



Comparison of extreme value theory approaches in temperature frequency analysis (case study: Arak plain in Iran)

Mahnoosh Moghaddasi^{1,2} · Sedigheh Anvari³ · Tayybeh Mohammadi¹

Received: 17 January 2022 / Accepted: 28 May 2022 / Published online: 9 June 2022
© Saudi Society for Geosciences 2022

Abstract

Worldwide research on climate change reveals that a slight temperature change can trigger extreme phenomena such as drought, heavy rain, and storm. This study aims to investigate and compare extreme value theory (EVT) methods in the non-stationary analysis of maximum monthly temperature in Arak plain of Iran over a 116-year period (1901–2016). To this end, the maximum temperature data were collected from the CRU (Climate Research Unit) gridded dataset for stations located in the study area including Sarugh, Shanagh, Arak, Ashtian, and Gavar. Then, using EVT methods including two methods of block maxima (BM) and peaks-over-threshold (POT), time series of maximum temperature were extracted. The MK (the Mann–Kendall), ADF (augmented Dickey–Fuller), KPSS (Kwiatkowski–Phillips–Schmidt–Shin), and PPT (Pettitt’s test) tests were used to examine the trends, stationarity, and homogeneity for the selected stations. Then, distribution parameters were computed based on the DE-MC (differential evolution Markov chain) method. The results showed that all the mentioned stations have a non-stationary trend and the highest difference in the maximum annual temperature was about 0.8 °C over a 2-year period. Moreover, the BM method showed the lowest temperature uncertainty in different return periods, which is recommended as the best method in this research.

Keywords Maximum temperature · Non-stationary · DE-MC · AMS · POT · Arak plain

Introduction

Extreme events are phenomena whose occurrence was not expected by humans, and man-made structures are incapable of handling them and so are often vulnerable because the ecosystems and physical structures of human societies are regulated by normal climatic conditions. Therefore, it can be concluded that climate change events are one of the most serious challenges across the world today (CCSP 2008). Heavy rains (floods), abrupt temperature changes (drought),

and unexpected seasons are examples of these events (Meehl et al. 2000). These events can occur at different time and location scales and may include one or more climatic variables such as temperature, rainfall, flow, and water level, which indicates their complexity.

Based on the special report of the intergovernmental panel on climate change (IPCC) on the management of the dangers of extreme events and natural hazards (Field et al. 2012), global warming can have severe effects on climatic extremes such as changes in their frequency, intensity, and spatial pattern. Also, several studies have been investigated the extreme temperature in Iran that they show increasing trend of temperature significantly (Zamani and Berndtsson 2018; Babaeian et al. 2019; Moghaddasi et al. 2022). The past few decades have witnessed a substantial increase in climatic extremes, including heavy precipitation events and extreme hot days (Alexander et al. 2006; Vose et al. 2005). Several studies have focused on cases of climate extremes like heavy rainfall and severely hot days in different temporal and spatial scales (Jakob 2013; AghaKouchak et al. 2013; Diffenbaugh and Giorgi 2012; Kharin et al. 2007; Easterling et al. 2000). Records show concordance of various extremes

Responsible Editor: Zhihua Zhang

✉ Mahnoosh Moghaddasi
mah_moghaddasi@hotmail.com

¹ Department of Water Science and Engineering, Faculty of Agriculture and Environment, Arak University, Arak, Iran

² Research Institute for Water Science and Engineering, Arak University, Arak, Iran

³ Department of Ecology, Institute of Science and High Technology and Environmental Science, Graduate University of Advanced Technology, Kerman, Iran

such as hot-dry and hot-humid conditions in the last decades of the second millennium. For example, the global mean surface temperature is predicted to increase by about 3 °C in the next century (Ejder et al. 2016). There is an even larger increase in the probability of extreme temperature events (Mearns et al. 1984; Katz and Brown 1992; Perkins et al. 2012) and the frequency and intensity of hydro-climatological extreme events (Mirza 2003; Linnenluecke et al. 2011; Young 2013; Tian et al. 2016). It is expected that the frequency, duration, and intensity of extreme heat events will increase in a future warmer climate (IPCC 2012).

Applying concepts such as return period and return level, under the assumption of a stationary climate, would yield valuable knowledge for planning, decision-making, and estimation of the impacts of unusual weather and climatic events (Rosbjerg and Madsen 1998). But, due to the change in the frequency of extreme values, there is a need for approaches and concepts that are applicable in the non-stationary analysis of climate and hydrological extreme values (Parey et al. 2010; Cooley 2013; Salas and Obeysekera 2013). Frequency analysis is used to investigate the probability of the occurrence of extreme events in a given return period (Gilroy and McCuen 2012). The regional frequency analysis of extreme values is usually determined by two approaches: annual maximum series (AMS), also called block maxima (BM), and peak-over threshold (POT). The generalized extreme value (GEV) and generalized Pareto distribution (GPD) are used as probability distributions for values selected with AMS and POT methods, respectively (Katz et al. 2002; Hawkes et al. 2008; AghaKouchak and Nasrollahi 2010; Li et al. 2015).

There are two approach for estimating of distribution parameters including Bayesian and classical. The choice between a Bayesian and classical approach is then often motivated by the problem at hand. The Bayesian approach allows for a more convenient way of dealing with parameter uncertainty when using estimation results for decision making, for example, in forecasting exercises. In a classical framework (such as MLE method), one often has to rely on bootstrapping techniques. Another advantage of the Bayesian approach arises if the model to be analyzed contains unobserved or latent variables such as, for example, unobserved states describing the stage of the business cycle, missing variables, or the unobserved volatility in a stochastic volatility model. The Bayesian analysis allows for a natural way to conduct inference on the unobserved variables, where again parameter uncertainty is taken into account. The MCMC and DE-MC methods are based on Bayesian approach (Paap 2002).

A great deal of research has been done on extreme temperatures, and some of which are mentioned in the following. Gao and Zheng (2018) utilized quantile regressions to illustrate the annual temperature extremes and to correlate

them with two other weather models of the western Pacific subtropical high (WPSH) and the Arctic Oscillation (AO) in 357 metrological stations in China. In this study, all statistical data such as WPSH (or AO) and prominent positive graphs between warm (or cold) temperature extremes have been analyzed in most of the metrological stations. Finally, the optimal model with the minimum Bayesian information criterion (BIC) was chosen from among 16 nominated constructed GEV distribution models. The periods 1961–1980 and 1991–2010 were computed and estimated based on the 20-year return levels of annual warm (or cold) extremes. The outcomes confirmed the deterministic effect of the trend and distribution changes on the 20-year return levels of variations in annual warm and cold extremes.

Raggad (2018) worked on two numerical studies using the maximum temperature information collected from 15 Saudi Arabian meteorological stations during 1985–2014. The amalgamation of those two approaches resulted in a model by Raggad, which ascertains the effects of diachronic changes on the evaluation of return level and justifies the utilization of the non-stationary generalized extreme value distribution method to modify most of the findings.

Gabda et al. (2018) presented a theory to deduce the distribution of the investigated temperature extremes and their subsequent changes over time by using the findings of climatological data. This research employed the annual maximum perceived temperatures from all 439 sites on a 25-km spatial network in the UK in the 1960–2009 period. It confirmed that using observed information would significantly reduce uncertainty in evaluating historical and future changes in extreme temperature. Aziz et al. (2020) investigated the temporal variability in yearly and seasonal extreme temperatures across Turkey using stationary and non-stationary frequency analysis. The analyses were conducted using generalized extreme value (GEV) and normal and Gumbel distributions for minimum and maximum temperatures during historical (1971–2016) and projection (2051–2100) periods. Magnitudes of non-stationary impacts (30-year return level) show strong spatial and seasonal variability. Notably, higher magnitudes are observed for minimum temperature (up to +10 °C) than maximum temperature (up to +4 °C). Such positive impacts are more significant particularly in eastern Turkey for yearly and seasonal scales. This effect shows greater regional variability in the historical period but, with increased temperature projection, it is more homogenous and larger in the future period for each region.

The main goal of the present study is to compare extreme value theory (EVT) approaches for frequency analysis of temperature in Arak plain, central Iran, as the case study. For this aim, the monthly maximum temperature was first collected from the CRU. In this regard, two commonly used approaches were used including block maxima (BM) or the annual maximum series (AMS) approach and the

peaks-over-threshold (POT) approach to extract time series of maximum temperature. Moreover, some nonparametric trend and stationarity tests were employed for the 1901–2016 period. By fitting the GEV model, the distribution parameters were estimated using the DE-MC method. It is worth noting that the most common estimation methods utilized in previous studies on extreme values were based on either the method of moments or the maximum likelihood. Therefore, the novelty of this research includes (1) temperature frequency analysis based on differential evolution Markov chain (DE-MC) approach and (2) application of two methods to derive time series of extreme maximum temperature. In addition, previous studies have not investigated the non-stationary behavior of extreme temperatures in the central Iranian plateau.

Methodology

Case study and dataset

Markazi province, located in the central part of Iran, is known as the industrial capital of this country. Its climate is warm and dry, with average annual rainfall and temperature of 261 mm and 14.6 °C, respectively. In this province, the most important and the largest plain is the Arak plain, known as the Meyghan desert, wherein enormous industrial and agricultural activities are based. The extent of this basin is 5460 km², and 3100 km² of which is flat and the rest is mountain-outs. Meyghan wetland is one of the 10 major wetlands in Iran, which is located in the middle of this area

(Fig. 1). The research temperature data were collected from the nearby synoptic stations, including Arak, Gavar, Sarugh, Ashtian, and Shanagh. The research datasets included Climate Research Unit (CRU) and station-based observations. The CRU gridded dataset produces time series of monthly climate variables from 1901 to 2016 with a 0.5-degree spatial resolution (New et al. 1999; Mitchell and Jones 2005). These datasets are generated from monthly ground-based climate variables over land and are interpolated through inverse distance weighting (IDW). This research work used the monthly maximum temperature from this dataset (<https://data.ceda.ac.uk>).

Methodology

An overview of the methodology of the current research is as follows (Fig. 2):

Extraction methods of extreme series

In EVT, there are two fundamental approaches, both widely used: the block maxima (BM) or annual maximum series (AMS) (a specific case of BM for yearly blocks) method and the peaks-over-threshold (POT) method (Ferreira and Haan 2015). The BM approach consists of dividing the observation period into nonover-lapping periods of equal size and restricts attention to the maximum observation in each period Gumbel (1958). In POT approach, one selects those of the initial observations that exceed a certain high threshold. The probability distribution of those selected

Fig. 1 The case study

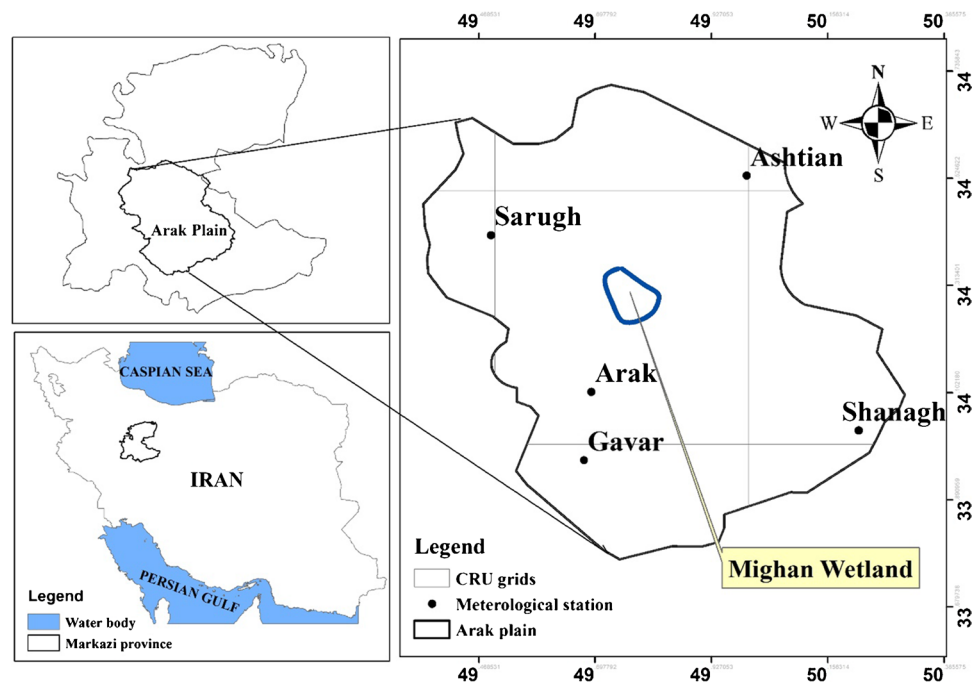
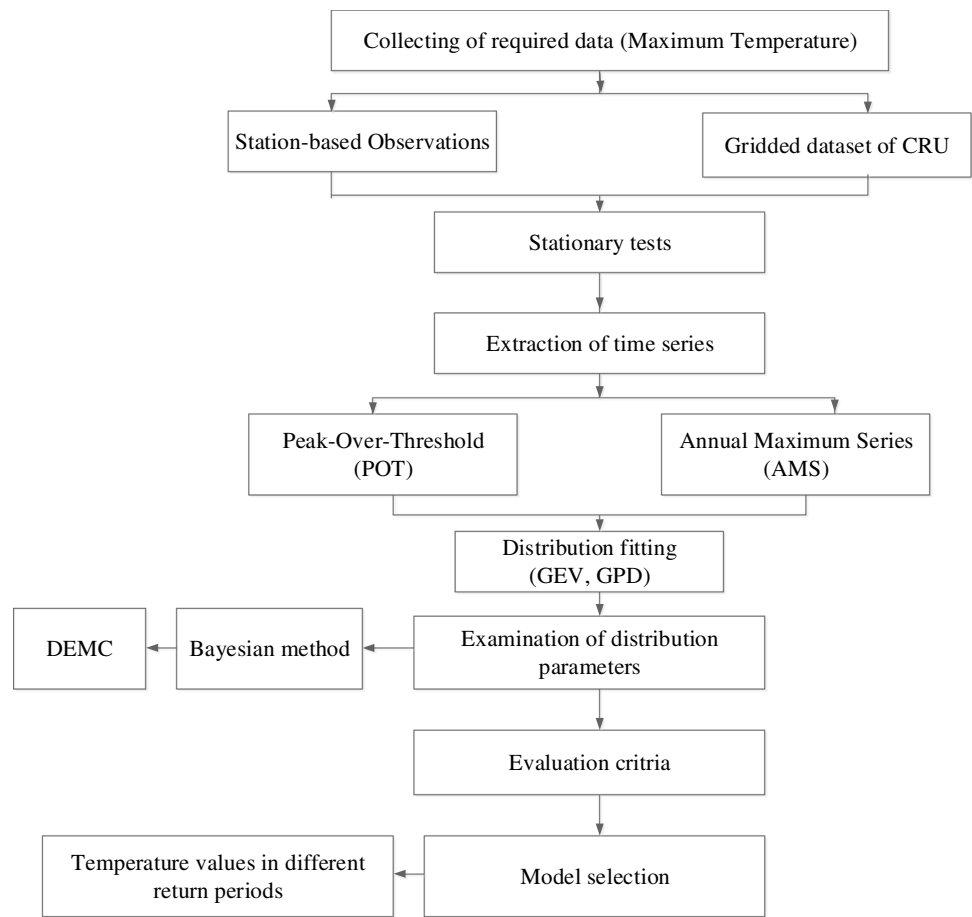


Fig. 2 The research flowchart



observations is approximately a generalized Pareto distribution (Pickands 1975).

Stationary analysis tests

Augmented Dickey-Fuller (ADF) test The augmented Dickey-Fuller (ADF) test can assess stationary. Here, H0 hypothesizes that the time series is stationary, while H1 signifies a non-stationary time series (Banerjee et al. 1993). ADF test was applied at each station to assess stationary in time series at the 95% level of confidence using Eq. (1).

$$\Delta x_n = \alpha + \beta t + \gamma x_{t-1} + \sum_{p=1}^p (\delta_t \Delta x_{t-p}) + \epsilon_t \quad (1)$$

where Δx is the differenced series at a lag of n years, α is the drift, β is the coefficient on a time trend, p is the lag order autoregressive process, γ represents the process root coefficient, δ_t represents the lag operator, and ϵ_t is the independent identical distribution residual term with mean=0 and variance $\sigma^2=0$.

Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test The null hypothesis here is the stationary of time series around mean

or a linear trend, while the alternative hypothesis assumes a non-stationary time series because of the existence of a unit root (Kwiatkowski et al. 1992). Time series Y_1, Y_2, \dots, Y_n can be decomposed to a deterministic trend (β_t), random walk (r_t), and stationary error (ϵ_t):

$$Y_t = r_t + \beta_t + \epsilon_t \quad (2)$$

$$r_t = r_{t-1} + u_t \quad (3)$$

For time series Y_t with a deterministic stationary trend, the null hypothesis is expressed as $\sigma_U^2 = 0$, meaning that the intercept is a fixed component, against the alternative of a positive σ_U^2 . For this case, the residual error $e_t = \epsilon_t$, where $e_t = Y_t - \hat{Y}_t$ and $Y_t = r_0 + \beta_t + \epsilon_t$ (H0). The H1 specifies that $e_t = r_t + \epsilon_t$, which means the process has a unit root. The general form for the KPSS test is as follows:

$$KPSS = \frac{1}{T^2} \sum_{t=1}^t \frac{s_t^2}{\hat{\sigma}_\infty^2} \quad (4)$$

$$s_t = \sum_{j=1}^t e_j \quad (5)$$

$$\hat{\sigma}_\infty^2 = \lim_{t \rightarrow \infty} \text{var} \left(\sum_{i=1}^t r_i \right) \tag{6}$$

Pettitt’s test The Pettitt’s test is based on the Mann–Whitney two-sample test (rank-based test) and allows the detection of a single shift at an unknown time t (Pettitt 1979). The null hypothesis is no change in the distribution of a sequence of random variables; the alternative hypothesis is that the distribution function $F1(x)$ of the random variables from X_1 to X_t is different from the distribution function $F2(x)$ of the random variables from X_{t+1} to X_T . Hence:

$$D_{ij} = \text{sgn} (X_i - X_j) = \begin{cases} -1 & (X_i - X_j) < 0 \\ 0 & (X_i - X_j) = 0 \\ +1 & (X_i - X_j) > 0 \end{cases} \tag{7}$$

where X_i and X_j are random variables with X_i following X_j in time. The test statistic $U_{i,T}$ depends on D_{ij} as:

$$U_{i,T} = \sum_{j=1}^t \sum_{j=t+1}^T D_{ij} \tag{8}$$

The statistic $U_{i,T}$ is the same as the Mann–Whitney statistic for analyzing when the two samples X_1, \dots, X_t and X_{t+1}, \dots, X_T arise from the same population. The test statistic $U_{i,T}$ is assessed for all random variables from 1 to T ; then the most significant change point is selected where the value of $|U_{i,T}|$ is the largest:

$$K_T = \max_{1 \leq t < T} |U_{i,T}| \tag{9}$$

A change-point occurs at time t when the statistic K_T is significantly different from zero at a given level. The approximate significant level is given by:

$$p = 2 \cdot \exp \left(\frac{-6k_T^2}{T^2 + T^3} \right) \tag{10}$$

Once the p -value is less than the pre-assigned significance level α , we can reject the null hypothesis and divide the data into two sub-series (before and after the location of the change-point) with two different distribution functions.

Extrem value analysis (EVA)

A statistical distribution is fitted to a series of observations based on which the magnitude and probability of the occurrence of the variable under study are determined. It is recommended to use GEV and GPD distributions, respectively, for fitting the best function distribution on AMS and POT data.

Generalized extreme value (GEV) The GEV distribution method can flexibly model different behaviors of extremes with three distribution parameters $\theta = (\mu, \sigma, \epsilon)$: (1) the

location parameter (μ) shows the center of the distribution; (2) the scale parameter (σ) defines the deviation size around the location parameter; and (3) the shape parameter (ξ) oversees the tail behavior of the GEV distribution (Delgado et al. 2010). This distribution is a three-parameter model incorporating Gumbel, Fréchet, and Weibull maxima extreme value distribution (Coles et al. 2001) as in the following equation:

$$Y_t \sim \text{GEV}(\mu_t, \sigma_t, \epsilon_t) \tag{11}$$

$$\text{GEV}(\mu_t, \sigma_t, \epsilon_t) = \begin{cases} \exp \left(\left[\left(1 + \xi \left(\frac{Y-\mu}{\sigma} \right) \right)^{\frac{-1}{\xi}} \right] \right) & \text{if } \epsilon = 0 \\ \exp \left(-e^{-\left(\frac{Y-\mu}{\sigma} \right)} \right) & \text{if } \epsilon = 0 \end{cases} \tag{12}$$

In non-stationary conditions, the parameters of the distribution function are time-dependent. Since modeling temporal changes in the shape parameter requires long-term observations that are often not available for practical applications, non-stationary behavior concerning the location parameter and the location and scale parameters is assumed to be a linear function of time.

$$\text{GEV}(\mu(t) = \mu_0 + \mu_1 t, \sigma, \epsilon) \tag{13}$$

$$\text{GEV}(\mu(t) = \mu_0 + \mu_1 t, \sigma(t) = \exp(\sigma_0 + \sigma_1 t), \epsilon) \tag{14}$$

Generalized Pareto distribution (GPD) The cumulative probability of the generalized Pareto distribution (GPD) is calculated from the following equation:

$$\text{Pr}[X < x] \approx G(x - u' \epsilon' \sigma) = \begin{cases} 1 - \left(1 - \frac{\xi(x-u)}{\sigma} \right)^{\frac{1}{\xi}} & \epsilon \neq 0 \\ 1 - \exp\left(-\frac{x-u}{\sigma}\right) & \epsilon = 0 \end{cases} \tag{15}$$

where u, ξ , and σ are threshold, scale, and shape parameters, respectively.

Threshold selection

In this context, two methods are used to select the thresholds, namely MRL plot and Hill plot:

Mean residual life plot (MRL plot) Davison and Smith (1990) introduced MRL plot to determine the threshold using the expectation of the GPD excesses:

$$e(u) = E(x - u | x > u) = \frac{\sigma_u + \epsilon_u}{1 - \epsilon}, \epsilon < 1 \tag{16}$$

where u denotes the threshold, ϵ denotes the shape parameter, and σ_u denotes the scale parameter corresponding to threshold u . The condition of $\epsilon < 1$ is defined to guarantee the existence of the expectation. Equation (1) shows that the expectation of excesses is linear in u with gradient $\epsilon / (1 - \epsilon)$ and intercept $\sigma_u / (1 - \epsilon)$. For a set of samples (X_1, X_2, \dots, X_n) , the empirical estimate of the mean of excesses can be obtained as follows:

$$e_n(u) = \frac{1}{N_u} \sum_{i=1}^{N_u} (X_i - u), X_i > u \tag{17}$$

where N_u denotes the number of exceedances over u . The set data of $\{u, e_n(u)\}$ represent the MRL plot. Generally, the value of u above which the plot is an approximately straight line can be selected as the optimal threshold.

Hill plot Let $X_{(1,n)} > X_{(2,n)} > \dots > X_{(n,n)}$ be associated with the descending order statistics of (X_1, X_2, \dots, X_n) , which are independent and identically distributed random variables. Assuming that the distribution of these random variables is heavy-tailed, the Hill estimator, a well-known estimator of ϵ , is defined as:

$$H_{k, n} = \frac{1}{k} \sum_{i=1}^k \log \left(\frac{X_{(i,n)}}{X_{(i+1,n)}} \right), k \leq n \tag{18}$$

The Hill estimator is a function of these extreme random variables $\{X_{(1,n)}, X_{(2,n)}, \dots, X_{(k,n)}\}$, which depends on the chosen threshold. A Hill plot is constructed by the Hill estimator of a range of k values versus the k value or the threshold. The value of $X_{k,n}$ above which the Hill estimator tends to be stable can be chosen as the optimal threshold (Hill 1975).

Estimating distribution parameters

In this section, the DE-MC was used to estimate the distribution parameters based on Bayesian theory. In the Bayesian method, inferences about model parameters are based on their posterior distribution, which is a combination of observed data and information from previous studies or personal experiences known as prior distribution (Renard et al. 2013; Cheng et al. 2014). Differential evolution (DE) is a simple genetic algorithm for numerical optimization in real parameter spaces. In a statistical context one would not just want the optimum but also its uncertainty. The uncertainty distribution can be obtained by a Bayesian analysis (after specifying prior and likelihood) using Markov chain Monte Carlo (MCMC) simulation. DE-MC is a population MCMC algorithm, in which multiple chains are run in parallel. DE-MC solves an important problem in MCMC, namely that of choosing an appropriate scale and orientation for the jumping distribution. In DE-MC, the jumps are simply a

multiple of the differences of two random parameter vectors that are currently in the population. Simulations and examples illustrate the potential of DE-MC. In the fact, the DE-MC combines the genetic algorithm called differential evolution (DE) for global optimization over real parameter space with Markov chain Monte Carlo (MCMC) so as to generate a sample from a target distribution. In Bayesian analysis, the target distribution is typically a high dimensional posterior distribution. Both DE and MCMC are enormously popular in a variety of scientific fields for their power and general. Briefly, the advantages of DE-MC over conventional MCMC are simplicity, speed of calculation, and convergence, even for nearly collinear parameters and multimodal densities (Ter Braak 2006).

Evaluation criteria

Criteria such as the coefficient of determination (R^2), root mean square error (RMSE), mean bias error (MBE), and convergence (ϵ) (Gelman and Shirley 2011) were used for data analysis and model evaluation. Their relationships are presented below:

$$RMSE = \sqrt{\frac{\sum_{m=1}^n (X_p - X_0)^2}{n}} \tag{19}$$

$$R^2 = \left[\frac{\frac{1}{n} \sum_{m=1}^n (X_p - \mu_0)(X_p - \mu_0)}{\sigma_{X_p} \times \sigma_{X_0}} \right]^2 \tag{20}$$

$$MBE = \frac{\sum_{m=1}^n (X_p - X_0)}{n} \tag{21}$$

$$\epsilon_i = \frac{|\sum_{i=9}^{i-5} OF_i - \sum_{i=4}^i OF_i|}{\sum_{i=4}^i OF_i} \tag{22}$$

where X_p and X_0 are the simulated and observed data, respectively, μ is the mean of the data population, σ is the standard deviation, n is the total number of data, i is the current iteration number (> 10), OF_i is the objective function value in the i^{th} iteration, and ϵ_i is the convergence value of the objective function in the i^{th} iteration. R^2 represents the linear relationship between simulated and observed data, which is between 0 and 1. The closer to 1, R^2 represents a stronger linear relationship between the simulated and observed data.

Return level and return period

The year m return level is the level where the number of expected events in an m year period is one. In the stationary

state, the level of return is the same for all years, and there is a one-to-one relationship between the level of return (multiple) and the period of return (related time interval). The return level is expressed as a function of the return period T:

$$T = \frac{1}{1 - p} \tag{23}$$

P is the occurrence probability in a defined year (assuming stationary). The return level in the stationary state is obtained as follows:

$$q_p = \left(\left(-\frac{1}{\ln p} \right)^\xi - 1 \right) \times \frac{\sigma}{\xi} + (\neq 0) \tag{24}$$

The model parameters are used to calculate the return level in the non-stationary state as follows:

$$\tilde{\mu} = Q_K(\mu_{t1}, \mu_{t2}, \dots, \mu_{tn}), (\mu(t) = \mu_{t1} + \mu_0) \tag{25}$$

$$q_p = \left(\left(-\frac{1}{\ln p} \right)^\xi - 1 \right) \times \frac{\sigma}{\xi} + \tilde{\mu} (\xi \neq 0) \tag{26}$$

where $k = 0.5$ is the median of location parameters and $Q_K(\mu_{t1}, \mu_{t2}, \dots, \mu_{tn})$, and $k = 0.95$ is related to the ninth percentile of location parameters (Cheng et al. 2014). It should be noted that NEVA packages were used to perform the calculations of frequency analysis (Cheng et al. 2014).

Results

Assessment of CRU dataset

To use CRU data, their correlation with station observation-based data should be examined. To this end, the maximum monthly temperature data of the stations in the study area were extracted and compared with the observational data of the Arak synoptic station in the same period (during from 1956 to 2010). It should be noted that the nearest grid ($x = 49.75, y = 34.25$) to Arak synoptic station was considered for comparing purposes. The comparison results showed that these data have a high correlation (0.98) and the lowest error (1.49) with observation station-based data (Fig. 3). The data had the necessary accuracy for other stations as well.

Stationary analysis tests

The Mann–Kendall test was used to examine the trend of time series of extreme maximum temperature values (Mann 1945; Kendall 1975). The results showed that the τ values were from 0.25 to 0.33 indicating the independence of the temperature data from the time (Table 1). In addition, the mean son’s slope was about 0.011 that showed a significant increasing trend. Finally, as the p -value in all selected stations was less than the significance level (5%), the null hypothesis of none of the trends was rejected (Fig. 4a).

Fig. 3 The variation of maximum monthly temperature (observational and CRU) during 1956–2010

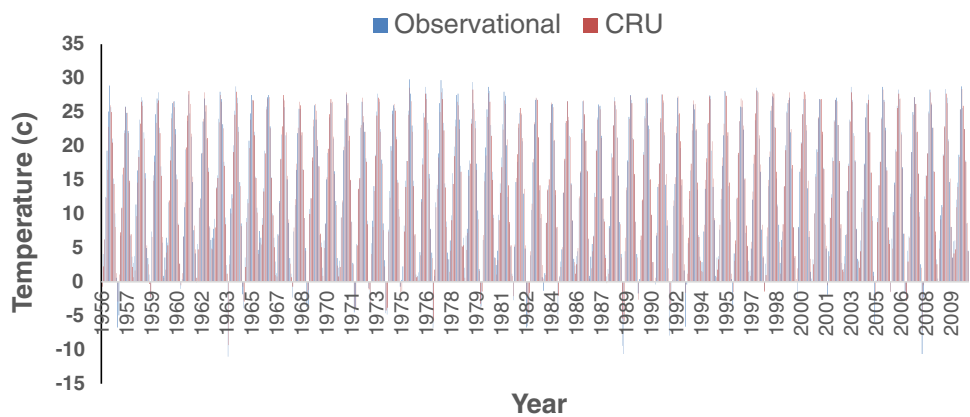


Table 1 Stationary tests for the selected stations

Test	Mankell			Stationary				Homogeneity	
	p -value	Sen’s slope	τ	p -value	ADF	p -value	KPSS	p -value	Changing point
Arak	0.0001	0.011	0.31	0.33	-2.15	0.03	0.16	0.0001	1957
Ashtian	0.0001	0.01	0.25	0.52	-1.69	0.02	0.17	0.0001	1958
Sarugh	0.0001	0.011	0.32	0.46	-1.86	0.02	0.19	0.0001	1958
Gavar	0.0001	0.012	0.33	0.24	-2.38	0.03	0.16	0.0001	1957
Shanagh	0.0001	0.011	0.31	0.36	-2.09	0.04	0.15	0.0001	1958

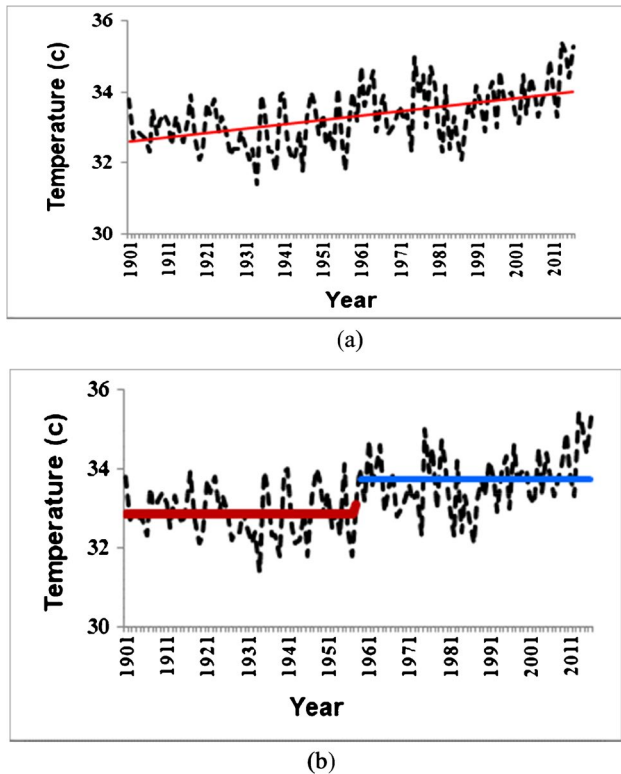


Fig. 4 The trend (a) and homogeneity (b) of extreme maximum temperature in Gavar station

As illustrated in Fig. 4, Gavar station shows an increasing trend. Then, KPSS and ADF tests were used to assess the stationarity of these time series. The results showed the time series of extreme temperature were non-stationary in

all selected stations (Table 1). For instance, in Gavar station, the ADF statistic was obtained at -2.38 , i.e., lower than the critical values (0.6). The p -value was greater than the (5%) significance level, and so it is concluded that the null hypothesis of the non-stationary state should be accepted. For the KPSS test, since the p -value was 0.03 and smaller than the (5%) significance level, and the KPSS statistic was greater than the critical value (0.16), the null hypothesis was rejected (i.e., the data were stationary). Therefore, the time series data in the Gavar station was non-stationary. In the following, the homogeneity of extreme temperature was assessed using Pettit’s test over the study area. The results show that all stations were non-homogeneous, as presented in Fig. 4b for the Gavar station.

Frequency analysis of GEV distribution

In this section, Q-Q probability plot was used to fit the GEV distribution to maximum temperature (for Gavra station, Fig. 5). As illustrated in this figure, the GEV theoretical distribution is in good relation with the one obtained from the empirical distribution. This issue is also confirmed by other stations.

In the following, distribution parameters were obtained using the DE-MC method (for Gavra station, Fig. 6). According to the white dashed line in this figure, the non-stationary behavior of maximum temperature can be expressed by parameters of $\mu_0 = 32.38$, $\mu_1 = 0.011$, scale = 0.75, and shape = -0.21 , which were obtained by averaging 5000 iterations.

As evaluation criteria, scale and shape parameters do not show a non-stationary behavior, unlike the location

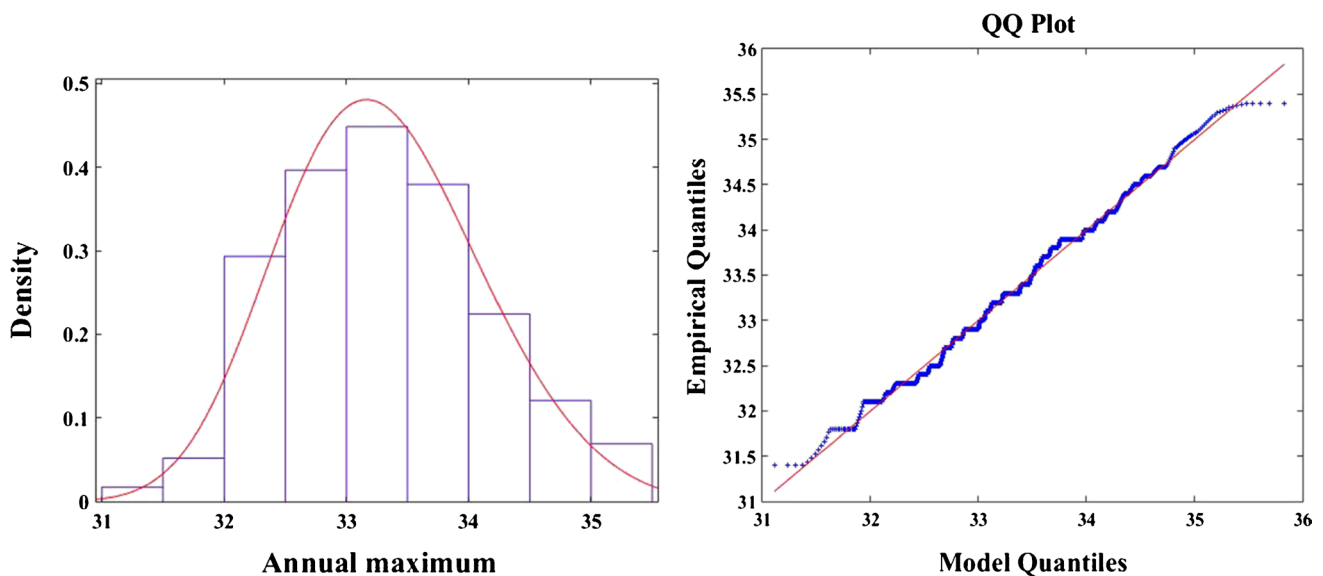


Fig. 5 The GEV distribution for extreme temperature in Gavar station

Fig. 6 DE-MC realizations of the GEV parameters with a Bayesian analysis in Gavar station

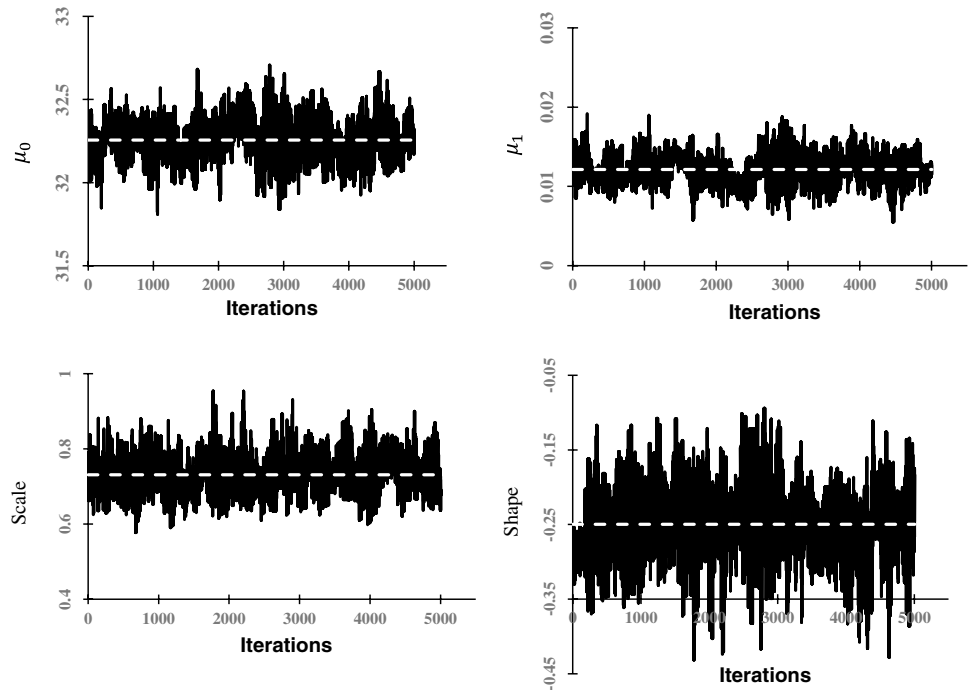


Table 2 The GEV distribution parameters and evaluation criteria for the selected station

Station	State	DE-MC	
		Convergence	Parameters (μ, σ, ϵ)
Arak	Stationary (GEV0)	0.3312	33.34, 0.77, -0.15
	Non-stationary (GEV1)	0.4454	34.75, 0.70, -0.19
Ashtian	Stationary (GEV0)	0.3827	33.64, 0.85, -0.23
	Non-stationary (GEV1)	0.4157	33.05, 0.79, -0.22
Sarugh	Stationary (GEV0)	0.3661	32.37, 0.77, -0.18
	Non-stationary (GEV1)	0.4161	31.72, 0.73, -0.32
Gavar	Stationary (GEV0)	0.3756	32.98, 0.79, -0.19
	Non-stationary (GEV1)	0.4193	32.38, 0.75, -0.21
Shanagh	Stationary (GEV0)	0.4067	34.19, 0.77, -0.149
	Non-stationary (GEV1)	0.4460	33.62, 0.70, -0.22

parameter. Therefore, only the non-stationary state with respect to μ (location parameter) was discussed by considering a linear function model: $\mu(t) = 32.38 + 0.011t$. According to Table 2, convergence values were higher in the non-stationary than the stationary state, indicating non-stationary behavior in the selected station.

Frequency analysis of GPD distribution

First, the threshold value was estimated using the mentioned methods (for Gavar station, Fig. 7). In the first method, the threshold value is the area at which the graphs begin to linearize with a high slope. In the second method (the Hill estimator), there is a relatively large deviation from the straight line. It should be noted, as long as data above the threshold follows the GPD distribution, the threshold selection is somewhat optional (Fig. 7). The results showed that the threshold values were 35.3, 32.7, 32.3, 33, and 34.4 for Arak, Ashtian, Sarugh, Gavar, and Shanagh stations, respectively (Fig. 8). Then, the GPD distribution parameters were obtained based on the DE-MC method (Table 3).

Frequency analysis

In this section, by comparing the convergence criteria, it can be said that GEV outperformed GPD in modeling extreme temperature. For instance, the convergence values in the Gavra station for both non-stationary GEV and GPD were 0.4193 and 0.3954, respectively. After detecting the GEV distribution as the best model and estimating its parameters, the values of the return level in the different return periods of 2, 10, 20, 50, and 100 years in the stationary and non-stationary state were determined. For example, in the Gavar station in the stationary state, the values of the return periods were constant and unchanged for all years, while in the non-stationary state, the temperature increases (Fig. 9).

Fig. 7 Threshold selection by two methods: **a** mean residual life, **b** Hill plot, in Gavar station

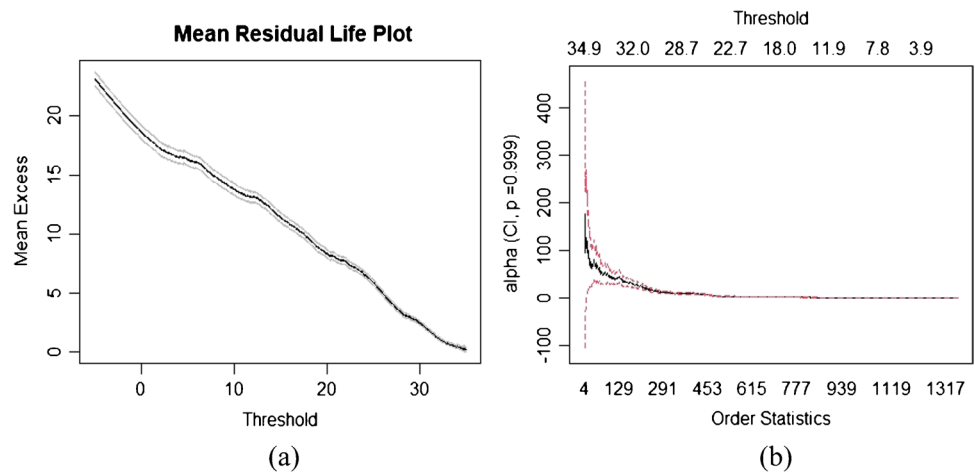


Fig. 8 The GPD distribution for extreme temperature in Gavar station

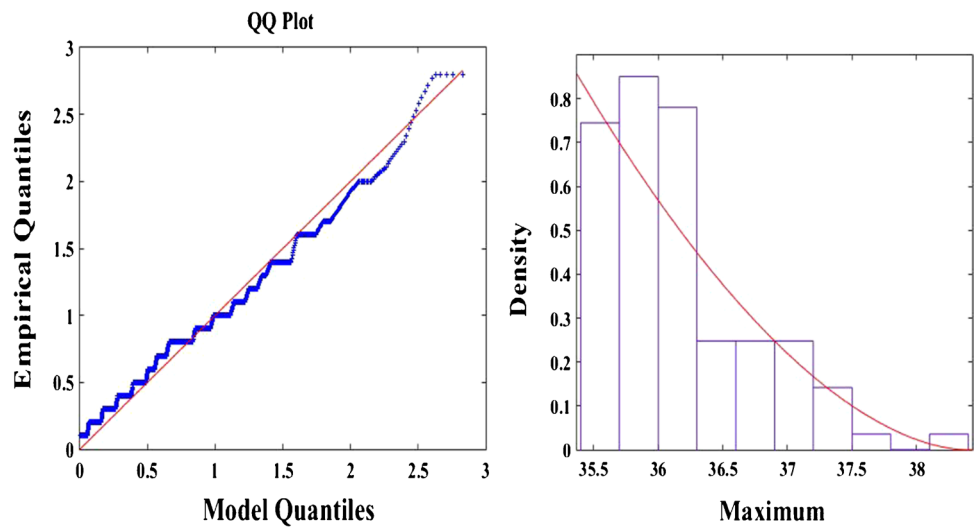


Table 3 The GPD distribution parameters and evaluation criteria for the selected station

Station	State	DE-MC		Threshold	Data above threshold
		Convergence	Parameters (σ, ϵ)		
Arak	Stationary (GEV0)	0.3587	1.1, -0.3	35.3	94
	Non-stationary (GEV1)	0.3912	1.16, -0.44		
Ashtian	Stationary (GEV0)	0.3641	1.08, -0.37	33.7	92
	Non-stationary (GEV1)	0.4104	1.13, -0.41		
Sarugh	Stationary (GEV0)	0.3521	1.01, -0.31	32.3	108
	Non-stationary (GEV1)	0.3845	1.17, -0.53		
Gavar	Stationary (GEV0)	0.3689	1.05, -0.35	33	95
	Non-stationary (GEV1)	0.3954	1.16, -0.52		
Shanagh	Stationary (GEV0)	0.3427	1.04, -0.35	34.4	74
	Non-stationary (GEV1)	0.4193	1.03, -0.36		

Fig. 9 Return periods of maximum temperature amount for stationary (a) and non-stationary (b) states in Gavar station

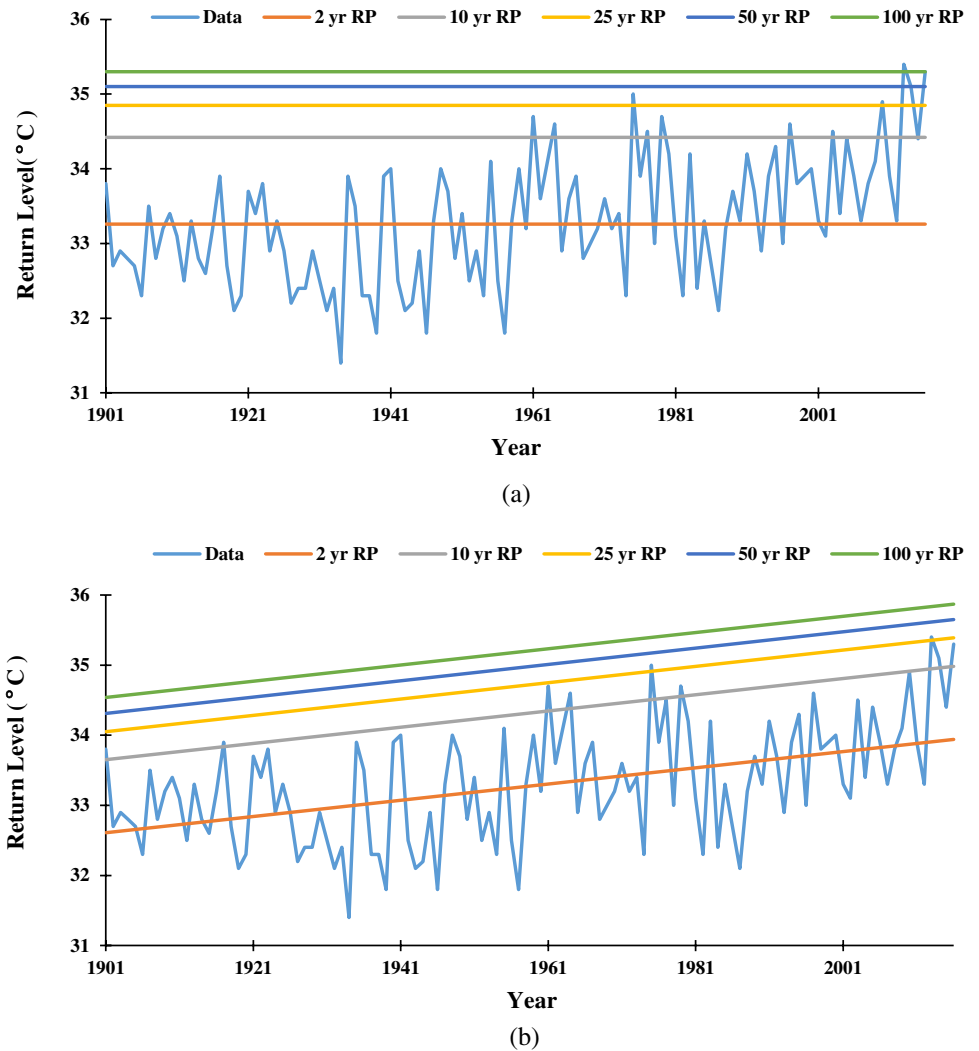


Table 4 The return level of extreme temperature in the non-stationary state based on AMS

Station/return period	2-year	100-year
Arak	34.98–36.26	36.82–38.1
Ashtian	33.38–34.5	35.25–36.36
Sarugh	32.02–33.38	33.54–34.9
Sanagh	33.87–35.15	35.74–37.01
Gavar	32.57–33.95	34.37–35.75

As indicated in Table 4, for the 100-year return period, the average value of temperature is 37.52 °C in the stationary state, whereas in the non-stationary state, the temperature changes from 34.37 °C (1901) to 35.75 °C (2016).

Conclusion

Extreme temperatures as a record-breaking phenomenon have been occurred all over Iran including different climatic regions. In this county, a positive trend in annual maximum temperature (AMT) has been occurred during recent decades which causes a non-stationary (NS) conditions in its extreme temperatures. The main objective of this study was to compare extreme value theory (EVT) methods in the non-stationary analysis of maximum temperature in Arak plain. To achieve this purpose, the maximum temperature was collected from the CRU gridded dataset for the selected stations of the study area during 1901–2016. Time series of maximum temperature were extracted using two methods of block maxima (BM) and peaks-over-threshold (POT). The results showed that the monthly temperature data extracted from the CRU climate database have good accuracy and validity for the region. Hence, the results of MK, ADF, KPSS, and Petit statistical tests showed that the maximum annual temperature of the selected stations has a trend and is non-stationary

and heterogeneous. Threshold selection methods (MRL plot and Hill plot) showed no difference among the threshold values. Therefore, the mean of threshold values was used as a selected threshold to extract time series. Moreover, the comparison of convergence evaluation criteria for the annual maximum temperature time series using AMS sampling was more accurate than the POT sampling method. Regarding the convergence criteria for the AMS approach, Arak and Shanagh stations were more non-stationary than other stations. A comparison of the maximum annual temperature values in stationary and non-stationary states in different return periods showed that the maximum temperature difference is about 0.8 °C in the short return period (2 years). Finally, the findings in this study indicate that the consideration of non-stationarity in extreme temperature time series is a necessity during return level estimations over the study area. Also, we recommend that for nonstationary, extreme value modeling is employed exogenous covariates such as large-scale climate modes and hydrological variables.

Author contribution MM and SA: conceptualization, methodology, technical investigation, writing, reviewing and editing, visualization, software, data curation, validation, and editing. TM: original draft preparation.

Data availability Available on request.

Code availability The software used in this research will be available (by the corresponding author), upon reasonable request.

Declarations

Ethics approval Not applicable, because this article does not contain any studies with human or animal subjects.

Consent to participate The research data were not prepared through a questionnaire.

Consent for publication There is no conflict of interest regarding the publication of this article.

Conflict of interest The authors declare no competing interests.

References

- AghaKouchak A, Easterling D, Hsu K, Schubert S, Sorooshian S (2013) *Extremes in a Changing Climate*. Springer, Netherlands
- Aziz R, Yucel I, Yozgatligil C (2020) Nonstationarity impacts on frequency analysis of yearly and seasonal extreme temperature in Turkey. *Atmos Res* 238:104875. <https://doi.org/10.1016/j.atmosres.2020.104875>
- Alexander LV, Zhang X, Peterson TC, Caesar J et al (2006) Global observed changes in daily climate extremes of temperature and precipitation. *J Geophys Res* 111(D5). <https://doi.org/10.1029/2005JD006290>
- AghaKouchak A, Nasrollahi N (2010) Semi-parametric and parametric inference of extreme value models for rainfall data. *Water Resour Manage* 24(6):1229–1249. <https://doi.org/10.1007/s11269-009-9493-3>
- Babaeian I, Karimian M, Modirian R, Mirzaei E (2019) Future climate change projection over Iran using CMIP5 data during 2020–2100. *NIVAR J Meteorol Organization* 43:61–70
- Banerjee A, Dolado JJ, Galbraith JW, Hendry D (1993) Co-integration, error correction, and the econometric analysis of non-stationary data. OUP Catalogue. Oxford University Press, Oxford. <https://doi.org/10.2307/2235236>
- Climate Change Science Program (2008) *Weather and climate extremes in a changing climate. Regions of Focus: North America, Hawaii, Caribbean, and US Pacific Islands*
- Coles S, Bawa J, Trenner L, Dorazio P (2001) *An introduction to statistical modeling of extreme values*, Springer, London, 208, p 208
- Cooley D (2013) Return periods and return levels under climate change, *Extremes in a Changing Climate*. Springer, Netherlands. https://doi.org/10.1007/978-94-007-4479-0_4
- Cheng L, AghaKouchak A, Gilleland E, Katz RW (2014) Non-stationary extreme value analysis in a changing climate. *Clim Change* 127(2):353–369. <https://doi.org/10.1007/s10584-014-1254-5>
- Davison AC, Smith RL (1990) Models for exceedances over high thresholds. *J Royal Stat Soc: Series B (Methodological)* 52(3):393–442. <https://doi.org/10.1111/j.2517-6161.1990.tb01796.x>
- Delgado JM, Apel H, Merz B (2010) Flood trends and variability in the Mekong river. *Hydrol Earth Syst Sci* 14(3):407–418
- Diffenbaugh NS, Giorgi F (2012) Climate change hotspots in the CMIP5 global climate model ensemble. *Climate Chang* 114(3):813–822. <https://doi.org/10.1007/s10584-012-0570-x>
- Easterling DR, Meehl GA, Parmesan C et al (2000) Climate extremes: observations, modeling, and impacts. *Science* 289(5487):2068–2074
- Ejder T, Kale S, Acar S, Hisar O, Mutlu F (2016) Effects of climate change on annual streamflow of Kocabaş Stream (Çanakkale, Turkey). *J Sci Res Rep* 11(4):1–11. <https://doi.org/10.9734/JSRR/2016/28052>
- Field CB, Barros V, Stocker TF, Dahe Q (Eds.) (2012) *Managing the risks of extreme events and disasters to advance climate change adaptation: special report of the intergovernmental panel on climate change*. Cambridge University Press
- Ferreira A, De Haan L (2015) On the block maxima method in extreme value theory: PWM estimators. *Ann Stat* 43(1):276–298
- Gabda D, Tawn J, Brown S (2018) A step towards efficient inference for trends in UK extreme temperatures through distributional linkage between observations and climate model data. *Natural Hazards*, 1–20 <https://doi.org/10.1007/s11069-018-3504-8>
- Gao M, Zheng H (2018) Nonstationary extreme value analysis of temperature extremes in China. *Stoch Env Res Risk Assess* 32(5):1299–1315. <https://doi.org/10.1007/s00477-017-1482-0>
- Gelman A, Shirley K (2011) Inference from simulations and monitoring convergence. *Handbook of markov chain monte carlo, Inference from Simulations and Monitoring Convergence* 163–174
- Gilroy KL, McCuen RH (2012) A nonstationary flood frequency analysis method to adjust for future climate change and urbanization. *J Hydrol* 414:40–48. <https://doi.org/10.1016/j.jhydrol.2011.10.009>
- Gumbel EJ (1958) *Statistics of extremes*. Columbia Univ. Press, New York
- Hawkes PJ, Gonzalez-Marco D, Sánchez-Arcilla A, Prinos P (2008) Best practice for the estimation of extremes: a review. *J Hydraul Res* 46:324–332. <https://doi.org/10.1080/00221686.2008.9521965>
- Hill BM (1975) A simple general approach to inference about the tail of a distribution. *The annals of statistics*, 1163–1174. <https://www.jstor.org/stable/2958370>
- IPCC (2012) Glossary of terms, in: *managing the risks of extreme events and disasters to advance climate change adaptation*. In: field CB, Barros V, Stocker TF, Qin D, Dokken DJ, Ebi KL...

- Midgley PM (Eds.) A special report of working groups I and II of the Intergovernmental Panel on Climate Change (IPCC). Cambridge University Press, Cambridge, UK, and New York, NY, USA, pp 555–564.
- Jakob D (2013) Nonstationarity in extremes and engineering design. *Extremes in a Changing Climate*, Netherlands https://doi.org/10.1007/978-94-007-4479-0_13
- Katz RW, Brown BG (1992) Extreme events in a changing climate: variability is more important than averages. *Clim Change* 21(3):289–302. <https://doi.org/10.1007/BF00139728>
- Katz RW, Parlange MB, Naveau P (2002) Statistics of extremes in hydrology. *Adv Water Resour* 25(8–12):1287–1304. [https://doi.org/10.1016/S0309-1708\(02\)00056-8](https://doi.org/10.1016/S0309-1708(02)00056-8)
- Kendall MG (1975) Rank correlation methods. Oxford University Press, New York, NY
- Kharin VV, Zwiers FW, Zhang X, Hegerl GC (2007) Changes in temperature and precipitation extremes in the IPCC ensemble of global coupled model simulations. *J Clim* 20(8):1419–1444. <https://doi.org/10.1175/JCLI4066.1>
- Kwiatkowski D, Phillips PC, Schmidt P, Shin Y (1992) Testing the null hypothesis of stationarity against the alternative of a unit root. *J Econom* 54(1–3):159–178. [https://doi.org/10.1016/0304-4076\(92\)90104-Y](https://doi.org/10.1016/0304-4076(92)90104-Y)
- Li L, Zhang L, Xia J, Gippel CJ, Wang R, Zeng S (2015) Implications of modelled climate and land cover changes on runoff in the middle route of the south to north water transfer project in China. *Water Resour Manage* 29(8):2563–2579. <https://doi.org/10.1007/s11269-015-957-3>
- Linnenluecke MK, Stathakis A, Griffiths A (2011) Firm relocation as adaptive response to climate change and weather extremes. *Glob Environ Chang* 21(1):123–133. <https://doi.org/10.1016/j.gloenvcha.2010.09.010>
- Mann HB (1945) Nonparametric tests against trend. *Econom: J Econom Soc* 245–259. <https://doi.org/10.2307/1907187>
- Meehl GA, Karl T, Easterling DR et al (2000) An introduction to trends in extreme weather and climate events: observations, socioeconomic impacts, terrestrial ecological impacts, and model projections. *Bull Am Meteorol Soc* 81(3):413–416
- Mearns LO, Katz RW, Schneider SH (1984) Extreme high-temperature events: changes in their probabilities with changes in mean temperature. *J Appl Meteorol Climatol* 23(12):1601–1613
- Mirza MMQ (2003) Climate change and extreme weather events: can developing countries adapt? *Climate Policy* 3(3):233–248. [https://doi.org/10.1016/S1469-3062\(03\)00052-4](https://doi.org/10.1016/S1469-3062(03)00052-4)
- Mitchell TD, Jones PD (2005) An improved method of constructing a database of monthly climate observations and associated high-resolution grids. *Int J Climatol: A J Royal Meteorol Soc* 25(6):693–712. <https://doi.org/10.1002/joc.1181>
- Moghaddasi M, Anvari S, Akhondi N (2022) A trade-off analysis of adaptive and non-adaptive future optimized rule curves based on simulation algorithm and hedging rules. *Theor Appl Climatol* 148:65–78. <https://doi.org/10.1007/s00704-022-03930-y>
- New M, Hulme M, Jones PD (1999) Representing twentieth-century-space-time climate variability. Part I: development of a 1961–90 mean monthly terrestrial climatology. *J Clim* 12:829–856
- Paap R (2002) What are the advantages of MCMC based inference in latent variable models? *Stat Neerl* 56(1):2–22. <https://doi.org/10.1111/1467-9574.00060>
- Parey S, Hoang TTH, Dacunha-Castelle D (2010) Different ways to compute temperature return levels in the climate change context. *Environmetrics* 21(7–8):698–718. <https://doi.org/10.1002/env.1060>
- Perkins SE, Alexander LV, Nairn JR (2012) Increasing frequency, intensity and duration of observed global heatwaves and warm spells. *Geophys Res Lett* 39(20)
- Pettitt AN (1979) A non-parametric approach to the change-point problem. *J Roy Stat Soc: Ser C (appl Stat)* 28(2):126–135
- Pickands J III (1975) Statistical inference using extreme order statistics. *Ann Statist* 3:119–131
- Raggad B (2018) Statistical assessment of changes in extreme maximum temperatures over Saudi Arabia, 1985–2014. *Theoret Appl Climatol* 132(3–4):1217–1235
- Renard B, Sun X, Lang M (2013) Bayesian methods for non-stationary extreme value analysis. *Extremes in a changing climate*. Springer, Dordrecht, pp 39–95
- Rosbjerg R, Madsen H (1998) Design with uncertain design values, Hydrology in a Changing Environment. Wiley, 3:155–163
- Salas JD, Obeysekera J (2013) Revisiting the concepts of return period and risk for nonstationary hydrologic extreme events. *J Hydrol Eng* 19(3):554–568. [https://doi.org/10.1061/\(ASCE\)HE.1943-5584.0000820](https://doi.org/10.1061/(ASCE)HE.1943-5584.0000820)
- Ter Braak CJ (2006) A Markov chain Monte Carlo version of the genetic algorithm differential evolution: easy Bayesian computing for real parameter spaces. *Stat Comput* 16(3):239–249. <https://doi.org/10.1007/s11222-006-8769-1>
- Tian P, Mu X, Liu J, Hu J, Gu C (2016) Impacts of climate variability and human activities on the changes of runoff and sediment load in a catchment of the Loess Plateau, China. *Adv Meteorol*. <https://doi.org/10.1155/2016/4724067>
- Vose RS, Easterling DR, Gleason B (2005) Maximum and minimum temperature trends for the globe: an update through 2004. *Geophys Res Lett* 32(23). <https://doi.org/10.2307/2892861>.
- Young AF (2013) Urban expansion and environmental risk in the São Paulo Metropolitan Area. *Climate Res* 57(1):73–80. <https://doi.org/10.3354/cr01161>
- Zamani R, Berndtsson R (2018) Evaluation of CMIP5 models for west and southwest Iran using TOPSISI-based method. *Theoret Appl Climatol* 137:533–543