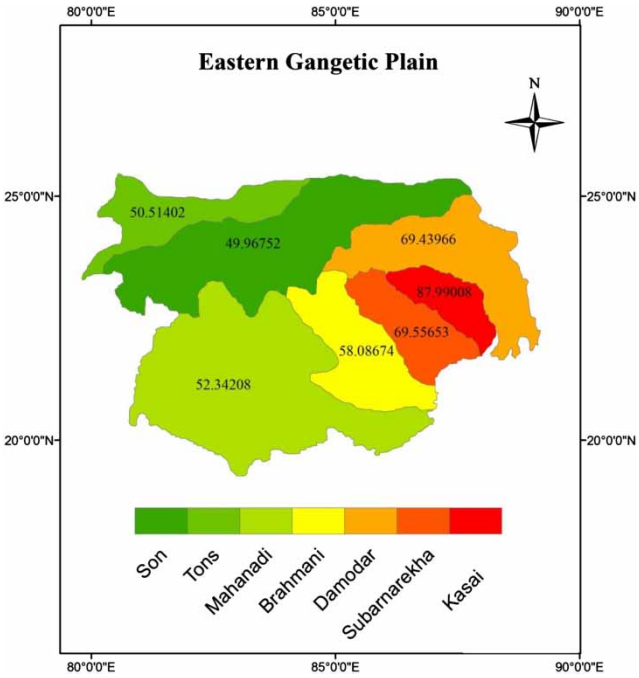

Figure 5 | The small range (2 year) return levels for the EIGP region.

return level, this trend continues with estimates reaching as large as 381 mm single day rainfall for the Tons basin followed by 343 mm for the Subarnarekha basin. These values are clearly expected to cause large scale damage. It can be observed that the estimated levels for the Brahmani, Subarnarekha and Tons river basins show an alarming rise, whereas the others show relatively more stable long term behavior.

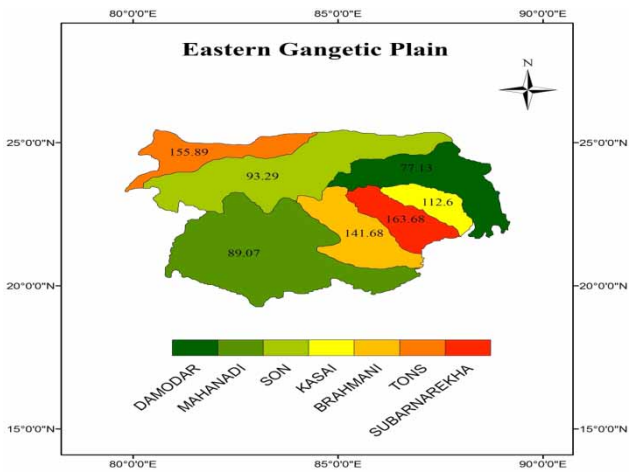

Figure 6 | The medium range (20 year) return levels for the EIGP region.

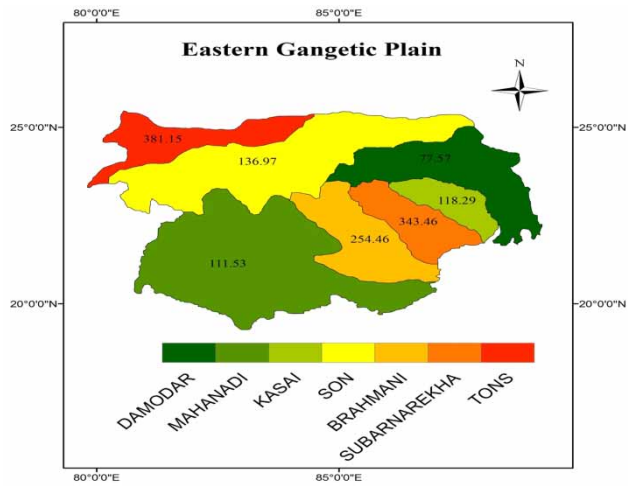

Figure 7 | The long range (100 year) return levels for the EIGP region.

CONCLUSION

The primary purpose of this paper was to present a characterization of the EIGP which have been known to exhibit behavior different from the adjoining areas. Statistical analysis and the extreme value theory are used to develop a threat map for the region which can aid policy makers in planning and designing long-term approaches to deal with situations arising from extreme precipitation events. Methodologically, MLE was used to estimate the parameter values for the fitted models, and then they were used to find the return levels which formed the basis of our risk map. The map presents a characterization which shows that the Brahmani, Subarnarekha and Tons basins are more prone to receive extreme rainfall events in the long run, which are much higher than those for the other basins. In fact, the other basins are expected to show a more stable behavior in the long run. As future work, the models can be improved by including the impact of covariates and possibly exploring other advanced methods to improve the forecasting results.

REFERENCES

Alpert, P., Ben-Gai, T., Baharad, A., Benjamini, Y., Yekutieli, D., Colacino, M., Diodato, L., Ramis, C., Homar, V., Romero, R., Michaelides, S. & Manes, A. 2002 The paradoxical increase of Mediterranean extreme daily rainfall in spite of decrease in total values. *Geophysical Research Letters* **29** (11), 31.1–31.4.

Balkema, A. & de Haan, L. 1974 Residual life time at great age. *Annals of Probability* **2**, 792–804.

Fauchereau, N., Trzaska, S., Roualt, M. & Richard, Y. 2003 Rainfall variability and changes in Southern Africa during the 20th century in the global warming context. *Natural Hazards* **29**, 139–154.

Fisher, R. A. & Tippett, L. H. C. 1928 Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Mathematical Proceedings of the Cambridge Philosophical Society* **24**, 180–190.

Gilleland, E. & Katz, R. W. 2011 New software to analyze how extremes change over time. *Eos Transactions American Geophysical Union* **92** (2), 13–14.

Gnedenko, B. 1943 Sur la distribution limite du terme maximum d'une serie aleatoire. *Annals of Mathematics* 423–453.

Goswami, B. N., Venugopal, V., Sengupta, D., Madhusoodanan, M. S. & Xavier Prince, K. 2006 Increasing trend of extreme rain events over India in a warming environment. *Science* **314** (5804), 1442–1445.

Groisman, P. Y., Knight, R. W. & Karl, T. R. 2001 Heavy precipitation and high streamflow in the contiguous United States: trends in the twentieth century. *Bulletin of the American Meteorological Society* **82**, 219–246.

Guhathakurta, P. & Rajeevan, M. 2006 *Trends in the Rainfall Pattern Over India, NCC Research Report No 2/2006, May 2006*. India Meteorological Department, p. 23.

Gumbel, E. J. 1941 The return period of flood flows. *Annals of Mathematical Statistics* **12**, 163–190.

Haylock, M. & Nicholls, N. 2000 Trends in extreme rainfall indices for an updated high quality data set for Australia, 1910–1998. *International Journal of Climatology* **20** (13), 1533–1541.

Hogg, R. V., McKean, J. & Craig, A. T. 2013 *Introduction to Mathematical Statistics.* Pearson, Upper Saddle River, NJ.

IPCC 2014 *Climate Change 2014 Synthesis Report Summary for Policymakers*. Available online at: https://www.ipcc.ch/pdf/assessment-report/ar5/syr/AR5_SYR_FINAL_SPM.pdf (accessed 7 April 2016).

Katz, R. W., Parlange, M. B. & Naveau, P. 2002 Statistics of extremes in hydrology. *Advances in Water Resources* **25**, 1287–1304.

Kendall, M. G. 1975 *Rank Correlation Methods,* 4th edn. Charles Griffin, London.

Leadbetter, M. R., Lindren, G. & Rootzen, H. 1983 *Extremes and Related Properties of Random Sequences and Processes.* Springer, New York.

Mann, H. B. 1945 Non-parametric tests against trend. *Econometrica* **13**, 163–171.

Manton, M. J., Della-Marta, P. M., Haylock, M. R., Hennessy, K. J., Nicholls, N., Chambers, L. E., Collins, D. A., Daw, G., Finet, A., Gunawan, D., Inape, K., Isobe, H., Kestin, T. S., Lefale, P., Leyu, C. H., Lwin, T., Maitrepierre, L., Ouprasitwong, N., Page, C. M., Pahalad, J., Plummer, N., Salinger, M. J., Suppiah, R., Tran, V. L., Trewin, B., Tibig, I. & Yee, D. 2001 Trends in extreme daily rainfall and temperature in Southeast Asia and the south pacific 1961–1998. *International Journal of Climatology* **21**, 269–284.

Mitra, A. K., Bohra, A. K., Rajeevan, M. N. & Krishnamurti, T. N. 2009 Daily Indian precipitation analysis formed from a merge of rain-gauge data with the TRMM TMPA satellite-derived rainfall estimates. *Journal of Meteorological Society of Japan* **87A**, 265–279.

Pal, I. & Al-Tabbaa, A. 2010 Regional changes in extreme monsoon rainfall deficit and excess in India. *Dynamics of Atmospheres and Oceans* **49** (2), 206–214.

Pickands, J. 1975 Statistical inference using extreme order statistics. *Annals of Statistics* **3**, 119–131.

Pingale, S. M., Khare, D., Jat, M. K. & Adamowski, J. 2014 Spatial and temporal trends of mean and extreme rainfall and temperature for the 33 urban centers of the arid and semi-arid state of Rajasthan, India. *Atmospheric Research* **138**, 73–90.

Prakash, S., Sathiyamoorthy, V., Mahesh, C. & Gairola, R. M. 2014 An evaluation of high-resolution multisatellite rainfall products over the Indian monsoon region. *International Journal of Remote Sensing* **35** (9), 3018–3035.

Rajeevan, M. & Bhate, J. 2009 A high resolution daily gridded rainfall dataset (1971–2005) for mesoscale meteorological studies. *Current Science* **96** (4), 558–562.

Rajeevan, M., Bhate, J., Kale, J. D. & Lal, B. 2006 High resolution daily gridded rainfall data for the Indian region: analysis of break and active monsoon spells. *Current Science* **91**, 296–306.

Rajendran, K., Sajani, S., Jayasankar, C. B. & Kitoh, A. 2013 How dependent is climate change projection of Indian summer monsoon rainfall and extreme events on model resolution? *Current Science* **104**, 1409–1418.

Roy, S. S. & Balling Jr, R. C. 2004 Trends in extreme daily precipitation indices in India. *International Journal of Climatology* **24** (4), 457–466.

Sabarish, R. M., Narasimhan, R., Chandhru, A. R., Suribabu, C. R., Sudharsan, J. & Nithiyanantham, S. 2017 Probability analysis for consecutive-day maximum rainfall for Tiruchirappalli City (south India, Asia). *Applied Water Science* **7** (2), 1033–1042. doi:10.1007/s13201-015-0307-x.

Sahai, A. K., Sharmila, S., Abhilash, S., Chattopadhyay, R., Borah, N., Krishna, R. P. M., Joseph, S., Roxy, M., De, S., Pattnaik, S. & Pillai, P. 2013 Simulation and extended range prediction of monsoon intraseasonal oscillations in NCEP CFS/GFS version 2 framework. *Current Science* **104** (10), 1394–1408.

Salinger, M. J. & Griffiths, G. M. 2001 Trends in New Zealand daily temperature and rainfall extremes. *International Journal of Climatology* **21**, 1437–1452.

Tabari, H. & Talaee, P. H. 2011 Temporal variability of precipitation over Iran: 1966–2005. *Journal of Hydrology* **396**, 313–320.

Vinnarasi, R. & Dhanya, C. T. 2015 Changes in the characteristics of extreme rainfall in Indian monsoon. *Journal of Agroecology and Natural Resource Management* **2** (3), 215–219.

First received 14 August 2016; accepted in revised form 30 January 2017. Available online 3 March 2017