



Generalised extreme value model with cyclic covariate structure for analysis of non-stationary hydrometeorological extremes

R S JAGTAP^{1,*}, V K GEDAM² and MOHAN M KALE²

¹Central Water and Power Research Station, Government of India, Khadakwasla, Pune 411 024, India.

²Department of Statistics & Centre for Advanced Studies in Statistics, Savitribai Phule Pune University, Pune 411 007, India.

*Corresponding author. e-mail: rsj.wadi@gmail.com

MS received 19 September 2017; revised 27 March 2018; accepted 27 March 2018;
published online 1 December 2018

Studies carried out recently on hydrometeorological extremes report the evidence of non-stationarity induced by potential long-term climatic fluctuations and anthropogenic factors. A critical examination of the stationarity assumption has been carried out and a non-stationary generalised extreme value model with cyclic covariate structure for modelling magnitude and variation of data series with some degrees of correlation for real-world applications is proposed. Interestingly, the sinusoidal function with periodicity around 30 yr has been derived as a suitable covariate structure to deal with the ambiguous nature of temporal trends and this could possibly be linked to ‘Sun cycles’. It has adequately explained the cyclic patterns recognised in the annual rainfall which are helpful for realistic estimation of quantiles. Various diagnostic plots and statistics support the usefulness of the proposed covariate structure to tackle potential non-stationarities in the data characterising extreme events in various fields such as hydrology, environment, finance, etc.

Keywords. Covariate; cyclicity; extreme values; trend; quantile.

1. Introduction

The extreme value theory (EVT) in its simplest form applies to independent and identically distributed (IID) random variables while the more advanced techniques allowing some degrees of correlation for coping with trends are necessary (van Nooijen and Kolechkina 2012). Only a few studies have been carried out recently, tackling the practical problems of moving from stationary to non-stationary frequency analyses in real-world applications. A stationary time series has a probability distribution invariant to temporal translations (Brillinger 2001). Stationarity implies that

a time series does not exhibit abrupt and slowly varying changes, periodicities or cyclicity. Several studies have examined the validity of the stationarity assumption for rainfall time series in the Midwest US (Kunkel *et al.* 2007; Pryor *et al.* 2009). Matalas (1997) and Koutsoyiannis (2006) gave an extensive discussion about the stationarity in hydrometeorological records. The inclusion of non-stationarity in flood studies significantly enhances the estimation of the return period; however, its application is questionable when the flood variable does not persistently confirm the significant non-stationarity. The importance of a comprehensive trend analysis during

flood studies has been demonstrated on unregulated daily streamflow series at the Kanawha Fall station on Kanawha River, USA, and at the Baltara station on Kosi River, India (Singh *et al.* 2015). The validity of the stationarity assumption is generally assessed by only testing the record for the presence of slowly varying changes, despite the fact that the presence of abrupt changes could have a large impact on the results of the trend analyses (Villarini *et al.* 2009). The long-term persistence could better explain some of the behaviours exhibited by time series such as statistically significant trends, even though no trends are present (Hurst 1951; Klemes 1974; Koutsoyiannis and Montanari 2007). We have investigated abrupt changes in the first two moments of the distribution of annual rainfall by means of the Pettitt test (Pettitt 1979). The presence of slowly varying changes has been examined using the Mann–Kendall test (Helsel and Hirsch 1993). The presence of ‘deterministic’ trends and change points in a time series in terms of long-term persistence has been checked using the Hurst exponent H by corrected R over S method.

There is a mention about potential non-stationarity induced by the presence of either trend or long-term climatic fluctuations in some hydrometeorological records due to natural climate variability or anthropogenic climate change (Brath *et al.* 1999; Jain and Lall 2001). A comprehensive study of climatic process over central India indicates a decline in the total rainfall during 1950–2015 despite the concurrent rise in the magnitude and frequency of extreme rainfall events (Goswami *et al.* 2006; Roxy *et al.* 2017). A study on changes in extreme rainfall characteristics over India indicates the stationary duration of extreme rainfall spells at most of the locations, and non-stationary associations between the intensity and frequency, and local changes in temperature at a large number of locations (Mondal and Mujumdar 2015). The extent of urbanisation plays a significant role in introducing non-stationarity in the characteristics of Indian Summer Monsoon Rainfall extremes (Singh *et al.* 2016). A study by Agilan and Umamahesh (2015) indicates an increasing trend in the intensity and frequency of daily extreme rainfall during monsoon and non-monsoon months at Hyderabad city that is attributed to global processes whereas the non-stationarity in sub-daily extreme rainfall is mostly associated with local processes. The assumption of randomness, independence and stationarity of data is not fully satisfied in the case of maximum daily rainfall in Mumbai (Jagtap 2014) wherein

the sample size was of the order of 100, resulting in varying estimates of extremes and possibility of non-stationarity in the climatic process. Limited hydrometeorological records pose a challenge for reliable assessment of the presence of non-stationarity. Distinguishing non-stationarity, sample variability and long-term climatic fluctuations becomes difficult as different statistical treatment is needed for coping with various types of variability. It is hypothesised that all natural systems are non-stationary, unequivocally and unconditionally. The issue, therefore, is not whether hydroclimatic systems are stationary or non-stationary but, rather, the question is whether the non-stationarity is substantial enough to require a complex deterministic characterisation of the process, or can a comparatively simple stationary stochastic model accurately represent the process (Lins and Cohn 2001; Cohn and Lins 2005)?

Research challenges include the statistical modelling of complex climatic extreme events which require increased collaboration between the climate scientists and the statisticians (Katz 2010). Researchers proposed abandoning of the stationarity hypothesis and called for innovative methods to provide estimates of hydrologic indicators that would be both reliable and useful for water management (Milly *et al.* 2008). Recent evidences of the impact of persistent modes of regional climate variability, coupled with the intensification of human activities, have led hydrologists to study flood regime without applying the hypothesis of stationarity (Lopez and Frances 2013). The generalised additive models for location, scale and shape are useful in modelling non-stationarity (Rigby and Stasinopoulos 2005). Extreme rainfall projections of Indian summer monsoon using statistical downscaling by Shashikantha and Prashanthi (2017) indicate that the future extremes are heterogeneous across regions and also highlight the impact of climate change on local hydrology for the management of extremes.

We have proposed non-stationary generalised extreme value (GEV) distribution with temporal cyclic covariate structure for modelling ambiguous trends in hydrometeorological quantiles, in the quest for addressing uncertainty issues for assessment of water resources in a better way. The paper is organised as follows. Section 1 is an introductory, which explains the necessity of new models; it also provides a brief literature survey. Section 2 details about the methodology comprising EVT in brief, proposed non-stationary GEV model with cyclic

covariate structure and properties of the estimator. Sections 3 and 4 give frequency analyses and upper tail properties of annual rainfall in the Mula–Mutha–Bhima (MMB) subbasin in India examined using the proposed model as the rainfall in the region is characterised by extreme intermittency typically driven by a very small number of intense events. In section 5, we summarise our findings.

2. Methodology

2.1 Classical extreme value modelling

The extreme value theorem forms the basis for estimation of magnitudes of hydrometeorological

where G is a non-degenerating CDF, then G belongs to one of the three families of distributions Gumbel, Fréchet and Weibull, which are members of extreme value distributions.

The EVT can be traced back to the pioneering work of Fisher and Tippet (1928), wherein the limits for the distributions of maxima of samples of IID random variables were shown to converge asymptotically to one of the three forms of extreme value distributions, called Type I, II and III (Chow *et al.* 1987). The three types of extreme value distributions are unified into the GEV distribution (Von Mises 1954; Jenkinson 1955) with the CDF given by

$$G(x) = \begin{cases} \exp \left\{ - \left[1 + \gamma \left(\frac{x - \alpha}{\beta} \right) \right]^{-1/\gamma} \right\}; & \gamma \neq 0, \beta > 0; 1 + \gamma \left(\frac{x - \alpha}{\beta} \right) > 0, \\ \exp \left\{ -\exp \left[- \left(\frac{x - \alpha}{\beta} \right) \right] \right\}; & \gamma = 0, \beta > 0. \end{cases} \quad (1)$$

The pdf of GEV distribution is

$$f(x|\alpha, \beta, \gamma) = \begin{cases} \frac{1}{\beta} \left[1 + \gamma \left(\frac{x - \alpha}{\beta} \right) \right]^{-((1/\gamma)+1)} \exp \left\{ - \left[1 + \gamma \left(\frac{x - \alpha}{\beta} \right) \right]^{-1/\gamma} \right\}; & \gamma \neq 0, \beta > 0; 1 + \gamma \left(\frac{x - \alpha}{\beta} \right) > 0 \\ \frac{1}{\beta} \exp \left[- \left(\frac{x - \alpha}{\beta} \right) \right] \exp \left\{ -\exp \left[- \left(\frac{x - \alpha}{\beta} \right) \right] \right\}; & \gamma = 0, \beta > 0. \end{cases} \quad (2)$$

sequences for various return periods used for deciding design criteria of engineering structures (Maidment 1993; Coles 2001). The extreme value theorem stated below gives the approximate asymptotic distribution of the sample maximum.

Theorem. Let X_1, X_2, \dots, X_n be a sequence of IID random variables from an unknown cumulative distribution function (CDF) F and let $X_{(n)}$ be a sample maximum. If there exists sequences of normalising constants $\{a_n > 0\}$ and $\{b_n\} \in R$ such that $\lim_{n \rightarrow \infty} P((X_{(n)} - b_n)/a_n \leq x) = G(x)$,

The location parameter α is related to the magnitude of the record, $\beta > 0$ is the scale parameter related to the record variability and γ is the shape parameter, which provides information about the heaviness of the tail of the distribution (El Adlouni *et al.* 2007). Parameters in the GEV distribution can be estimated by the principle of maximum likelihood (ML), where the goal is to identify the most likely set of parameters by maximising the likelihood function. The likelihood function of a sequence of n IID observations $\underline{X} = \{X_t: t = 1, 2, \dots, n\}$ from the GEV distribution can be written as:

$$L(\alpha, \beta, \gamma|\underline{X}) = \begin{cases} \frac{1}{\beta^n} \prod_{t=1}^n \left[1 + \gamma \left(\frac{x_t - \alpha}{\beta} \right) \right]^{-((1/\gamma)+1)} \exp \left\{ - \sum_{t=1}^n \left[1 + \gamma \left(\frac{x_t - \alpha}{\beta} \right) \right]^{-1/\gamma} \right\}; & \gamma \neq 0, \beta > 0; 1 + \gamma \left(\frac{x_t - \alpha}{\beta} \right) > 0, \\ \frac{1}{\beta^n} \exp \left[- \sum_{t=1}^n \left(\frac{x_t - \alpha}{\beta} \right) \right] \exp \left\{ - \sum_{t=1}^n \exp \left[- \left(\frac{x_t - \alpha}{\beta} \right) \right] \right\}. & \gamma = 0, \beta > 0. \end{cases} \quad (3)$$

To obtain ML estimates of the parameters α, β and γ , we differentiate the log likelihood function $\ln L(\alpha, \beta, \gamma | \underline{X})$ with respect to each of the parameters and equate to zero to get a system of simultaneous equations.

Quantile: The return level x_T , the magnitude of a random variable which can be equalled or exceeded at least once in a span of T years can be computed by solving the expression $F(x_T) = P(X \leq x_T) = 1 - (1/T) = p$, where p being the probability of non-exceedance. The return level, x_T using the GEV distribution yields

$$x_T = \begin{cases} \alpha + \frac{\beta}{\gamma} [1 - (1 - \ln p)^\gamma], & \gamma \neq 0, \\ \alpha - \beta \ln(-\ln p), & \gamma = 0. \end{cases} \quad (4)$$

2.2 Non-stationary GEV model with cyclic covariate structure

Stationary frequency analysis implies that the realisation $\underline{X} = \{X_t : t = 1, 2, \dots, n\}$ of the random variable X is IID with stationary CDF $F_X(x, \theta)$, where θ is a vector valued parameter. Whereas non-stationary analysis is carried out under the assumption that the observations are independent but not necessarily identically distributed (I/NID) with a non-stationary CDF $F_X(x, \theta(Z))$, parameters θ are not constant but change as a function of some covariate Z . The ideal representation for the covariate Z could be $Z(t)$, i.e., covariate as a function of time t and parameter denoted by $\theta(Z(t))$. In the absence of any data on covariate, the covariate $Z(t)$ can simply

be taken as time t . In our present work, we have proposed the analysis when parameters θ of an extreme value distribution change as a function of time t , viz., $\theta(t)$. We have thus considered a non-stationary distribution $F_X(x, \theta(t))$ for the extreme value analysis.

The law of variation of the parameters should reflect reasonable predictable physical mechanisms to guarantee that the patterns observed in the period of record are not just an effect of fluctuations of stationary processes whose dynamics evolve over longer time scales (Serinaldi and Kilsby 2015).

We have considered various functional forms for parameters $\theta(t)$, viz., $\theta_0 + \theta_1 t$, $\theta_0 + \theta_1 t + \theta_2 t^2$, $\theta_0 + \theta_1 \ln t$, $\theta_0 + \theta_1 (1/t)$, $\theta_0 + \theta_1 \sin(t)$, $\theta_0 + \theta_1 \sin((t/15)2\pi)$, $\theta_0 + \theta_1 \sin((t/30)2\pi)$, $\theta_0 + \theta_1 \text{abs}[\sin((t/30)2\pi)]$, etc. as given in figure 1 to represent the non-stationary behaviour through the magnitude and variation patterns. Different functional relationships between location and scale parameters of GEV distribution and covariate (which is time) have been developed and used in the extreme value analysis.

Let $\underline{X} = \{X_t : t = 1, 2, \dots, n\}$ be a sequence of n I/NID realisations from a non-stationary GEV distribution $F_X(x, \theta(t))$. Out of the various combinations of proposed non-stationarities for modelling cyclical patterns, a mathematical formulation for one case having the location, $\alpha(t) = \alpha_0 + \alpha_1 \sin((t/30)2\pi)$ and scale, $\beta(t) = \exp^{\beta_0 + \beta_1 t}$; $t = 1, 2, \dots, n$, has been derived and presented below. The ML method has been used to estimate the parameters of the proposed model. The likelihood function in this case is given by

$$L(\alpha_0, \alpha_1, \beta_0, \beta_1, \gamma | \underline{X}) = \prod_{t=1}^n f(X_t | \alpha_0, \alpha_1, \beta_0, \beta_1, \gamma), \quad (5)$$

$$L(\alpha_0, \alpha_1, \beta_0, \beta_1, \gamma | \underline{X}) = \begin{cases} \frac{1}{(\exp^{\beta_0 + \beta_1 t})^n} \prod_{t=1}^n \left[1 + \gamma \left(\frac{x_t - \alpha_0 - \alpha_1 \sin((t/30)2\pi)}{\exp^{\beta_0 + \beta_1 t}} \right) \right]^{-(1/\gamma + 1)} \\ \quad \times \exp \left\{ - \sum_{t=1}^n \left[1 + \gamma \left(\frac{x_t - \alpha_0 - \alpha_1 \sin((t/30)2\pi)}{\exp^{\beta_0 + \beta_1 t}} \right) \right]^{-1/\gamma} \right\}; & \gamma \neq 0 \\ \frac{1}{(\exp^{\beta_0 + \beta_1 t})^n} \exp \left\{ - \sum_{t=1}^n \left(\frac{x_t - \alpha_0 - \alpha_1 \sin((t/30)2\pi)}{\exp^{\beta_0 + \beta_1 t}} \right) \right\} \\ \quad \times \exp \left\{ - \sum_{t=1}^n \exp \left[- \left(\frac{x_t - \alpha_0 - \alpha_1 \sin((t/30)2\pi)}{\exp^{\beta_0 + \beta_1 t}} \right) \right] \right\}; & \gamma = 0. \end{cases} \quad (6)$$

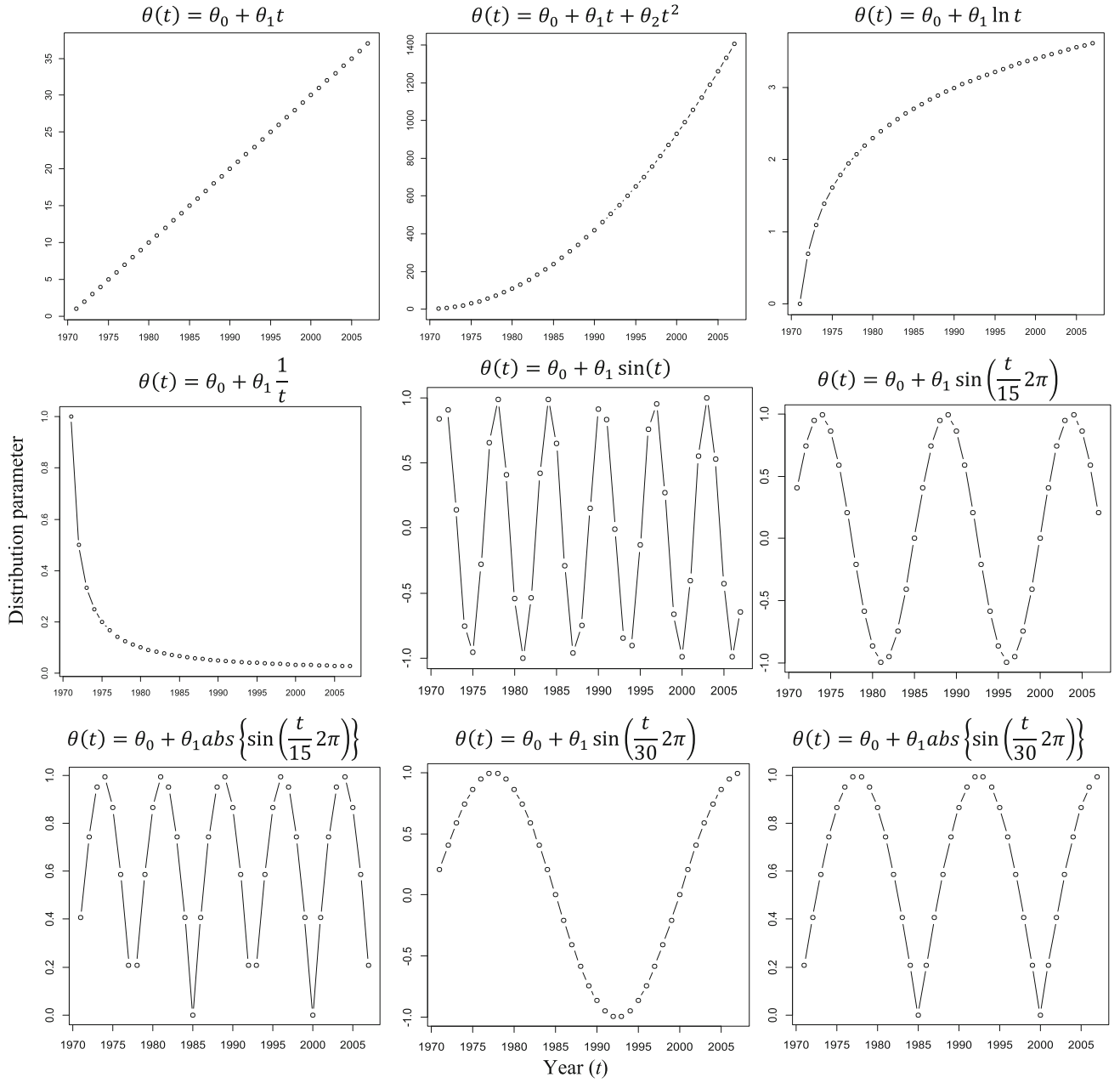


Figure 1. Various forms of parameter $\theta(t)$ of GEV distribution as a function of covariate t .

Take the logarithm of both sides as it is easier to handle, we get

$$\ln L(\alpha_0, \alpha_1, \beta_0, \beta_1, \gamma | \underline{X}) = \begin{cases} -n(\beta_0 + \beta_1 t) - \left(\frac{1}{\gamma} + 1\right) \sum_{t=1}^n \ln \left[1 + \gamma \left(\frac{x_t - \alpha_0 - \alpha_1 \sin((t/30)2\pi)}{\exp^{\beta_0 + \beta_1 t}} \right) \right] \\ - \sum_{t=1}^n \left[1 + \gamma \left(\frac{x_t - \alpha_0 - \alpha_1 \sin((t/30)2\pi)}{\exp^{\beta_0 + \beta_1 t}} \right) \right]^{-1/\gamma}; & \gamma \neq 0, \\ -n(\beta_0 + \beta_1 t) - \sum_{t=1}^n \left(\frac{x_t - \alpha_0 - \alpha_1 \sin((t/30)2\pi)}{\exp^{\beta_0 + \beta_1 t}} \right) \\ - \sum_{t=1}^n \exp \left[- \left(\frac{x_t - \alpha_0 - \alpha_1 \sin((t/30)2\pi)}{\exp^{\beta_0 + \beta_1 t}} \right) \right]. & \gamma = 0. \end{cases} \quad (7)$$

To obtain MLEs of the parameters $\alpha_0, \alpha_1, \beta_0, \beta_1$ and γ , we differentiate the log likelihood function $\ln L(\alpha_0, \alpha_1, \beta_0, \beta_1, \gamma | \underline{X})$ with respect to each of the parameters and equate to zero to get a system of simultaneous equations. For $\gamma \neq 0$, these equations are given below:

$$\frac{\partial \ln L}{\partial \alpha_0} = \left(\frac{1}{\gamma} + 1 \right) \sum_{t=1}^n \frac{\gamma}{\exp^{\beta_0 + \beta_1 t}} \frac{1}{\left[1 + \gamma \left(\frac{x_t - \alpha_0 - \alpha_1 \sin((t/30)2\pi)}{\exp^{\beta_0 + \beta_1 t}} \right) \right]} = 0, \quad (8)$$

$$\begin{aligned} \frac{\partial \ln L}{\partial \alpha_1} &= \left(\frac{1}{\gamma} + 1 \right) \sum_{t=1}^n \frac{\gamma \sin((t/30)2\pi)}{\exp^{\beta_0 + \beta_1 t}} \frac{1}{\left[1 + \gamma \left(\frac{x_t - \alpha_0 - \alpha_1 \sin((t/30)2\pi)}{\exp^{\beta_0 + \beta_1 t}} \right) \right]} \\ &\quad - \sum_{t=1}^n \left(\frac{1}{\gamma} \right) \left[1 + \gamma \left(\frac{x_t - \alpha_0 - \alpha_1 \sin((t/30)2\pi)}{\exp^{\beta_0 + \beta_1 t}} \right) \right]^{-((1/\gamma)+1)} \frac{\gamma \sin((t/30)2\pi)}{\exp^{\beta_0 + \beta_1 t}} = 0, \end{aligned} \quad (9)$$

$$\begin{aligned} \frac{\partial \ln L}{\partial \beta_0} &= -n - \left(\frac{1}{\gamma} + 1 \right) \sum_{t=1}^n \frac{1}{\left[1 + \gamma \left(\frac{x_t - \alpha_0 - \alpha_1 \sin((t/30)2\pi)}{\exp^{\beta_0 + \beta_1 t}} \right) \right]} \\ &\quad \times \left[1 - \gamma \left(\frac{x_t - \alpha_0 - \alpha_1 \sin((t/30)2\pi)}{\exp^{\beta_0 + \beta_1 t}} \right) \right] - \sum_{t=1}^n \left(\frac{1}{\gamma} \right) \left[1 + \gamma \left(\frac{x_t - \alpha_0 - \alpha_1 \sin((t/30)2\pi)}{\exp^{\beta_0 + \beta_1 t}} \right) \right]^{-((1/\gamma)+1)} \\ &\quad \times \left[1 - \gamma \left(\frac{x_t - \alpha_0 - \alpha_1 \sin((t/30)2\pi)}{\exp^{\beta_0 + \beta_1 t}} \right) \right] = 0, \end{aligned} \quad (10)$$

$$\begin{aligned} \frac{\partial \ln L}{\partial \beta_1} &= -nt - \left(\frac{1}{\gamma} + 1 \right) \sum_{t=1}^n \frac{1}{\left[1 + \gamma \left(\frac{x_t - \alpha_0 - \alpha_1 \sin((t/30)2\pi)}{\exp^{\beta_0 + \beta_1 t}} \right) \right]} \\ &\quad \times \left[1 - \gamma t \left(\frac{x_t - \alpha_0 - \alpha_1 \sin((t/30)2\pi)}{\exp^{\beta_0 + \beta_1 t}} \right) \right] + \sum_{t=1}^n \left(\frac{1}{\gamma} \right) \left[1 + \gamma \left(\frac{x_t - \alpha_0 - \alpha_1 \sin((t/30)2\pi)}{\exp^{\beta_0 + \beta_1 t}} \right) \right]^{-((1/\gamma)+1)} \\ &\quad \times \left[1 - \gamma t \left(\frac{x_t - \alpha_0 - \alpha_1 \sin((t/30)2\pi)}{\exp^{\beta_0 + \beta_1 t}} \right) \right] = 0, \end{aligned} \quad (11)$$

$$\begin{aligned} \frac{\partial \ln L}{\partial \gamma} &= \sum_{t=1}^n \left(\frac{1}{\gamma^2} \right) \ln \left[1 + \gamma \left(\frac{x_t - \alpha_0 - \alpha_1 \sin((t/30)2\pi)}{\exp^{\beta_0 + \beta_1 t}} \right) \right] \\ &\quad + \left(\frac{1}{\gamma} + 1 \right) \frac{1}{\left[1 + \gamma \left(\frac{x_t - \alpha_0 - \alpha_1 \sin((t/30)2\pi)}{\exp^{\beta_0 + \beta_1 t}} \right) \right]} \left(\frac{x_t - \alpha_0 - \alpha_1 \sin((t/30)2\pi)}{\exp^{\beta_0 + \beta_1 t}} \right) \\ &\quad - \sum_{t=1}^n \left(\frac{1}{\gamma} \right) \frac{1}{\left[1 + \gamma \left(\frac{x_t - \alpha_0 - \alpha_1 \sin((t/30)2\pi)}{\exp^{\beta_0 + \beta_1 t}} \right) \right]} \left(\frac{x_t - \alpha_0 - \alpha_1 \sin((t/30)2\pi)}{\exp^{\beta_0 + \beta_1 t}} \right) + \left(\frac{1}{\gamma^2} \right) \\ &\quad \times \ln \left[1 + \gamma \left(\frac{x_t - \alpha_0 - \alpha_1 \sin((t/30)2\pi)}{\exp^{\beta_0 + \beta_1 t}} \right) \right] = 0. \end{aligned} \quad (12)$$

On similar lines, ML equations for the case $\gamma = 0$ can also be derived. However, since no closed-form solution exists for the above system of simultaneous equations, the MLEs are obtained by the

direct maximisation of the log likelihood based on numerical algorithms. Optimisation of log likelihood function has been done by developing codes using *extRemes* package in *R* software.

2.3 Properties of model parameter estimators

Under certain regularity conditions, the MLEs will have desirable asymptotic properties viz., consistency, asymptotic normality, invariance and asymptotic efficiency. The log likelihood of the proposed non-stationary GEV model with cyclic covariate structure is essentially a likelihood of the GEV distribution with parameters varying with time. Since the range of the GEV distribution depends on unknown parameters α, β and γ , the regularity conditions for ML estimation are not necessarily satisfied. Nevertheless, the desirable asymptotic properties of efficiency and normality of MLEs hold for the GEV distribution for $\gamma < 0.5$ (Smith 1985).

2.4 Estimation of extreme quantiles

After obtaining the parameter estimates of the non-stationary GEV model described in section 2.2, estimates of quantile for various return periods T have been computed as given below:

$$F\left(x_T|\hat{\alpha}_0, \hat{\alpha}_1, \hat{\beta}_0, \hat{\beta}_1, \hat{\gamma}\right) = 1 - \frac{1}{T} = p, \quad (13)$$

$$\text{for } \gamma \neq 0, \exp\left\{-\left[1 + \gamma\left(\frac{x_T - \alpha_0 - \alpha_1 \sin((t/30)2\pi)}{\exp^{\beta_0 + \beta_1 t}}\right)\right]^{-1/\gamma}\right\} = p, \quad (14)$$

$$\text{for } \gamma = 0, \exp\left\{-\exp\left[-\left(\frac{x_T - \alpha_0 - \alpha_1 \sin((t/30)2\pi)}{\exp^{\beta_0 + \beta_1 t}}\right)\right]\right\} = p. \quad (15)$$

On simplifying, we get

$$x_T = \begin{cases} \alpha_0 + \alpha_1 \sin((t/30)2\pi) + \frac{\exp^{\beta_0 + \beta_1 t} [(-\ln p)^{-\gamma} - 1]}{\gamma}; & \gamma \neq 0, \\ \alpha_0 + \alpha_1 \sin((t/30)2\pi) + \exp^{\beta_0 + \beta_1 t} [-\ln(-\ln p)]; & \gamma = 0, \end{cases} \quad t = 1, 2, \dots \quad (16)$$

The notion of return period becomes ambiguous when we move from stationary to the non-stationary environment; nonetheless, it can still be defined for operational purposes (Serinaldi and Kilsby 2015). It may be noted that the quantile estimates will be changing over time t under the non-stationary GEV model setup.

2.5 Models considered for comparison

We have considered various covariate structures viz., linear, quadratic, reciprocal, logarithmic, sinusoidal (cyclic), as deliberated in section 2.2,

through location, scale and shape parameters $\theta(t)$ under the proposed non-stationary GEV model so as to appropriately explain the non-stationarity behaviour in the environmental variable under study. In the absence of any concrete support for incorporating non-stationarity in the shape parameter, it has not been included in the final shortlisted models. The model performance has been compared with the stationary GEV distribution. The log likelihood, MLEs and quantile estimates can be obtained for the non-stationary GEV models with various covariate structures on similar lines as deliberated in sections 2.2–2.4.

2.6 Criterion for model selection

The validity or suitability of the model has been assessed using various graphical diagnostics like empirical *vs.* modelled density plot, quantile–quantile plot, return level plot and the quantitative assessment has been made using the

negative log likelihood (NLL) and the Akaike information criterion (AICc) corrected for small sample sizes. The likelihood function tends to be a most sensitive criterion of the deviation of the model parameters from the true values (Cramer 1946; Akaike 1974). Alternatively, we seek a minimum of the $NLL = -\ln L$. The $AICc = AIC + (2k(k+1)/(n-k-1))$, where $AIC = 2k - 2\ln L$, where k being the number of independently adjusted parameters within the model, with a penalty for additional parameters in more complicated models. Most parsimonious model with infimum AICc value will be selected.

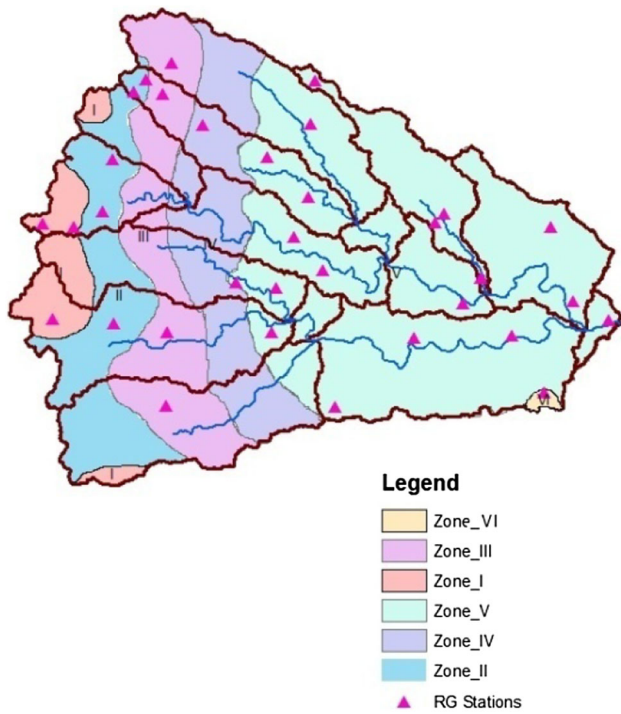


Figure 2. Isohyetal zones, watersheds and rain gauging stations in the MMB basin.

3. Data

The MMB subbasin located in the north-western part of the Krishna basin in India with a catchment area of 6317 km² is selected as study area. The mountainous western part receives very good rainfall, which drastically decreases towards the eastern part which is a rainfall deficit region thus the basin characterises very high temporal and spatial variation. Figure 2 shows a watershed map of the MMB basin stratified into six zones I–VI by drawing isohyetal lines of 2500, 2000, 1500, 1000 and 500 mm. Annual rainfall for the period 1971–2007 for six isohyetal zones and the whole MMB basin have been compiled based on data from 33 representative

rain gauge stations maintained by the Water Resources Department, Maharashtra State (Jagtap *et al.* 2011). Summary of statistical characteristics of annual rainfall in the MMB basin is given in table 1.

3.1 Preliminary data analysis

The non-stationarity in the assessment of extremal behaviour of annual rainfall, which is a vital parameter for water resources planning, has been discussed. The GEV distribution is used as the annual rainfall in different zones of study region is characterised by extreme intermittency typically driven by a very small number of intense events depicting non-symmetric, heavy-tailed probability distribution. A preliminary investigation carried out for annual rainfall in the MMB basin indicates deviation from stationarity assumption, which could be attributed to natural climate variability or anthropogenic climate change and its severe impact on extreme value estimates possibly fuelled due to highly uncertain nature of monsoon-type climate of the region. Various graphical diagnostics of stationary GEV distribution fitted to annual rainfall data in six isohyetal zones confirm the invalidity of the model assumptions and inappropriateness of the stationary model. The two-cycle pattern depicted in figure 3 for representative data series in the MMB basin indicates the presence of non-stationary behaviour. The assumption of randomness, independence and stationarity of annual rainfall series has been quantitatively checked using the Wald–Wolfowitz two-sided run test (Wald and Wolfowitz 1943), since its violation could result in the detection of a statistically significant trend, even if no trend was present (Cox and Stuart 1955). The test results given in table 2 show clear evidence against the stationarity of data series except for zones I and III at 5% significance level.

Table 1. Summary statistics of annual rainfall (mm) in the MMB basin.

	Isohyetal zone						MMB basin
	I	II	III	IV	V	VI	
<i>n</i>	37	37	37	37	37	25	37
Min	2208	1141	890.8	201.2	307.9	70.3	744.7
Max	5941	4010	3128.5	2445.0	1135.7	714.4	2274.4
Average	3866	2465	1874.6	1089.3	671.0	333.7	1367.4
SD	947.3	700.6	596.3	482.5	172.4	162.9	329.9
CV	24.5	28.4	31.8	44.3	25.7	48.8	24.1
Skewness	0.28	0.65	0.56	0.42	0.32	0.58	0.74
Kurtosis	−0.77	0.09	−0.64	0.84	0.82	0.00	1.08

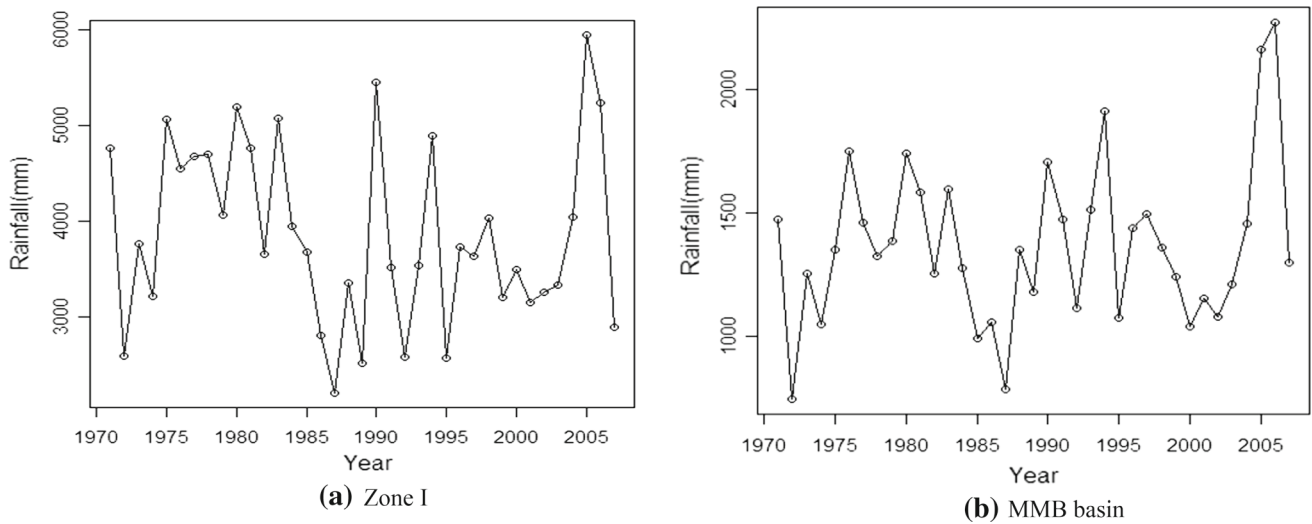


Figure 3. Temporal variation in annual rainfall over the MMB basin.

Table 2. Testing randomness, independence and stationarity of annual rainfall in the MMB basin using Wald–Wolfowitz run test.

	Isohyetal zone (basin)						
	I	II	III	IV	V	VI	MMB
Test statistic	0.906	2.570	0.687	3.258	2.709	2.544	2.087
<i>p</i> -value	0.365	0.010	0.492	0.001	0.007	0.011	0.037

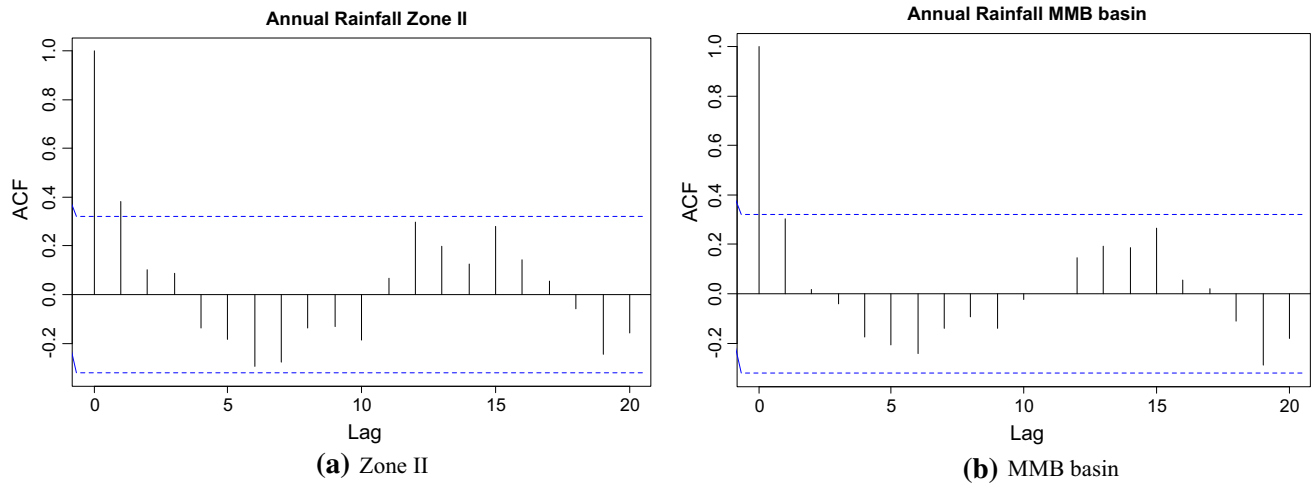


Figure 4. Lag-one autocorrelation of annual rainfall in the MMB basin.

The driving force to undertake the problem further can revolve around the following main issues:

1. There is a huge amount of uncertainty in the estimation of hydrometeorological quantiles. To explain the additional uncertainty caused due to non-stationarity, there is a need for detailed exploratory data analysis.
2. There is a need for developing more consistent non-stationary frequency analysis methods that

can account for the transient nature of a changing climate, lack of data and information characterising hydrometeorological variables.

3.2 Exploratory data analysis

The stationarity assumption is investigated further in this section. In four out of seven data series in each zone of the MMB basin, we found that the lag-one autocorrelation was significantly

Table 3. *Pettitt non-parametric test to detect a shift in the central tendency of annual rainfall in the MMB basin.*

	Isohyetal zone (basin)						MMB
	I	II	III	IV	V	VI	
Pettitt test statistic (K)	148	122	82	150	94	82	76
p -value	0.160	0.359	0.921	0.149	0.722	0.167	1.000
Probable change point (yr)	1984	1989	2002	1979	2003	2003	2003

Table 4. *Mann–Kendall test to detect a monotonic trend in annual rainfall in the MMB basin.*

	Isohyetal zone (basin)						MMB
	I	II	III	IV	V	VI	
Mann–Kendall test statistic (T)	−0.120	0.108	−0.009	0.288	0.074	0.260	0.045
p -value	0.302	0.353	0.948	0.012	0.530	0.072	0.704

Table 5. *Hurst exponent H for annual rainfall in the MMB basin.*

	Isohyetal zone (basin)						MMB
	I	II	III	IV	V	VI	
Corrected R over S Hurst exponent (H)	0.781	0.608	0.595	0.654	0.581	0.561	0.559
Corrected empirical Hurst exponent (H)	0.565	0.239	0.248	0.358	0.282	0.479	0.107

different from zero, although of the order of 0.4, as depicted in figure 4. The presence of any abrupt changes in the central tendency and the variance of the annual rainfall distribution are tested using the Pettitt test. The p -values and probable change points have been given in table 3. Although the test has identified some probable change points for each rainfall series but the changes detected are not significant. The identified behaviour although insignificant could be linked to the change in rainfall regime. The presence of slowly varying changes is examined by means of the Mann–Kendall trend test (Mann 1945; Kendall 1970). Test results in table 4 show that out of seven data series without a statistically significant change point in location, a statistically significant increasing monotonic trend has been detected in zone IV at 5% significance level and in zone VI at 10% significance level. The main difference between abrupt and gradual changes is that with the former the time series remains in the same regime until another abrupt change occurs. Slowly varying changes will tend to persist in the future.

We have computed the Hurst exponent H for examination of long-term persistence, which can result in apparent trends in time series that are actually stationary; the results are summarised

in table 5. Hurst exponents estimated using the corrected empirical method are smaller than or equal to 0.5 suggests that the time series do not exhibit long-term persistence. Whereas Hurst exponents estimated using the corrected R over S method indicate H is consistent with long-term persistence although the estimated H may not be significantly larger than 0.5. It may be mentioned that the estimates of H are very different by different approaches due to the limited sample sizes, so it will not be desirable to provide a conclusive statement.

4. Results using the proposed model

Identification of covariate structure to aptly represent the non-stationary behaviour in the environmental data series and to capture ambiguous trends, periodicities or cyclicities is a formidable task. Various prospective function relationships for distribution parameters and covariate time as described in section 2.2 have been considered for analysis of extremes. Initially, we have incorporated covariate on an individual basis for location, scale and shape parameter in the GEV distribution. Complexities increase in the

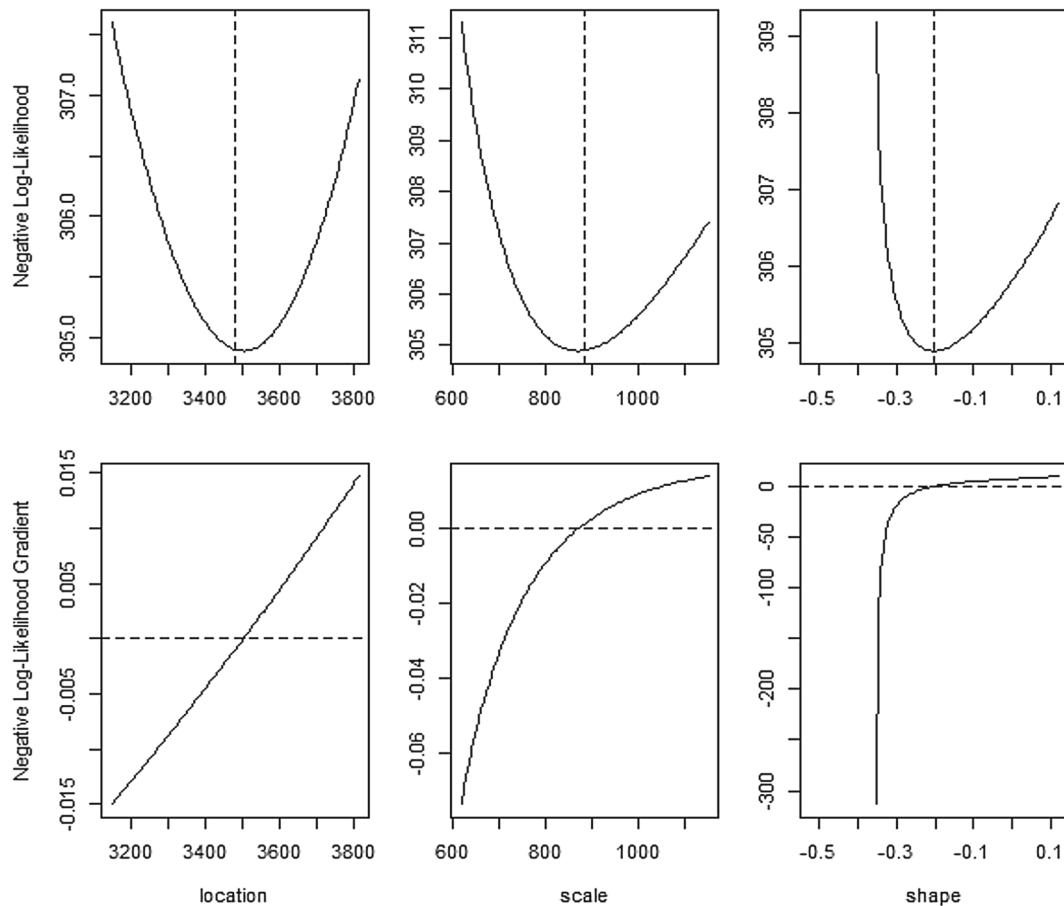


Figure 5. Traces of NLL for GEV parameters whereby other parameters held fixed at MLE values and respective gradient for annual rainfall in the MMB basin zone I.

optimisation of log likelihood as we endeavour covariate simultaneously for various combinations of parameters. The trace plot of NLL for each parameter of GEV distribution, whereby other parameters are held fixed at the MLE values, and the corresponding gradient negative log-likelihoods are useful in diagnosing any fitting problems that might arise in practice. The trace plot in the case of stationary GEV model for annual rainfall in the MMB basin zone I is given in figure 5.

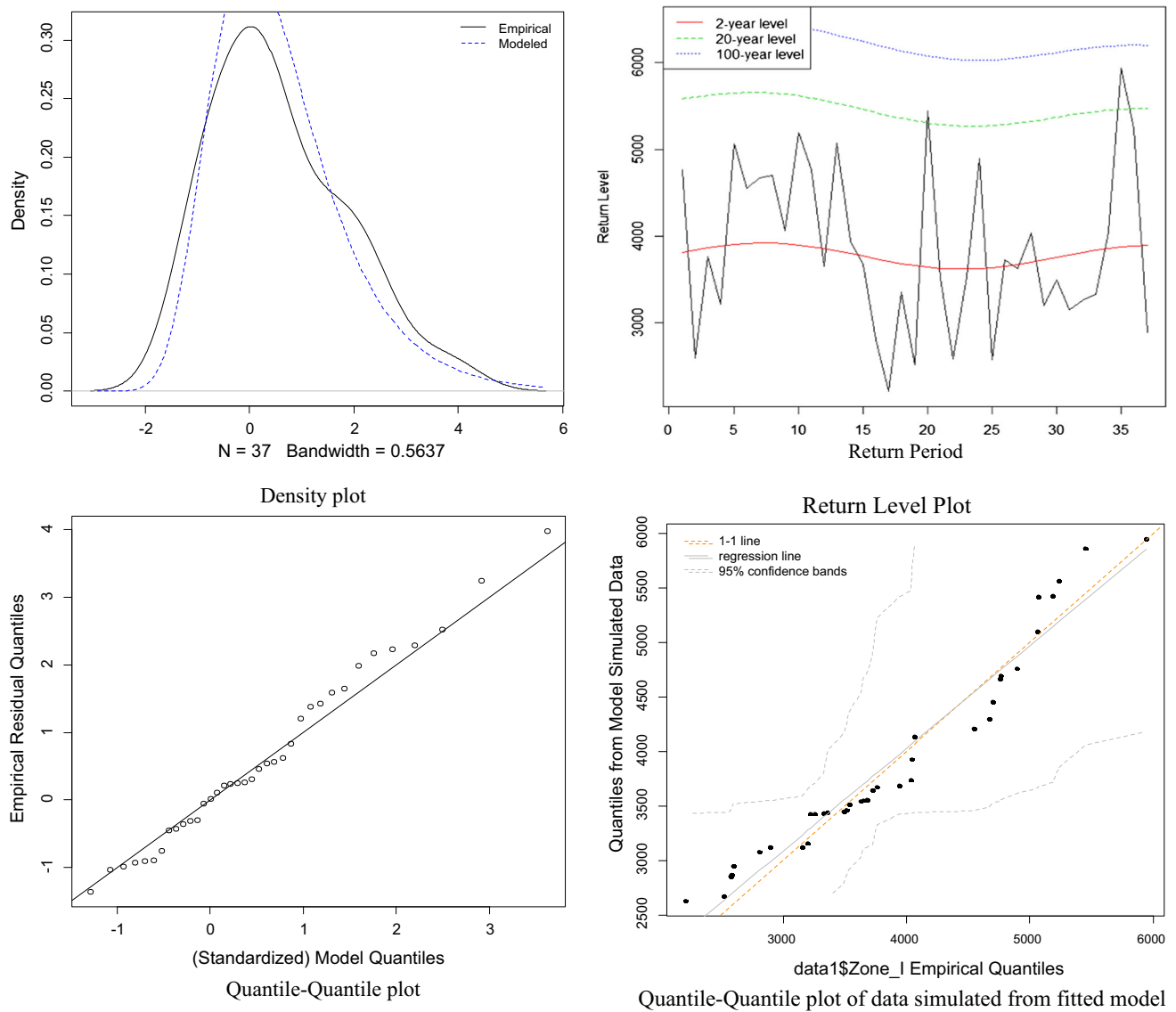
Data analysis has been carried out using the codes developed by *extRemes* package in *R* software (R Core Team 2015). Details about some attempted non-stationary extreme value models, and their corresponding NLL and AICc statistics are given in table 6. It indicates that the magnitudes of annual rainfall when modelled using cyclic trends as non-stationary location parameter show better NLL and AICc values. It is further observed that the variation in the annual rainfall also comprises some cyclic patterns; hence, the non-stationary scale parameter with sinusoidal

function has been incorporated into the analysis. Moreover, we had incorporated non-stationarity in the shape parameter also. The linear, logarithmic and reciprocal function is used to model the changing behaviour in the shape or heaviness in the tail of the distribution. It is seen that one may get better values of NLL and AICc by incorporating non-stationarity in the shape parameter; however, various diagnostic plots show the unstable behaviour of the corresponding log likelihood. Thus, in the absence of any concrete support for incorporating non-stationarity in the shape parameter, it will remain just a mathematical exercise and hence, it has not been included in the final shortlisted models. Interestingly, we get sinusoidal as a suitable covariate function of time representing the cyclic pattern of approximately 30 yr period for the hydrometeorological data series.

It can be inferred from table 6 that there is a considerable gain in using non-stationary EVA over the plain stationary GEV model as could be seen from the enhanced NLL and AICc statistics that

Table 6. Diagnostics of non-stationary GEV models with different covariate structures for annual rainfall in the MMB basin zone I.

Location $\alpha(t)$	Scale $\beta(t)$	Shape $\gamma(t)$	NLL	AICc
α	β	γ	304.91	616.54
$\alpha_0 + \alpha_1 \sin((t/30)2\pi)$	$\beta_0 + \beta_1 \text{abs}\{\sin((t/30)2\pi)\}$	γ	300.42	612.77
$\alpha_0 + \alpha_1 \sin((t/30)2\pi)$	$\beta_0 + \beta_1 \sin((t/30)2\pi)$	γ	301.64	615.21
$\alpha_0 + \alpha_1 \sin((t/30)2\pi)$	β	γ	302.98	615.22
α	$e^{\beta_0 + \beta_1 \text{abs}\{\sin((t/30)2\pi)\}}$	γ	303.28	615.81
$\alpha_0 + \alpha_1 \sin((t/30)2\pi)$	$\beta_0 + \beta_1 t$	γ	302.50	616.93
$\alpha_0 + \alpha_1 \sin((t/30)2\pi)$	$e^{\beta_0 + \beta_1 \ln t}$	γ	302.78	617.50
$\alpha_0 + \alpha_1 \text{abs}\{\sin((t/30)2\pi)\}$	$e^{\beta_0 + \beta_1 \text{abs}\{\sin((t/30)2\pi)\}}$	γ	303.03	617.99
$\alpha_0 + \alpha_1 \text{abs}\{\sin((t/30)2\pi)\}$	β	γ	304.61	618.46
α	$e^{\beta_0 + \beta_1 \sin((t/30)2\pi)}$	γ	304.82	618.90
$\alpha_0 + \alpha_1 t$	β	γ	304.92	619.09

Figure 6. Diagnostic plot for non-stationary GEV model with location $\alpha(t) = \alpha_0 + \alpha_1 \sin((t/30)2\pi)$ and scale $\beta(t) = e^{\beta_0 + \beta_1 t}$ for extremal analysis of annual rainfall in the MMB basin zone I.

ultimately results in the reduction of uncertainty. In other words, one can see that the reliability in the quantile estimates of annual rainfall in the MMB basin has been improved as we employ the non-stationary extreme value modelling. The density plot of empirical and modelled densities, quantile–quantile plot of empirical and modelled quantiles and that of the data simulated from the fitted model, return level plot corresponding to 2, 20 and 100 yr return periods for the fitted non-stationary GEV distribution for annual rainfall in the MMB basin zone I for assessing the adequacy of one combination of the proposed model is shown in figure 6.

Table 7 gives the NLL and AICc statistics for shortlisted GEV non-stationary models along with the stationary model for all annual rainfall data series in the MMB basin. Incorporation of sinusoidal covariates with a frequency of 30 yrs for capturing possible cyclic trends in magnitudes and variation of annual rainfall have given enhanced model fits in respect all the seven data series as compared with the plain GEV stationary model. The extreme value estimates of annual rainfall corresponding to 2, 20 and 100 yr return periods for six isohyetal zones and the MMB basin as a whole using the non-stationary GEV model are shown in figure 7. It indicated that the estimated extreme values of environmental variables corresponding to various return periods are changing over time due to the existence of potential non-stationarity.

The proposed non-stationary GEV model $F_X(x, \theta(t))$ with cyclic covariate structure $\alpha(t) = \alpha_0 + \alpha_1 \text{abs}\{\sin((t/30)2\pi)\}$, $t = 1, 2, \dots, n$, for location and $\beta(t) = \exp^{\beta_0 + \beta_1 \text{abs}\{\sin((t/30)2\pi)\}}$ for scale has performed reasonably well for the majority of the hydrometeorological data series under the study indicating sinusoidal pattern with a period of around 30 yr. Although the approach used in the investigation is basically a data-driven approach, the cyclic patterns pinpointed in the outcome can possibly be linked to the ‘Sun cycles’ which could be an important driving mechanism for the occurrence of rainfall on earth.

For ascertaining the applicability of the proposed non-stationary GEV model with the sinusoidal covariate of the period around 30 yr for location and scale, it has been used for analysis of the extremal behaviour of annual rainfall over the Tapi basin with catchment area of 63,923 km² located in central India. The rainfall data has been accessed from the source India Water Resources Information System website: www.india-wris.nrsc.gov.in.

Table 7. Final non-stationary GEV models with covariate structures shortlisted based on NLL, AICc for extremal analysis of annual rainfall in the MMB basin.

Location $\alpha(t)$	Scale $\beta(t)$	Shape $\gamma(t)$	Criterion	Isohyetal zone						MMB basin
				I	II	III	IV	V	VI	
$\alpha_0 + \alpha_1 \sin((t/30)2\pi)$	$\beta_0 + \beta_1 \text{abs}\{\sin((t/30)2\pi)\}$	γ	NLL	300.42	292.80	286.00	272.69	242.07	159.07	264.21
			AICc	612.77	597.53	583.93	557.32	496.08	331.29	540.36
$\alpha_0 + \alpha_1 \text{abs}\{\sin((t/30)2\pi)\}$	$e^{\beta_0 + \beta_1 \text{abs}\{\sin((t/30)2\pi)\}}$	γ	NLL	303.03	290.52	285.18	276.08	236.21	158.55	262.20
			AICc	617.99	592.97	582.29	564.10	484.36	330.25	536.33
$\alpha_0 + \alpha_1 \sin((t/30)2\pi)$	$\beta_0 + \beta_1 \sin((t/30)2\pi)$	γ	NLL	301.64	292.44	285.97	277.89	239.34	159.20	264.74
			AICc	615.21	596.82	583.87	567.71	490.62	331.55	541.41
$\alpha_0 + \alpha_1 t$	$e^{\beta_0 + \beta_1 \text{abs}\{\sin((t/30)2\pi)\}}$	γ	NLL	303.09	292.51	286.58	274.84	240.49	156.31	264.01
			AICc	618.11	596.96	585.10	561.62	492.91	325.78	539.95
$\alpha_0 + \alpha_1 \sin((t/30)2\pi)$	$e^{\beta_0 + \beta_1 t}$	γ	NLL	303.44	291.51	285.99	280.30	239.91	160.27	264.50
			AICc	618.82	594.95	583.92	572.54	491.75	333.70	540.94
α	β	γ	NLL	304.91	292.84	286.61	280.30	242.47	161.35	265.20
			AICc	616.54	592.41	579.95	567.33	491.66	329.85	537.13

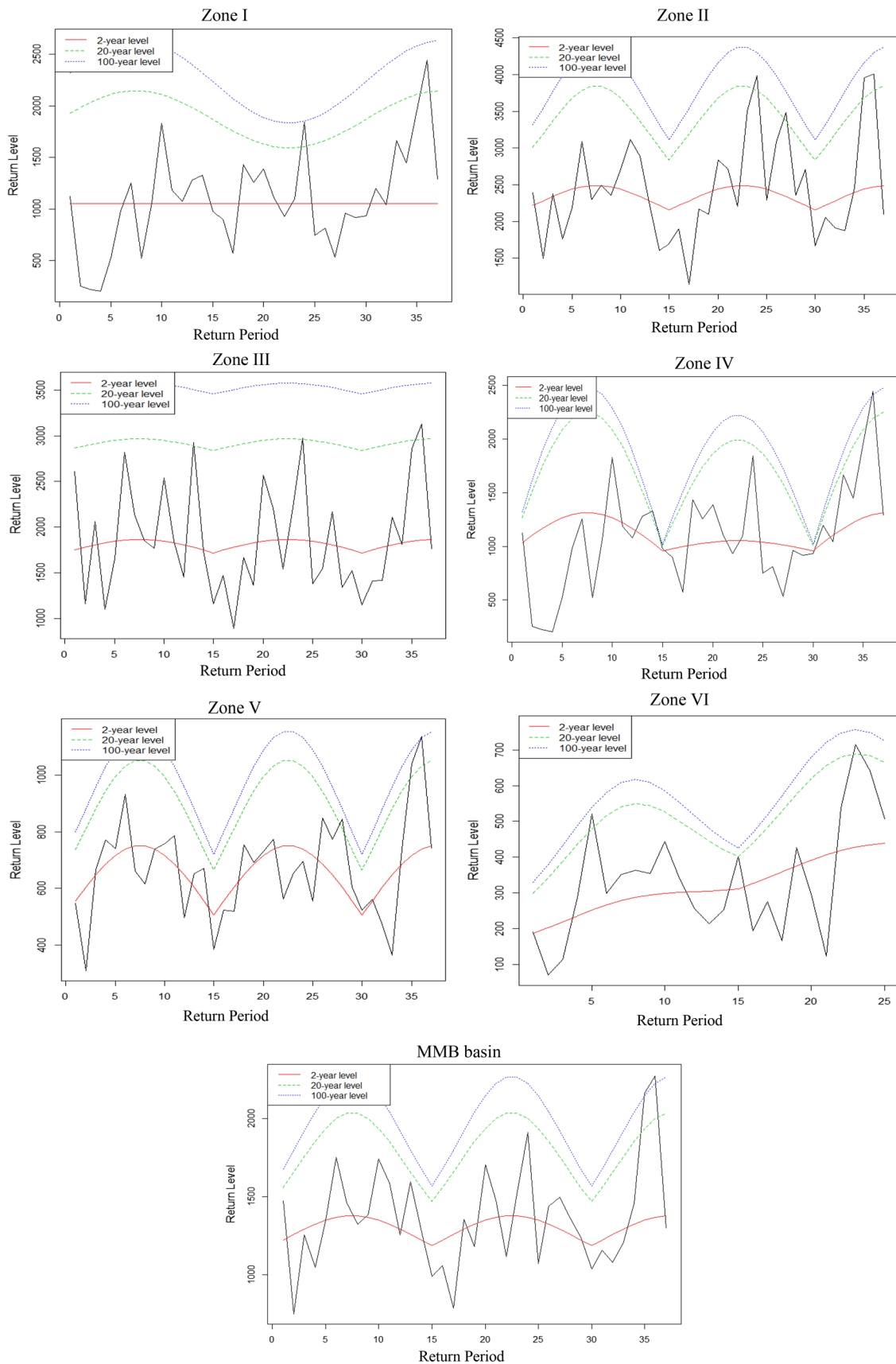


Figure 7. Return level plots using non-stationary GEV model for annual rainfall in the MMB basin.

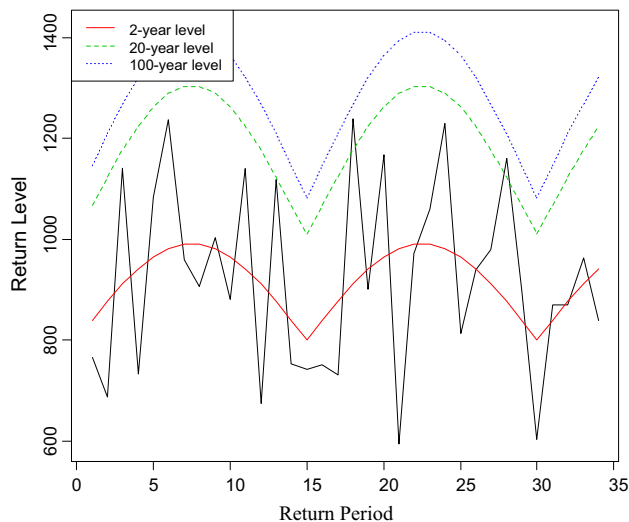


Figure 8. Return level plots using non-stationary GEV model for annual rainfall in the Tapi basin.

The whole Tapi basin has been divided into Upper, Middle and Lower sub-basins and the rainfall data is based on meteorological stations maintained by India Meteorological Department (IMD), Central Water Commission (CWC) and Indian Space Research Organization (ISRO). The NLL and AICc statistical criteria indicate that the proposed non-stationary GEV model with cyclic covariate performs better than the plain stationary GEV model which thus ultimately results in the reduction of uncertainty. The return level plot corresponding to 2, 20 and 100 yr return periods for the fitted non-stationary GEV for annual rainfall in the Tapi basin is given in figure 8.

5. Conclusion

We have made a critical examination of non-stationary frequency analysis essential for real-world applications. Reliable information on the extremal behaviour of some random quantities in our environment such as annual rainfall is a vital input for planning socio-economic activities. Preliminary data analysis, various statistical tests as well as exploratory data analysis indicate the presence of non-stationarity and in-homogeneity that need to be dealt in the extreme value analysis. The proposed non-stationary GEV model with cyclic covariate structure for magnitude and variation of data series has resulted in improved extreme value estimation thereby fall in overall uncertainty. Interestingly, the sinusoidal function with periodicity around 30 yrs has been originated as a suitable covariate structure to deal with ambiguous nature

of temporal trends and has effectively explained the cyclic patterns seen in the environmental random variables as illustrated for annual rainfall series. It could be linked to ‘Sun cycles’ possibly an important driving mechanism for the occurrence of rainfall on earth. Various diagnostic plots and statistics support the usefulness of the proposed covariate structure to tackle potential non-stationarity in the hydrometeorological data series.

One of the limitations may possibly be the complexities in the estimation with increased number of parameters in the probability model. The increasing data length would lead to better model selection and reduction in uncertainty. Reliable information on physical laws driving climatic processes leading to peculiar rainfall patterns would aid in the proper detection of the covariate structure. Parameter estimation methods which address complexities and standard guidelines for practitioners for the inclusion of non-stationary analyses to avoid misuse are indispensable.

Acknowledgment

Authors wish to thank the Director, CWPRS, Pune, for encouragement and continuous support.

References

- Agilan V and Umamahesh N V 2015 Detection and attribution of non-stationarity in intensity and frequency of daily and 4-h extreme rainfall of Hyderabad, India; *J. Hydrol.* **530** 677–697.
- Akaike H 1974 A new look at the statistical model identification; *IEEE Trans. Autom. Control* **19** 716–723.
- Brath A, Castellarin A and Montanari A 1999 Detecting non-stationarity in extreme rainfall data observed in Northern Italy; In: *Proceedings of EGS – Plinius conference on Mediterranean storms*, Maratea, pp. 219–231.
- Brillinger D R 2001 *Time series: Data analysis and theory*; Society for Industrial and Applied Mathematics, Philadelphia, PA.
- Chow V T, Maidment D R and Mays L W 1987 *Applied hydrology* (Civil Engineering Series), McGraw-Hill International editions.
- Cohn T A and Lins H F 2005 Nature’s style: Naturally trendy; *Geophys. Res. Lett.* **32** L23402, <https://doi.org/10.1029/2005GL024476>.
- Coles S 2001 *An introduction to statistical modelling of extreme values*; Springer, London.
- Cox D R and Stuart A 1955 Some quick sign tests for trend in location and dispersion; *Biometrika* **42** 80–95.
- Cramer H 1946 *Mathematical methods of statistics*; Princeton Univ. Press, Princeton, NJ.
- El Adlouni S, Ouarda T B M J, Zhang X, Roy R and Bobee B 2007 Generalized maximum likelihood estimators for the

- non-stationary generalized extreme value model; *Water Resour. Res.* **43** W03410.
- Fisher R and Tippett L 1928 Limiting forms of the frequency distribution of the largest or smallest member of a sample; *Proc. Cambridge Philos. Soc.* **24** 180–190.
- Goswami B N, Venugopal V, Sengupta D, Madhusoodanan M S and Prince K X 2006 Increasing trend of extreme rain events over India in a warming environment; *Science* **314** 1442–1445.
- Helsel D R and Hirsch R M 1993 *Statistical methods in water resources*; Elsevier, <https://www.elsevier.com/books/statistical-methods-in-water-resources/helsel/978-0-444-88528-9>.
- Hurst H E 1951 Long-term storage capacity of reservoirs; *Trans. Trans Am. Soc. Civil Eng.* **116** 770–808.
- Jagtap R S 2014 Effect of record length and recent past events on extreme precipitation analysis; *Curr. Sci.* **106**(5) 698–707.
- Jagtap R S, Roy S K, Mathew F T and Pawar P K 2011 Entropy based assessment of rain gauge network in a river basin; *J. Appl. Hydrol.* **24**(1&2) 74–85.
- Jain S and Lall U 2001 Floods in a changing climate: Does the past represent the future? *Water Resour. Res.* **37** 3193–3205.
- Jenkinson A F 1955 The frequency distribution of the annual maximum (or minimum) values of meteorological elements; *Quart. J. Roy. Meteorol. Soc.* **81** 158–171.
- Katz R W 2010 Statistics of extremes in climate change; *Clim. Change* **100** 71–76, <https://doi.org/10.1007/s10584-010-9834-5>.
- Kendall M G 1970 *Rank correlation methods* (4th edn), Charles Griffin, London.
- Klemes V 1974 The Hurst phenomenon: A puzzle? *Water Resour. Res.* **10**(4) 675–688.
- Koutsoyiannis D 2006 Non-stationarity vs. scaling in hydrology; *J. Hydrol.* **324**(1–4) 239–254.
- Koutsoyiannis D and Montanari A 2007 Statistical analysis of hydroclimatic time series: Uncertainty and insights; *Water Resour. Res.* **43** w05429.
- Kunkel K E, Karl T R and Easterling D R 2007 A Monte Carlo assessment of uncertainties in heavy precipitation frequency variations; *J. Hydrometeorol.* **8** 1152–1160.
- Lins H F and Cohn T A 2001 Stationarity: Wanted dead or alive? *J. Am. Water Res. Assoc. JAWRA* **47**(3).
- Lopez J and Frances F 2013 Non-stationary flood frequency analysis in continental Spanish rivers using climate and reservoir indices as external covariates; *Hydrol. Earth Syst. Sci.* **17** 3189–3203, <https://doi.org/10.5194/hess-17-3189-2013>.
- Maidment D R 1993 *Handbook of hydrology*; McGraw-Hill Professional, New York.
- Mann H B 1945 Non-parametric tests against trend; *Econometrica* **13** 245–259.
- Matalas N C 1997 Stochastic hydrology in the context of climate change; *Clim. Change* **37** 89–101.
- Milly P C D, Betancourt J, Falkenmark M, Hirsch R M, Kundzewicz Z W, Lettenmaier D P and Stouffer R J 2008 Stationarity is dead: Whither water management? *Science* **319** 573–574.
- Mondal A and Mujumdar P P 2015 Modeling non-stationarity in intensity, duration and frequency of extreme rainfall over India; *J. Hydrol.* **521** 217–231.
- Pettitt A N 1979 A non-parametric approach to the change point problem; *J. R. Stat. Soc. C Appl. Stat.* **28** 126–135.
- Pryor S C, Howe J A and Kunkel K E 2009 How spatially coherent and statistically robust are temporal changes in extreme precipitation in the contiguous USA? *Int. J. Climatol.* **29** 31–45.
- R Core Team 2015 R: *A language and environment for statistical computing*; R Foundation for Statistical Computing, Vienna, Austria, <https://www.R-project.org/>.
- Rigby R A and Stasinopoulos D M 2005 Generalized additive models for location, scale and shape; *J. R. Stat. Soc. Ser. C Appl. Stat.* **54** 507–554.
- Roxy M K, Ghosh S, Amey P, Athulya R, Milind M, Raghu M, Pascal T and Rajeevan M 2017 A three-fold rise in widespread extreme rain events over central India; *Nat. Commun.* **8** 708, <https://doi.org/10.1038/s41467-017-00744-9>.
- Serinaldi F and Kilsby C G 2015 Stationarity is undead: Uncertainty dominates the distribution of extremes; *Adv. Water Res.* **77** 17–36.
- Shashikantha K and Prashanthi M 2017 Modeling of extreme rainfall projections of Indian monsoon under climate change; *Indian J. Sci. Res.* **15**(2) 15–19.
- Singh J, Vittal H, Singh T, Karmakar S and Ghosh S 2015 A framework for investigating the diagnostic trend in stationary and nonstationary flood frequency analyses under changing climate; *J. Clim. Change* **1**(1–2) 47–65.
- Singh J, Vittal H, Karmakar S, Ghosh S and Niyogi D 2016 Urbanization causes nonstationarity in Indian summer monsoon rainfall extremes; *Geophys. Res. Lett.* **43**(21).
- Smith R L 1985 Maximum likelihood estimation in a class of non-regular cases; *Biometrika* **72** 67–90.
- Villarini G, Serinaldi F, Smith J A and Krajewski W F 2009 On the stationarity of annual flood peaks in the continental United States during the 20th century; *Water Resour. Res.* **45** W08417, <https://doi.org/10.1029/2008WR007645>.
- van Nooijen R R P and Kolechkina A G 2012 Estimates of extremes in the best of all possible worlds; In: *3rd STAHY international workshop on statistical methods for hydrology and water resources management*, 1–2 October, Tunis, Tunisia.
- Von Mises R 1954 The distribution of the greatest values of n ; *American Math. Soc.: Providence RI* **II** 271–294.
- Wald A and Wolfowitz J 1943 An exact test for randomness in the non-parametric case based on serial correlation; *Ann. Math. Stat.* **14**(4) 378–388.