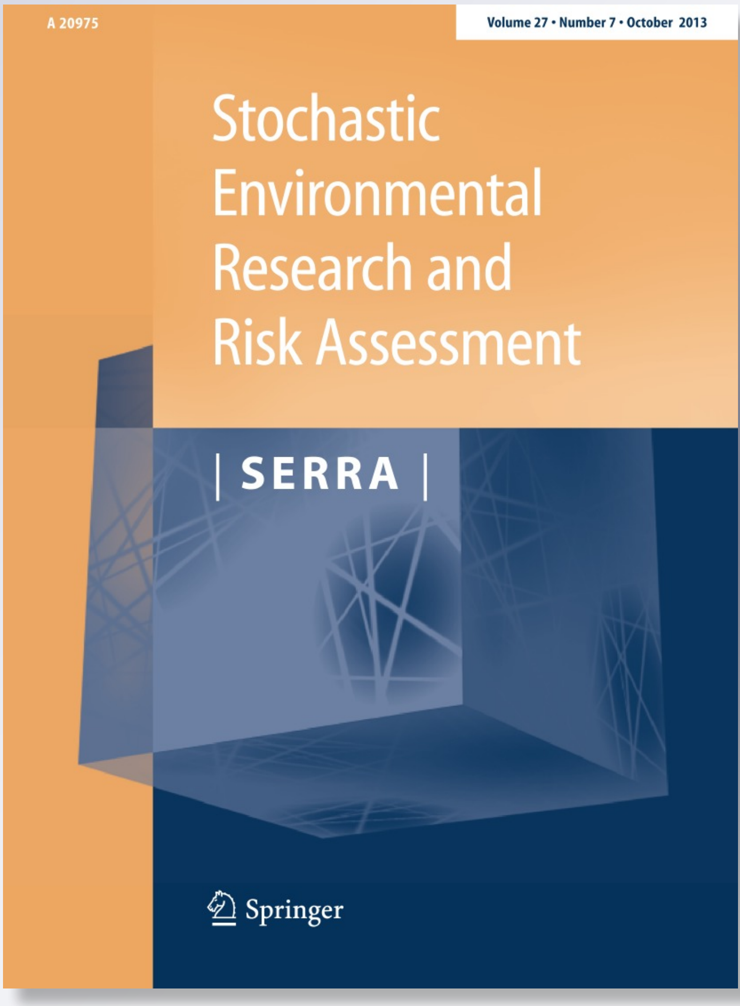
Frequency analysis of climate extreme events in Zanjan, Iran

Saeed Jahanbaksh Asl, Ali Mohammad Khorshiddoust, Yagob Dinpashoh & Fatemeh Sarafrouzeh

Stochastic Environmental Research and Risk Assessment

ISSN 1436-3240
Volume 27
Number 7

Stoch Environ Res Risk Assess (2013)
27:1637-1650
DOI 10.1007/s00477-013-0701-6



Springer

Stoch Environ Res Risk Assess (2013) 27:1637–1650
DOI 10.1007/s00477-013-0701-6

ORIGINAL PAPER

Frequency analysis of climate extreme events in Zanjan, Iran

Saeed Jahanbaksh Asl · Ali Mohammad Khorshiddoust · Yagob Dinpashoh · Fatemeh Sarafrouzeh

Published online: 1 March 2013

Abstract In this study, generalized extreme value distribution (GEV) and generalized Pareto distribution (GPD) were fitted to the maximum and minimum temperature, maximum wind speed, and maximum precipitation series of Zanjan. Maximum (minimum) daily and absolute annual observations of Zanjan station from 1961 to 2011 were used. The parameters of the distributions were estimated using the maximum likelihood estimation method. Quantiles corresponding to 2, 5, 10, 25, 50, and 100 years return periods were calculated. It was found that both candidate distributions fitted to extreme events series, were statistically reasonable. Most of the observations from 1961 to 2011 were found to fall within 1–10 years return period. Low extremal index (θ) values were found for excess maximum and minimum temperatures over a high threshold, indicating the occurrence of consecutively high peaks. For the purpose of filtering the dependent observations to obtain a set of approximately independent threshold excesses, a declustering method was performed, which separated the excesses into clusters, then the de-clustered peaks were fitted to the GPD. In both models, values of the shape parameters of extreme precipitation and extreme wind speed were close to zero. The shape parameter was less negative in the GPD than the GEV. This leads to significantly lower return period estimates for high extremes with the GPD model.

S. Jahanbaksh Asl · A. M. Khorshiddoust · F. Sarafrouzeh (✉)
Department of Physical Geography, Faculty of Geography, University of Tabriz, 29 Bahman Blvd, Tabriz, Iran
e-mail: fa_sa777@yahoo.com

Y. Dinpashoh
Department of Water Engineering, University of Tabriz, Tabriz, Iran

Keywords Climate extreme events · Generalized extreme value · Generalized Pareto · Maximum likelihood · Return period

1 Introduction

Extreme events are important aspects of any climate. Changes in the magnitude and frequency of climatic extremes would have environmental and socio-economic consequences. It is therefore, of great interest to analyze the extreme events. A widely used theorem for extreme events follows one of three types of distribution, including Gumbel, Frechet and Weibull. These distributions can be written in a single expression as a family of distributions referred to as the generalized extreme value (GEV) distribution (Coles 2001; Reiss and Thomas 2001; Beirlant et al. 2004). If daily data are available, the generalized Pareto distribution (GPD) is proposed for application. The GPD approach has advantages with respect to the GEV approach because it adapts better to heavy-tailed distributions and permits the consideration of more extreme cases (Unkasevic and Tosic 2009). GEV and GPD have been used for the analysis of extreme values in meteorology, climatology, hydrology and evaluation of damage caused by such events (Katz et al. 2002; Smith 2003; Renald et al. 2006; de Oliveira et al. 2010; Van Den Brink et al. 2004; Payer and Kuchenhoff 2004; Kharin and Zwiers 2000; Holmes and Moriarty 1999; Sanabria and Cechet 2007; Kharin and Zwiers 2005; Abaurrea et al. 2007; Parey et al. 2007; Laurent and Parey 2007; Hashmi et al. 2011). Tagavi and Mohammadi (2007) analyzed return periods of climate extreme events using ETCCDMI indices in 16 synoptic stations of Iran. Trends and return periods of extreme indices in two periods (1961–1990 and 1990–2000) were determined. Asakereh (2012) studied frequency distribution change of extreme precipitation in Zanjan. In this paper GEV were fitted to precipitation equal or less than 5th percentile, precipitation equal or more than 95th percentile, total of maximum precipitation and total precipitation of five greatest precipitations in Zanjan for all time series as well as for two halves of periods (1961–1983 and 1984–2006). Results indicated that there is a reduction in lower and higher percentiles in the case of precipitation data of Zanjan. But quantiles of precipitation were not calculated. He did not include other extreme events such as temperature and wind speed. Moreover he didn't analyze data by the GPD method.

In this study, the frequency of absolute annual and daily values of maximum wind speed, maximum precipitation, maximum and minimum temperature in Zanjan were analyzed by using GEV and GPD. For performing the analysis, the R software package extRemes (Gilleland and Katz 2005a) was used. The paper is organized as follows: a description of the data and methodology used in this study is given in Sect. 2. Section 3 presents the results and discussion, and conclusions are presented in Sect. 4.

2 Data and methodology

Zanjan synoptic station is located in the northwest of Iran 48°29′ E and 36°41′ N, 1,663 m above the sea level. Data of daily maximum wind speed during 1961–2009 and data of daily precipitation, maximum and minimum temperature during 1961–2011 were used. These dataset have been supplied by the Islamic Republic of Iran Meteorological Organization.

GEV and GPD distributions were employed for frequency analysis of extreme events in the study area. The historical cornerstone of extreme value theory is the GEV distribution which classically models block maxima (Naveau et al. 2005). GEV distribution has been widely used for modeling natural phenomenon (Hosking et al. 1985). Let x_1, x_2, x_3,…, x_n, be a sequence of independent and identically distributed random variables, where x_j is the maximum value occurring in the jth year. The distribution of the extreme event magnitudes, x_j, is usually approximated by GEV distribution. The cumulative probability function (CDF) for the GEV distribution is given by:

$$F(x) = \exp - \left\{ - \left[1 - \xi \left(\frac{z-\mu}{\sigma} \right) \right]^{\frac{1}{\xi}} \right\}, \tag{1}$$

where μ, σ and ξ are location, scale and shape parameters, respectively. The shape parameter, ξ, determines whether or not the distribution has an upper bound. The former is true whenever $\xi < 0$; which corresponds to a reversed Weibull distribution, while there is no upper limit for $\xi > 0$; which is of Frechet type and $\xi = 0$ being interpreted as $x \to 0$ yielding a Gumbel distribution. Maximum likelihood method (MLM) was used for estimation of GEV parameters. This is due to the fact that MLM approach is considered to be the most efficient method for parameter estimation (Rao and Hamed 2000). Furthermore, this method provides the smallest sampling variance of the estimated parameters, and hence of the estimated quantiles, compared to other methods (Rao and Hamed 2000). The likelihood function for GEV is given by:

$$L = \prod_{i=1}^{N} \left\{ \frac{1}{\sigma} \left[1 - \xi \left(\frac{x_i - \mu}{\sigma} \right) \right]^{\frac{1}{\xi} - 1} e^{-\left[1-\xi\left(\frac{x_i-\mu}{\sigma}\right)\right]^{\frac{1}{\xi}}} \right\}, \tag{2}$$

where N is the number of observations. For the case of GEV distribution given in (1), the predicted return level (Z_p) corresponding to T-years return period is:

$$Z_p = \hat{\mu} - \frac{\hat{\sigma}}{\xi} \left[1 - \left\{ - \log \left(1 - \frac{1}{T} \right) \right\}^{-\xi} \right]. \tag{3}$$

Using the variance of the return levels we can calculate a confidence interval for the return levels. The confidence interval obtained by using the equation $\hat{Z}_P \pm 1.96\,\mathrm{var}(\hat{Z}_P)$ is based on asymptotic normality and so this will give symmetric confidence intervals around the mean. This method is called Delta method. However, this confidence interval does not take into account that with limited data we will know more about events closer in the future than further into the future, the more we extrapolate from the data the less confident we will be in our estimates. This can be rectified by taking the profile log-likelihood of the return levels. From this we can see that as we estimate the return level further into the future the profile log-likelihood becomes more skewed as we are less confident about the future (Nemeth 2011). Applying the methods to temperature data, Gilleland and Katz (2005b) found out that the profile-likelihood method gives better results because it considers the asymmetry of the data. Compared to other uncertainty calculation methods, it has the advantage that it utilizes more information from the sample, especially the information provided by the most extreme events. Using this method, it is also possible to obtain asymmetric uncertainty estimates, which, are more precise and should be used in situations where it is necessary to obtain accurate confidence intervals (Ceppi et al. 2008).

The profile-likelihood method assumes an χ^2 distribution for the samples and maximizes the log-likelihood function to calculate the appropriate confidence interval (Sanabria and Cechet 2007). To obtain the profile likelihood for the shape parameter, we fixed $\xi = \xi_0$ and maximized the log-likelihood with respect to the remaining parameters, μ and σ (Coles 2001). This was repeated for a range of values of ξ_0. The corresponding maximized values of the log-likelihood constitute the profile log-likelihood for ξ that is used to obtain approximate confidence intervals. We can obtain confidence intervals for any specified return level Z_p. This requires a re-parameterization of the GEV model, so that Z_p is one of the model parameters, after which the profile log-likelihood is obtained by maximization with respect to the remaining parameters in the usual way. Re-parameterization is straightforward as follows:

$$\mu = Z_p + \frac{\sigma}{\xi}\left[1 - \left\{-\log\left(1 - \frac{1}{T}\right)\right\}^{-\xi}\right]. \quad (4)$$

So that replacement of μ in Eq. (3) with Eq. (4) has the desired effect of expressing the GEV model in terms of the parameters (Z_p, σ, ξ).

Taking only the maximum value of each year is apparently a high loss of information. Therefore, if continuous daily records are available, peaks over threshold (POT) method should be preferred to block maxima method. POT approach takes into account all values, which are greater than a given threshold value. For a sufficiently high threshold u, the distribution function of the excess values of x over u converges to the GPD, its CDF is given by:

$$F(x) = 1 - \left(1 + \xi\frac{x-u}{\sigma}\right)^{-\frac{1}{\xi}}, \quad (5)$$

where σ is a scale parameter and ξ is a shape parameter (Lin 2003). The POT model requires the exceedances be mutually independent. However, the fact that extremes may have a tendency to cluster in stationary series and temporal independence is an unrealistic assumption means that some change of practice is needed. In particular, the meteorological variables tend to present successive dependent extreme values, so a technique is need to consider successive extremes as belonging to the same event (de Oliveira et al. 2010).

Figure 1 shows a section of the Zanjan daily maximum temperature series that is approximately stationary. A threshold of 25 °C has also been added. Threshold exceedances are seen to occur in groups, implying that one extremely hot day is likely to be followed by another. The distribution of any one of the threshold excesses might be modeled using the GPD, but the clustering induces dependence in the observations, invalidating the log-likelihood (Coles 2001).

It is necessary to identify clusters of exceedances and to apply the methods of the extreme value analysis to the peak values within each cluster (Kysely et al. 2010). The most widely adopted method for dealing with this issue is declustering (Ceppi et al. 2008), which filters the dependent observations to obtain a set of threshold excesses that are approximately independent. This is done by defining an

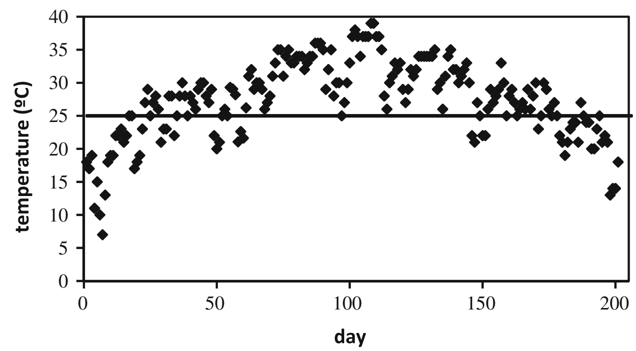


Fig. 1 Portion of Zanjan daily maximum temperature time series

empirical rule to identify clusters of exceedances and also the cluster maxima; these maxima are called declustered peaks and are assumed to be independent (Ceppi et al. 2008). Coles (2001) recommended GPD for fitting such series. Therefore, in the present study, the declustered peaks were fitted to the GPD. Declustering were done by, first, setting a threshold and defining clusters to be wherever there are consecutive exceedances of this threshold. Then setting a run length r (minimum separation) between each cluster; a cluster is terminated whenever the separation between two threshold exceedances is greater than the run length. This should ensure that the declustered peaks are independent from each other.

The major limitation of the GPD in practical works is the selection of the appropriate threshold 'u' for the given data set. Values above the threshold level are considered for fitting the GPD. For this reason, the GPD is very sensitive to the threshold selection (Sanabria and Cechet 2010). The experience suggests that a very high threshold resulting in a small POT sample size would increase the sampling uncertainty (variance) associated with a quantile estimate. On the other hand, as threshold is lowered to include more data, quantile bias tends to increase. In this sense, it is expected that an optimal threshold might exist that would minimize both bias and variance (Caers and Maes 1998; Dougherty and Corotis 1998). Two visual techniques have been developed in recent years to solve this problem. The first one is called The mean residual life plot (MRL) also known as the conditional mean exceedance (CME) method (Lechner et al. 1992); the second one is the model based check (MBC). In the MRL technique, the mean excesses were plotted against threshold u. The locus of points is:

$$\left\{\left(u, \frac{1}{n_u}\sum_{i=1}^{n_u}\left(x_{(i)} - u\right)\right) : u < x_{max}\right\}, \quad (6)$$

where $x_{(l)},\ldots,x_{(n_u)}$ consist of the n_u observations that exceed u, and x_{max} is the largest of the X_i. The linearity of the CME plot can be used as an indicator of the appropriateness of the GPD model. Hence, an appropriate threshold can be chosen by selecting the lowest value above which the CME graph is approximately a straight line (Lin 2003). If a sample follows a GPD, the shape parameter (ξ) and the modified scale parameter ($\delta^* = \delta_u - \xi_u$) remain constant when u increases. In the MBC method, we can draw the maximum likelihood estimates of the modified scale (σ^*) and shape parameters with respect to threshold u and search for "domains of stability" where they remain roughly constant. As we want to be in the asymptotic domain, we are interested in the highest domain of stability. And as we want to have as much information as possible, we choose the lowest threshold of this highest domain (Mazas and Hamm 2011).

Like GEV, MLM was used for estimation of GPD parameters. The log-likelihood function for GPD is:

$$\log L(x;\, \sigma,\, \xi) = -N \log \sigma - (1-\xi)\sum_{i=1}^{N} y_i, \tag{7}$$

where

$$y_i = -\xi^{-1}\log(1 - \xi x_i/\sigma). \tag{8}$$

Quantiles of the GPD distribution are given in terms of parameters by Coles (2001):

$$\hat{x}_m = u + \frac{\sigma}{\xi}\left[(m\zeta_u\theta)^{\xi} - 1\right], \tag{9}$$

where $\hat{x}_m$ is the estimated m-observation return level, ζ_u is the probability of an exceedance of u, and θ is the extremal index, that,

$$\zeta_u = \frac{n_u}{N}, \tag{10}$$

$$\theta = \frac{n_c}{n_u}, \tag{11}$$

where n_u is the number of exceedances of the threshold u, n_c is the number of clusters obtained above u and N is the number of complete observations. Return levels for 2, 5, 10, 25, 50, and 100 years were calculated. Q–Q plots and χ^2 test were used to evaluate the goodness of fit of GEV and GPD. The R software package, extRemes, from the NCAR, was used to perform the extreme modeling following Gilleland and Katz (2005a, 2006).

3 Results and discussion

3.1 GEV

Figure 2 shows the absolute annual values of extremes in Zanjan station. As it can be seen from Fig. 2a, b, the lowest value of daily minimum temperature in Zanjan during the study period is −30 °C and the highest value of daily maximum temperature is 40 °C. Furthermore, it can be seen from Fig. 2c, d, that the highest value of mean daily maximum wind speed is 23 m/s. Also the highest value of daily maximum precipitation is 50.6 mm. GEV was fitted to the absolute annual values of these four extremes.

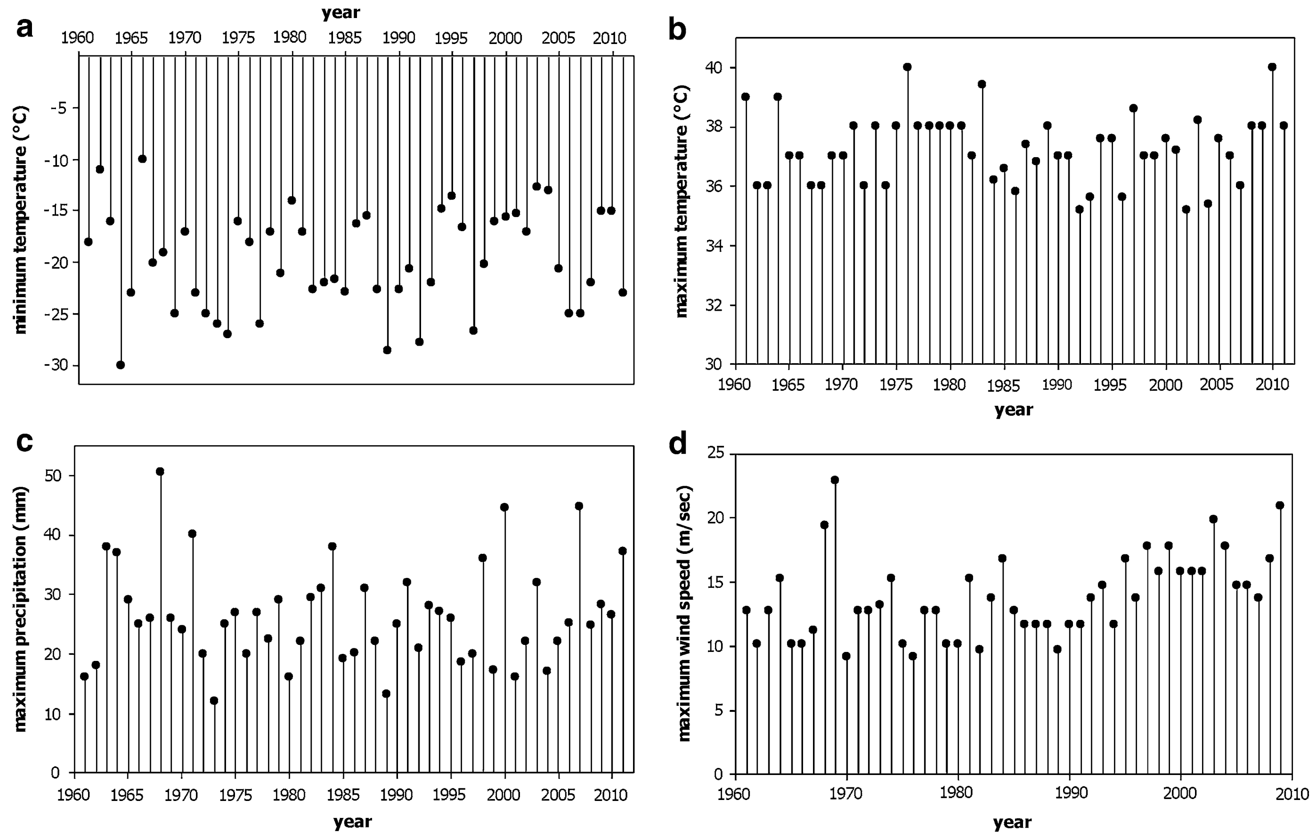
The values of GEV parameters were estimated with log-likelihood method for annual extreme events. For example for annual maximum wind series of Zanjan, maximum likelihood leads to the estimate $\mu = 12.397$, $\sigma = 2.591$ and $\xi = -0.018$. The approximate variance–covariance matrix of the parameter estimates is:

$$V = \begin{bmatrix} 0.18811 & 0.06015 & -0.02594 \\ 0.06015 & 0.10563 & -0.01971 \\ -0.02594 & -0.01971 & 0.01830 \end{bmatrix}.$$

The diagonals of the variance–covariance matrix correspond to the variances of the individual parameters of (μ, σ, ξ). Taking square roots, the standard errors are 0.434, 0.325 and 0.135 for μ, σ and ξ, respectively. Combining estimates and standard errors, 95 % confidence intervals for each parameter were approximated. According to the value of shape parameter (ξ) and its 95 % confidence interval, it can be concluded that for maximum precipitation, Gumbel hypothesis cannot be rejected, because the confidence intervals of ξ include zero. Accordingly the estimated parameters with standard errors for annual minimum and maximum temperature and maximum precipitation series of Zanjan were calculated and are presented in Table 1.

Estimates and confidence intervals for return levels were calculated. For example, to estimate the 10-year return level of maximum wind speed, $p = 1/10$ and find $Z_{0.1} = 18.1$ and $\mathrm{Var}(Z_{0.1}) = 0.510204$, a 95 % confidence interval for $Z_{0.1}$ is evaluated as $18.1 \pm 1.96 \times \sqrt{0.510204} = [16.7,\ 19.5]$. The corresponding estimates for the various return periods with a 95 % confidence interval were calculated. Figure 3 shows return level plots with 95 % confidence bounds for the GEV fit to the annual extreme events at Zanjan station. These plots indicate that most of observations fall within the mentioned bounds.

In order to achieve better estimate of the confidence intervals, we used the profile likelihood. Figures 4, 5, 6, 7 show the profile log-likelihood for the 10 and 100-year return levels of annual extreme events in Zanjan. These figures indicate symmetry for the 10-year return levels (left) and asymmetry for the 100-year return levels (right). Discrepancy between left and right arises because of asymmetry in the profile log-likelihood surface, the extent of which increases with increasing return period, since the data provide increasingly weaker information about high levels of the process. Table 2 represents various return levels for extreme events with confidence intervals in Zanjan.



Figs. 2 Annual absolute maximum value of, **a** minimum temperature, **b** maximum temperature, **c** precipitation and **d** wind speed at Zanjan

Table 1 Estimated parameters of GEV with standard errors in parentheses at Zanjan

Extreme events	μ	σ	ξ
Minimum temperature	−18.14 (0.76)	4.77 (0.56)	−0.30 (0.12)
Maximum temperature	36.79 (0.17)	1.09 (0.12)	−0.20 (0.10)
Maximum precipitation	22.67 (1.07)	6.79 (0.77)	−0.04 (0.11)
Maximum wind speed	12.40 (0.43)	2.59 (0.33)	−0.02 (0.14)

The *Q–Q* plots for assessing the accuracy of the GEV model fitted to the extreme events of Zanjan are shown in Fig. 8. These plots revealed the appropriateness of the GEV model. Each set of plotted points are near-linear. According to Fig. 8, all four plots of extreme events support the appropriateness of the GEV model. Testing the goodness of fit of the GEV for extreme events was done using the χ^2 goodness-of-fit test (Table 5). Results indicated that in all cases the computed χ^2 were less than the corresponding $\chi^2_{n,0.05}$ value extracted from χ^2 table.

3.2 GPD

The GPD distribution uses more information than GEV with the highest or the lowest value in a year. This approach contrasts with the block maxima approach through the characterization of an observation as an extreme if it exceeds a high threshold. Zanjan data cover extreme daily temperatures and precipitations consisting of 18,627 data points from 1961 to 2011 and extreme daily maximum wind speeds consisting of 17,897 data points from 1961 to 2009.

As mentioned earlier, the POT model requires that the exceedances be mutually independent. However, in the case of temperature and wind speed data series of Zanjan, this assumption may be violated because of serial correlation. Hence declustering of temperature and wind speed extremes was done. Then the GPD distribution was fitted to the declustered peaks of daily maximum and minimum temperatures and daily maximum wind speed.

In order to select suitable threshold, two techniques namely MRL plot also known as the CME method and MBC plot were used. Figures 9, 10, 11, 12 show the MRL plots and MBC plots for the Zanjan data series where the dotted lines show the 95 % confidence interval. The MRL plot should be approximately linear for valid values of threshold u. In the MBC technique, the GPD is fitted across a range of different thresholds (u) and then the stability of the modified scale and shape parameters against u is evaluated.

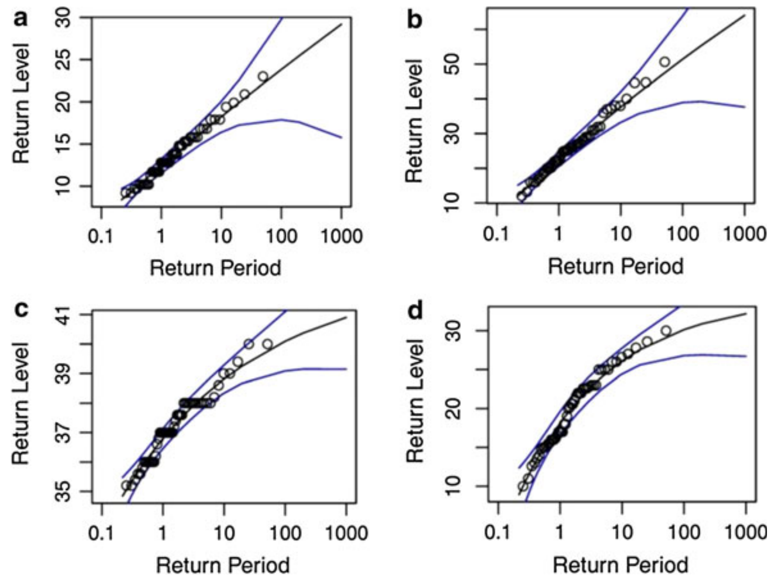


Fig. 3 Return level plots for the GEV fit to the **a** annual maximum wind speed, **b** annual maximum precipitation, **c** annual maximum temperature and **d** annual minimum temperature at Zanjan

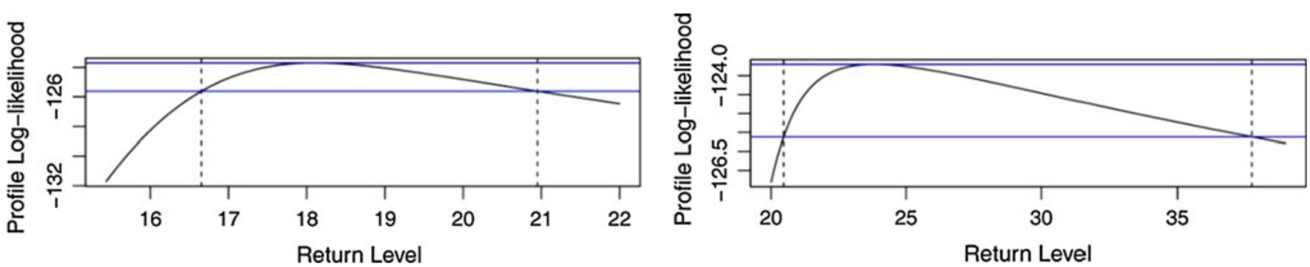


Fig. 4 Profile log-likelihoods of the GEV fit to absolute maximum wind speed at Zanjan for the 10- (*left*), and 100-year return level (*right*). The 95 % confidence interval is marked by *vertical dashed lines*. The *upper horizontal line* indicates the maximum of the profile log-likelihood, while the *lower one* intersects the profile log-likelihood values and the 95 % confidence interval

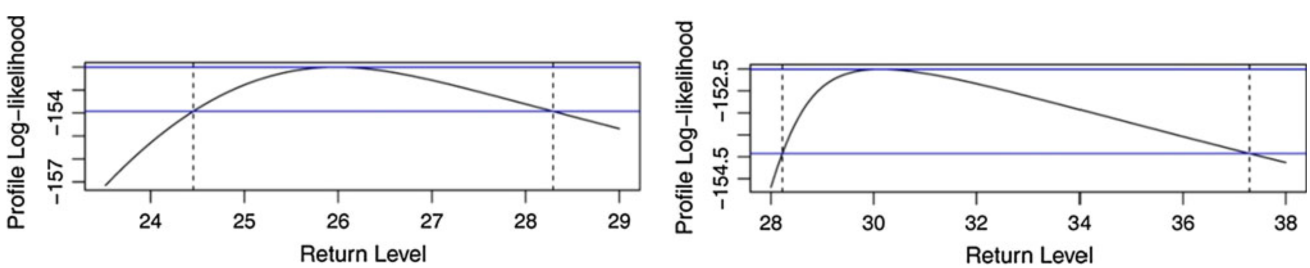


Fig. 5 As Fig. 4 but for minimum temperature

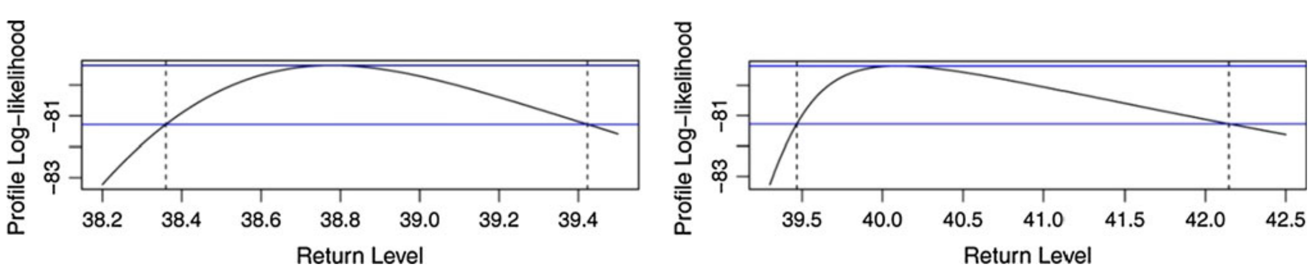


Fig. 6 As Fig. 4 but for maximum temperature

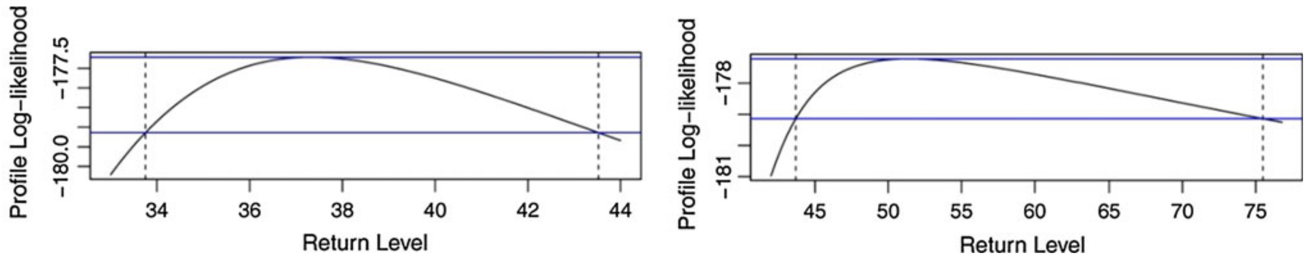
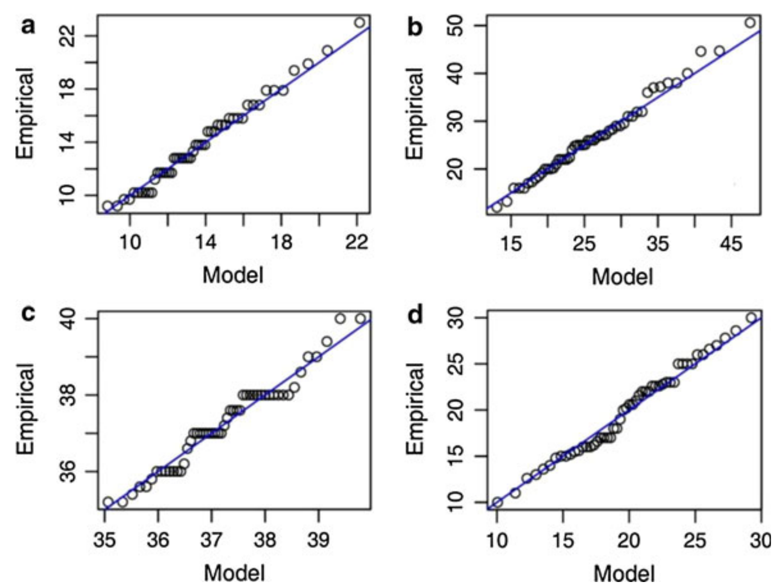


Fig. 7 As Fig. 4 but for maximum precipitation

Table 2 Return levels for annual extreme events with confidence intervals in parenthesis at Zanjan

Return period (years)	Return level			
	Precipitation (mm)	T_{min} (°C)	T_{max} (°C)	Wind (m/s)
2	25.1 (22.9, 27.6)	−19.8 (−18.3, −21.4)	37.2 (36.8, 37.5)	13.3 (12.4, 14.3)
5	32.6 (29.7, 36.4)	−23.9 (−22.4, −25.6)	38.2 (37.8, 38.7)	16.2 (15, 17.8)
10	37.3 (33.7, 43.5)	−26 (−24.5, −28.3)	38.8 (38.4, 39.4)	18.1 (16.6, 20.9)
25	43.1 (38.3, 54.5)	−28 (−26.4, −31.9)	39.4 (38.9, 40.5)	20.4 (18.4, 26.3)
50	47.3 (41.2, 64.3)	−29.2 (−27.4, −34.6)	39.7 (39.2, 41.3)	22.1 (19.5, 31.4)
100	51.4 (43.7, 75.5)	−30.1 (−28.2, −37.3)	40.1 (39.5, 42.1)	23.8 (20.5, 37.7)

Fig. 8 *Q–Q* plots for the GEV fit to the **a** annual maximum wind speed, **b** annual maximum precipitation, **c** annual maximum temperature and **d** annual minimum temperature at Zanjan



According to Figs. 9, 10, 11, 12(right), MBC plots indicate the perturbations of the parameters are small until the chosen thresholds were $u_0^{'} = 25$ °C, $u_0 = -6$ °C, $u_0 = 9$ m/s and $u_0 = 11$ mm for maximum temperature, minimum temperature, maximum wind speed and precipitation, respectively. As it can be seen in Figs. 9, 10, 11, 12(left) the MRL plots clearly show a linear form for thresholds higher than selected threshold u_0. Hence, the selected thresholds appear to be valid, reasonably.

Figure 13 indicates daily extreme events above threshold u in Zanjan during the study period. Autocorrelation in the series was first eliminated by using a standard declustering technique; however no declustering was performed for precipitation extremes, because data series were approximately independent. After declustering maximum likelihood estimates of parameters with standard errors were calculated (Table 3). The shape parameter ξ, which defines the behavior of the tail of the GPD, is negative for minimum and maximum temperature and close to zero for maximum precipitation and maximum wind speed. According to the value of shape parameter (ξ) and its 95 % confidence interval, it can be concluded that for daily maximum wind speed and maximum precipitation, exponential distribution hypothesis cannot be rejected, because

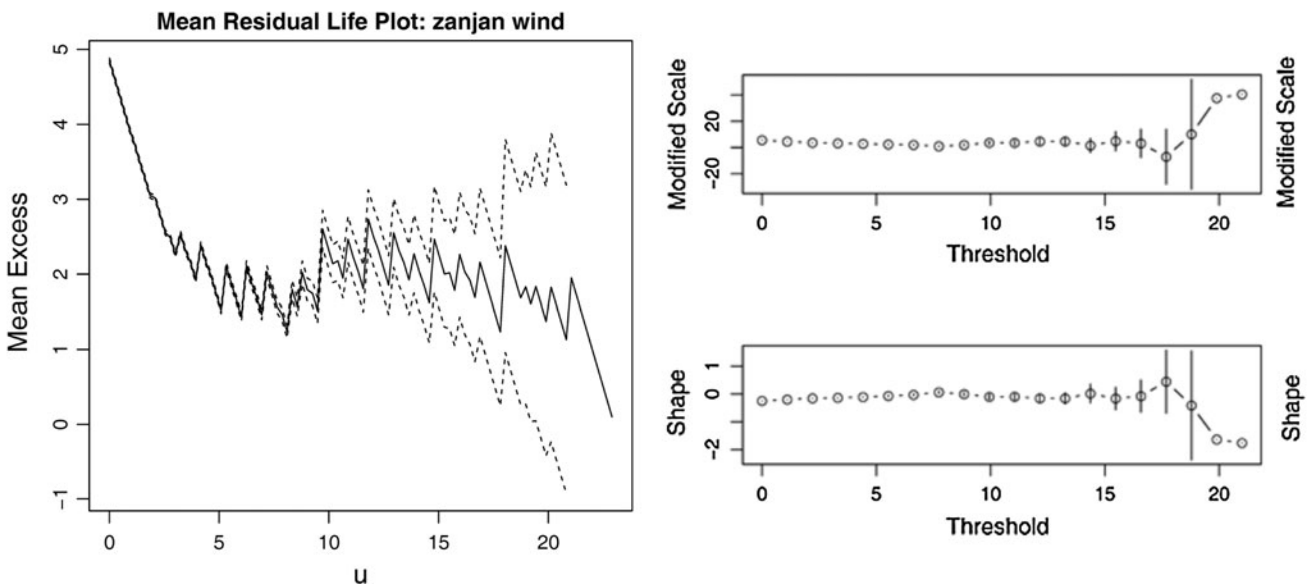


Fig. 9 MRL plots (*left*) and MBC plot (*right*) for maximum daily wind speed

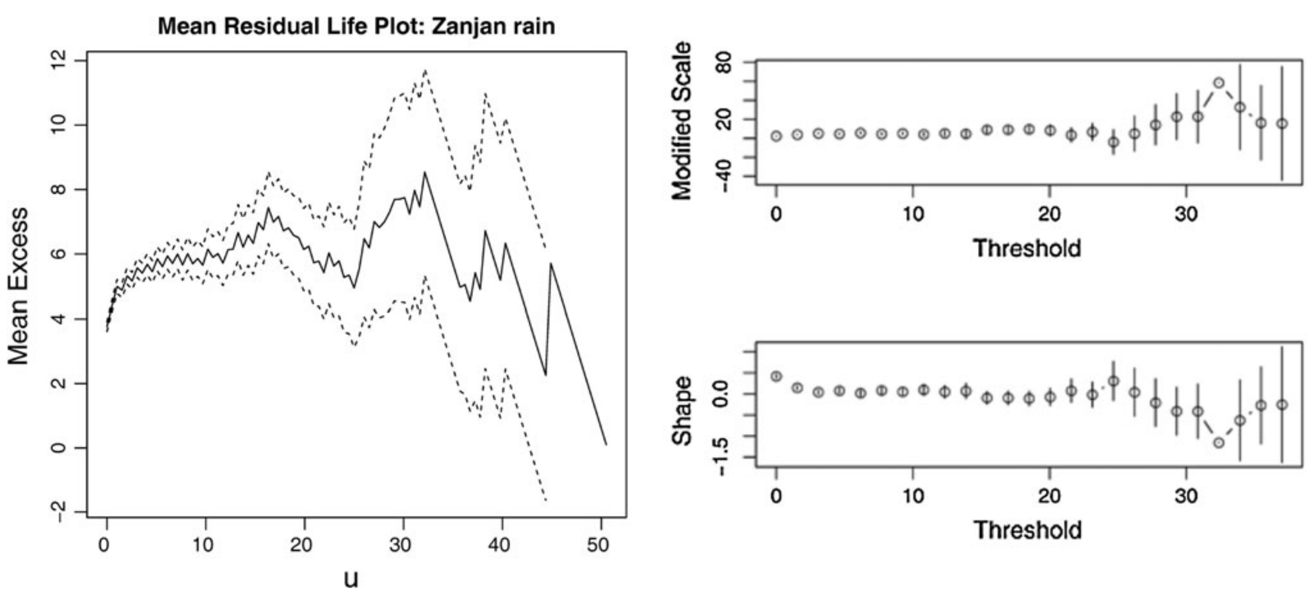


Fig. 10 MRL plots (*left*) and MBC plot (*right*) for daily precipitation

the confidence intervals of ξ include zero. About minimum and maximum temperature, the distribution has a bounded upper tail (Beta type).

Estimates and confidence intervals for various return levels of daily extreme events of Zanjan station were calculated. The return level function, which assigns the daily extreme event level x_m that is, exceeded on average once every m daily observations, was calculated from Eq. (9). For example, in order to estimate the 10-year return level of daily maximum wind speed, we have:

$$x_m = \frac{1.84}{0.06}\left[(365 \times 10 \times 0.035 \times 0.817)^{0.06} - 1\right] = 18.95,$$

with approximate variance:

$$\mathrm{Var}(\zeta_u) = \zeta_u(1 - \zeta_u)/17,897 = 1.89 \times 10^{-6}.$$

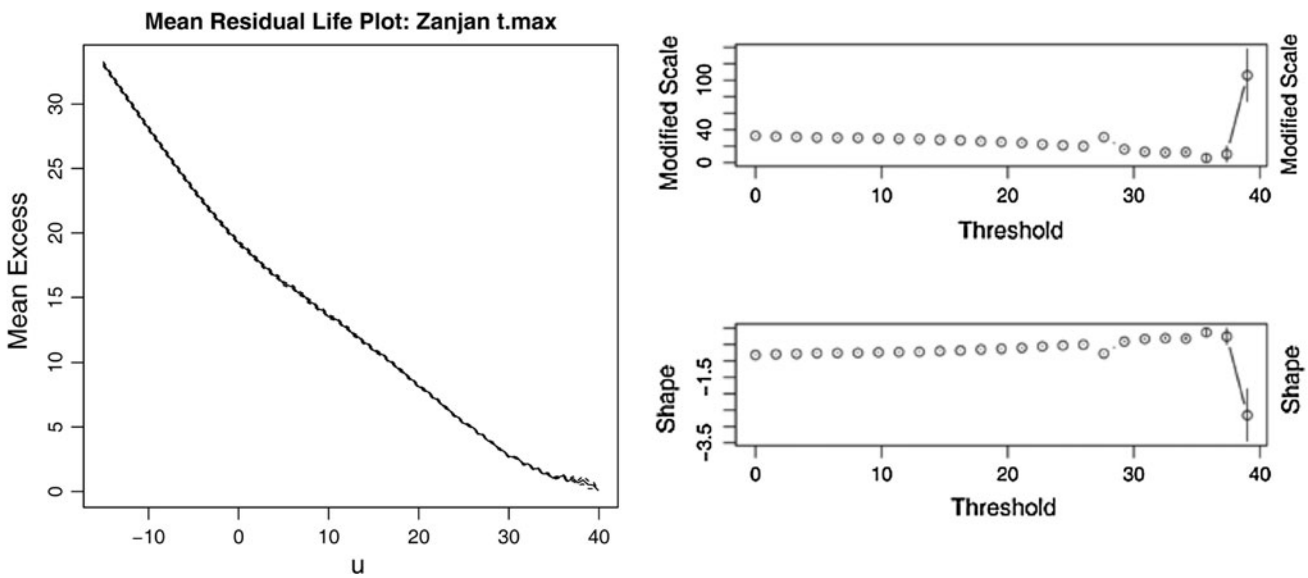


Fig. 11 MRL plots (*left*) and MBC plot (*right*) for maximum daily temperature

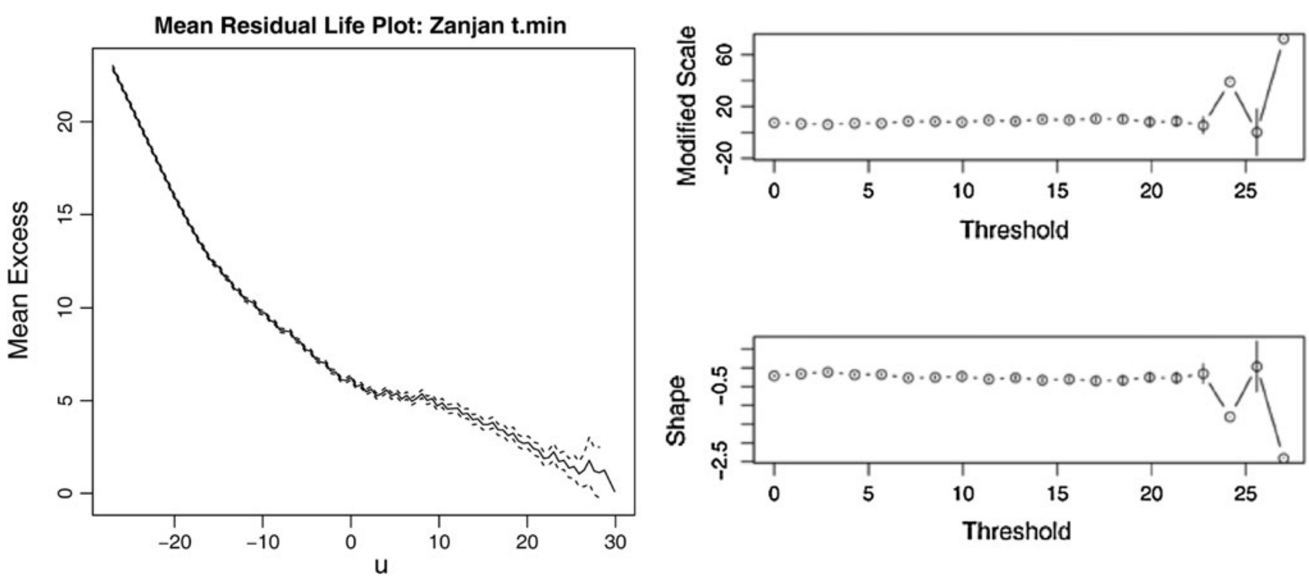


Fig. 12 MRL plots (*left*) and MBC plot (*right*) for minimum daily temperature at Zanjan (minimum temperature is negated)

Hence, the complete variance–covariance matrix for (ζ, σ, ξ) is:

$$V = \begin{bmatrix} 1.89 \times 10^{-6} & 0 & 0 \\ 0 & 0.014390 & -0.004033 \\ 0 & -0.004033 & 0.002316 \end{bmatrix}.$$

So $\mathrm{Var}(X_m) = 0.510204$. Hence, a 95 % confidence interval for X_m is evaluated as $18.95 \pm 1.96 \times \sqrt{0.510204} = [17.6, 20.4]$. The corresponding estimates for various return periods with a 95 % confidence interval were calculated. Return level plots for the fitted GPD to extreme events at Zanjan station are shown in Fig. 14. Return level curves seem to be nonlinear, though in the case of extreme wind speed and extreme precipitation series, the estimated curve expected to be close to linear, because the estimate of ξ is close to zero. The confidence intervals on the return level plot indicate that the model departures are not large, but the very large uncertainties accrue once the model is extrapolated to higher levels. Because the confidence interval substantially increases as the return periods increase indicating a higher degree of uncertainty when making inferences far beyond the range of the data (51 years). The confidence intervals suggest that return periods calculated beyond 1,000 years are too unreliable for use in practical applications (Sanabria and Cechet 2007).

In order to achieve better estimate of the confidence intervals, the profile likelihood was used. Figures 15, 16,

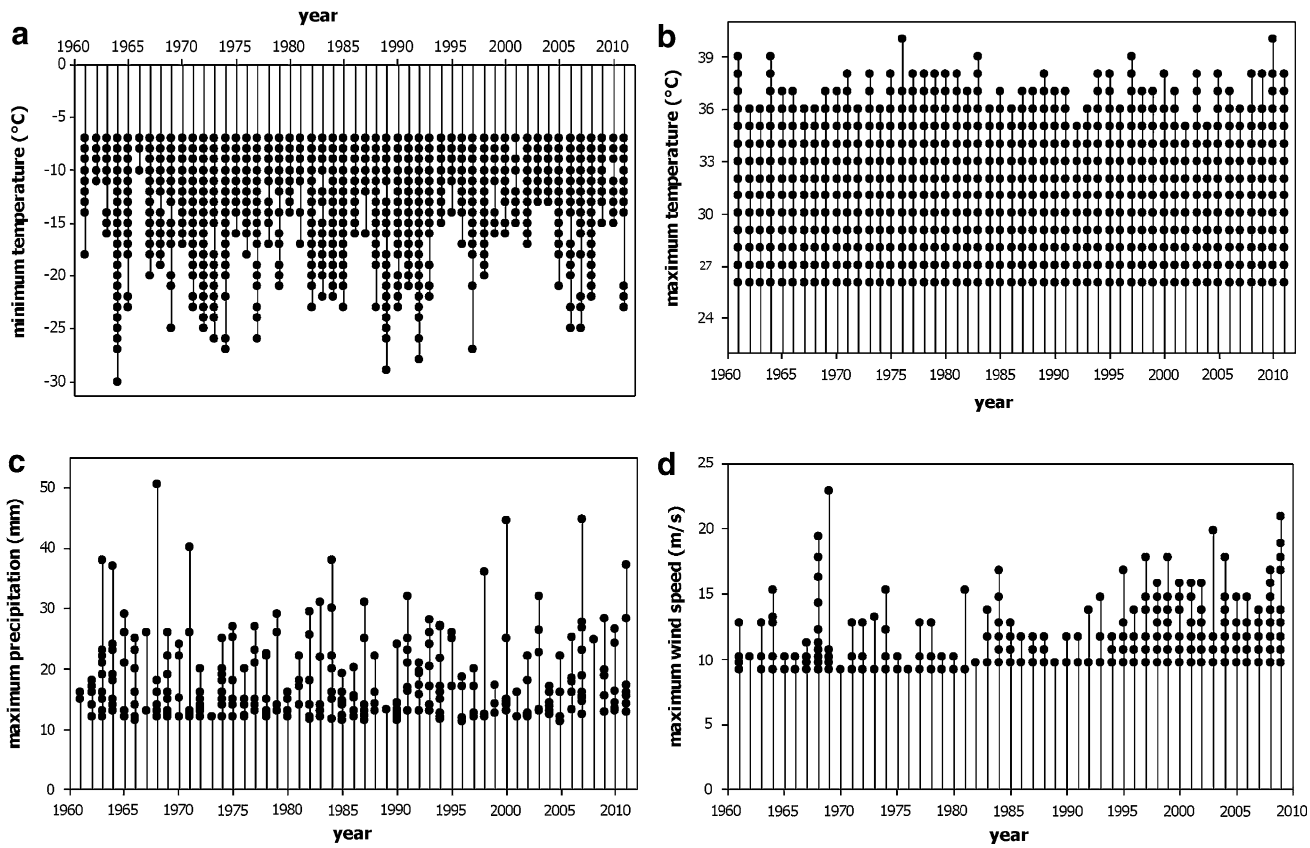


Fig. 13 Daily maximum extreme events above threshold u at Zanjan during the study period. **a** Maximum temperature, **b** minimum temperature, **c** precipitation and **d** maximum wind speed

Table 3 The estimated features of GPD to extreme events at Zanjan

Extreme events	u	r	n_c	ζ_u	θ	σ	ξ
Minimum temperature	−6	1	587	0.113	0.279	6.17 (0.35)	−0.16 (0.04)
Maximum temperature	25	1	563	0.331	0.091	7.29 (0.41)	−0.46 (0.04)
Maximum precipitation	11	–	335	0.018	1	5.91 (0.48)	0.03 (0.06)
Maximum wind speed	9	1	518	0.035	0.817	1.84 (0.12)	0.06 (0.05)

17, 18 show the profile log-likelihood for the 10- and 100-year return levels of daily extreme events in Zanjan. The profiles log-likelihood for the 100-year return levels are more asymmetric than the ones obtained by the GEV fit. Table 4 shows various return levels for extreme events with confidence intervals in Zanjan.

Figure 19 shows Q–Q plots for the GPD fit to the extreme events in Zanjan during the study period. As it can be seen, the Q–Q plots are near-linear that indicate the validity of the fitted model. The result of χ^2 goodness-of-fit test revealed that the GPD fit data acceptable at $\alpha = 0.05$ level (Table 5).

In order to compare the two extreme value analysis methods, the shape parameters of two models were compared. One of the main differences lies in the fact that the shape parameter is less negative in the GPD case (Table 6). In other words, the curve is less concave. This leads to significantly lower return period estimates for high extremes with the GPD model. Table 7 shows a comparison of the return period estimates for the observed highest extreme events at Zanjan station. As it can be inferred from Table 7, the return periods obtained from GEV are much greater than GPD.

Table 8, indicates the values ratio of 100-year return periods to values of 2-year return periods for GEV and GPD models. In this table, it can be seen how much extreme events in 100-year return periods, would be larger in comparison for 2-year return periods. According to

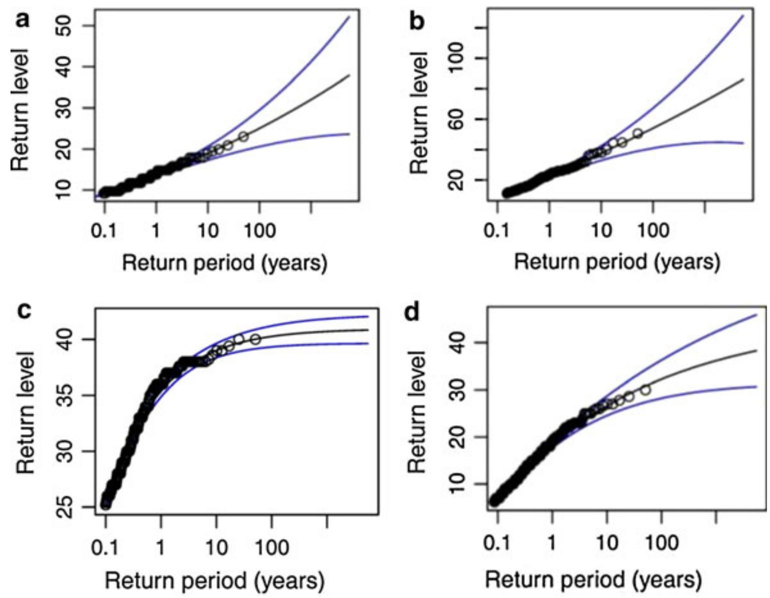


Fig. 14 Return level plots for the GPD fit to the **a** daily maximum wind speed, **b** daily precipitation, **c** daily maximum temperature and **d** daily minimum temperature at Zanjan

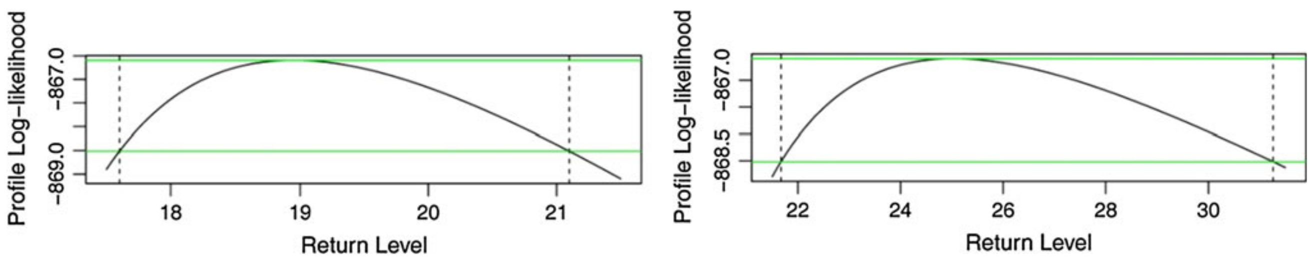


Fig. 15 Profile log-likelihoods of the GPD fit to daily maximum wind speed at Zanjan for the: 10- (*left*) and 100-year return level (*right*). The 95 % confidence interval is marked by *vertical dashed lines*. The *upper horizontal line* indicates the maximum of the profile log-likelihood, while the *lower one* intersects the profile log-likelihood values and the 95 % confidence interval

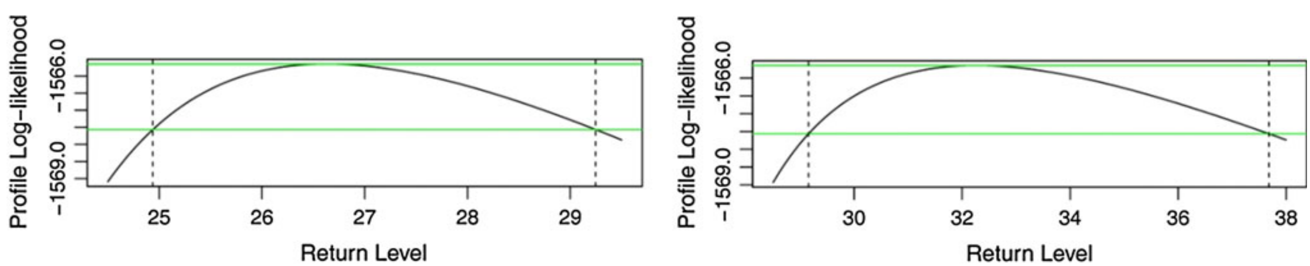


Fig. 16 As Fig. 15 but for daily minimum temperature

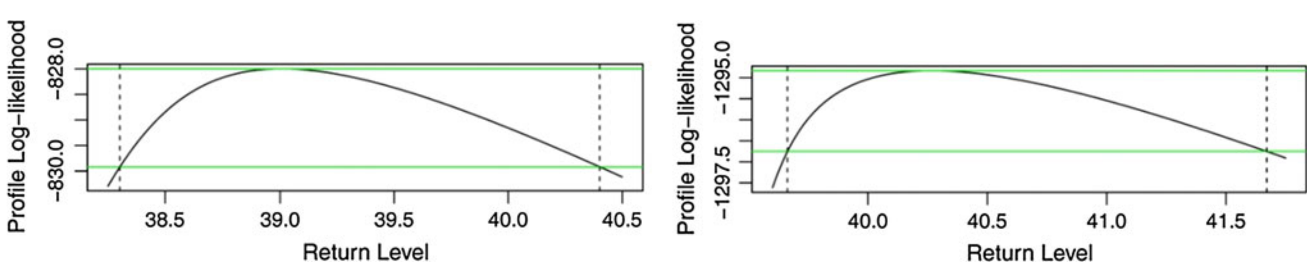


Fig. 17 As Fig. 15 but for daily maximum temperature

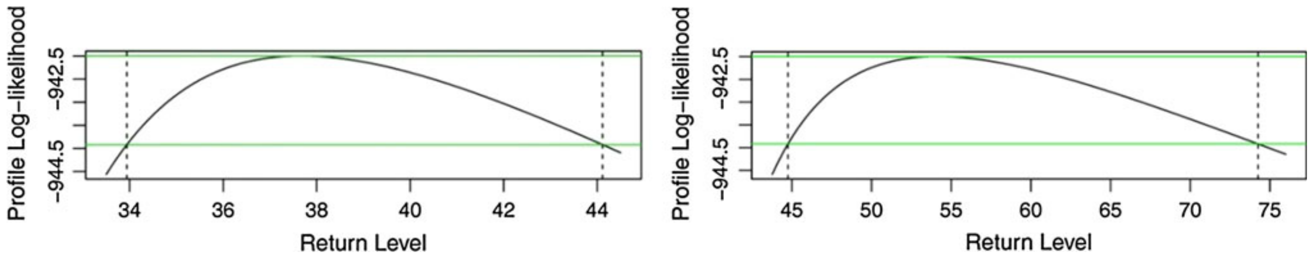


Fig. 18 As Fig. 15 but for daily maximum precipitation

Table 4 Return levels for daily extreme events with confidence intervals at Zanjan

Return period (years)	Return level			
	Precipitation (mm)	T_{min} (°C)	T_{max} (°C)	Wind (m/s)
2	26.9 (25.3, 29.1)	−21.3 (−20.3, −22.5)	36.9 (36.4, 37.4)	15.2 (14.6, 16)
5	33 (30.3, 37.1)	−24.5 (−23.1, −26.4)	38.3 (37.8, 39)	17.3 (16.3, 18.7)
10	37.7 (33.9, 44.1)	−26.6 (−24.9, −29.2)	39 (38.5, 39.9)	19 (17.6, 21.1)
25	44.1 (38.5, 54.7)	−29.1 (−26.9, −32.7)	39.7 (39.1, 40.8)	21.3 (19.3, 24.7)
50	49 (41.7, 63.9)	−30.8 (−28.1, −35.3)	40 (39.4, 41.3)	23.1 (20.5, 27.8)
100	54.1 (44.8, 74.3)	−32.3 (−29.2, −37.7)	40.3 (39.7, 41.7)	25 (21.7, 31.3)

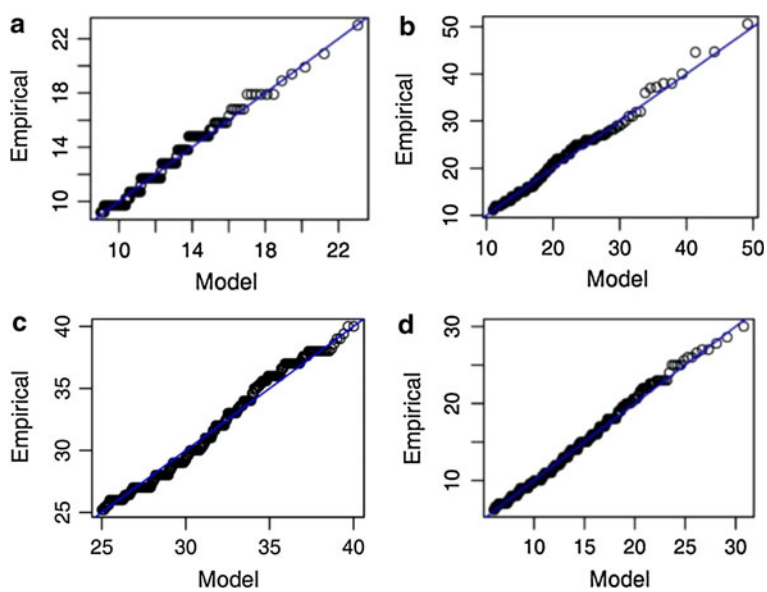


Fig. 19 *Q–Q* plots for the GPD fit to the **a** daily maximum wind speed, **b** daily precipitation, **c** daily maximum temperature and **d** daily minimum temperature at Zanjan

Table 5 The resulting values of χ^2 test for GEV and GPD models (critical value of χ^2 is 3.84 for $\alpha = 0.05$ at 1 df)

Extreme events	GEV	GPD
Minimum temperature	2.63	1.34
Maximum temperature	3.21	2.83
Maximum precipitation	0.37	0.11
Maximum wind speed	1.84	0.018

Table 6 Comparison of the ξ parameters of the GPD and GEV models fitted to extreme events at Zanjan

Extreme events	GEV	GPD
Minimum temperature	−0.30	−0.16
Maximum temperature	−0.20	−0.46
Maximum precipitation	−0.04	0.03
Maximum wind speed	−0.02	0.06

Table 7 Comparison of the return periods of peak extreme events obtained from the GPD and the GEV model at Zanjan

Extreme events and dates	Return level	Return period (years)	
		GEV	GPD
Minimum temperature (12 January 1964)	−30 °C	90	36
Maximum temperature (30/31 July 1976) (9 July 2010)	40 °C	83	50
Maximum precipitation (25 May 1968)	50.6 mm	87	62
Maximum wind speed (1 April 1969)	23 m/s	71	48

Table 8 Ratio of Trs 100–2 for GEV and GPD fitting to extreme events at Zanjan

Extreme events	GEV	GPD
Minimum temperature	1.52	0.66
Maximum temperature	1.08	1.05
Maximum precipitation	2.04	0.55
Maximum wind speed	1.79	0.61

Table 8, the reciprocal values for GEV were greater than those of the GPD.

4 Conclusions

In this paper the climate extreme events consisting of maximum temperature, minimum temperature, maximum precipitation and maximum wind speed in Zanjan were analyzed. GEV and GPD were fitted to the absolute maximum annual values, and maximum daily values of extreme events, respectively.

As the GPD model requires that the exceedances be independent, declustering was performed to obtain a set of threshold excesses that are approximately independent. The GPD was then fitted to the declustered peaks. GEV fitting showed that the Weibull distribution was more suitable for precipitation and wind speed extremes, whereas the Gumbel distribution appeared to be a better model for temperature extremes in Zanjan. The best fitting GPD was found to be the beta distribution for temperature extremes, and exponential distribution for precipitation and wind speed extremes in Zanjan.

Return levels of extreme values for 100 years for GEV distribution yielded values closer to the maxima already recorded, than the GPD, indicating rare information used in GEV model. The shape parameter was less negative in the GPD than the GEV. This leads to significantly lower return period estimates for high extremes with the GPD model. The Q–Q plots and χ^2 goodness-of-fit test indicated the validity of fitted models. GPD fits appear more satisfactory than those for the GEV, providing better results for the return levels of the return period of 100 years. Because the profiles log-likelihood obtained by GPD for the 100-year return levels were more symmetric than the ones obtained by the GEV fit. Also this was confirmed by the Q–Q plots and the results of χ^2 goodness-of-fit test. In other words, it is possible to obtain reliable station-based climate extreme values.

References

Abaurrea J, Asín J, Cebrián AC, Centelles A (2007) Modeling and forecasting extreme hot events in the central Ebro Valley, a continental-Mediterranean area. Glob Planet Change 57:43–58

Asakereh H (2012) Frequency distribution change of extreme precipitation in Zanjan City. Geogr Environ Plan J, 23rd Year, 45(1):13-17

Beirlant J, Goegebeur Y, Segers J, Teugels J (2004) Statistics of extremes: theory and applications. Wiley series in probability and statistics. Chichester

Caers J, Maes MA (1998) Identifying tail bounds and end points of random variables. Struct Saf 20:1–23

Ceppi P, Della-Marta PM, Appenzeller C (2008) Extreme value analysis of wind speed observations over Switzerland, Arbeitsberichte der MeteoSchweiz, No: 219. Bundesamt für Meteorologie und Klimatologie, MeteoSchweiz

Coles SG (2001) An introduction to statistical extreme values. Springer, London

de Oliveira MMF, Ebecken NFF, de Oliveira JLF, Gilleland E (2010) Generalized extreme wind speed distributions in South America over the Atlantic Ocean region. Theor Appl Climatol 104:377–385. doi:10.1007/s00704-010-0350-3

Dougherty AM, Corotis RB (1998) Extreme wind estimation: theoretical considerations. In: Proceedings of the 7th international conference on structural safety and reliability

Gilleland E, Katz RW (2005a) Extremes toolkit (extRemes): weather and climate applications of extreme value statistics. http://www.isse.ucar.edu/extremevalues/tutorial.pdf

Gilleland E, Katz RW (2005b) Analysing seasonal to interannual extreme weather and climate variability with the extremes toolkit. National Center for Atmospheric Research (NCAR), Boulder

Gilleland E, Katz RW (2006) Analyzing seasonal to interannual extreme weather and climate variability with the extremes toolkit. Research Applications Laboratory National Center for Atmospheric Research (NCAR), Boulder

Hashmi MZ, Shamseldin AY, Melville BW (2011) Comparison of SDSM and LARS-WG for simulation and downscaling of extreme precipitation events in a watershed. Stoch Environ Res Risk Assess 25(4):475–484

Holmes JD, Moriarty WW (1999) Application of the generalized Pareto distribution to extreme value analysis in wind engineering. J Wind Eng Ind Aerodyn 83:1–10

Hosking JRM, Wallis JR, Wood EF (1985) Estimation of the general extreme value distribution by the method of probability weighted moments. Technometrics 27(3):251–261

Katz RW, Parlange MB, Naveau P (2002) Statistics of extremes in hydrology. Adv Water Resour 25:1287–1304

Kharin VV, Zwiers FW (2000) Changes in the extremes in an ensemble of transient climate simulations with a coupled atmosphere–ocean GCM. J Clim 13:3760–3788

Kharin VV, Zwiers FW (2005) Estimating extremes in transient climate change simulations. J Clim 18:1156–1173

Kysely J, Picek J, Beranova R (2010) Estimating extremes in climate change simulations using the peaks-over-threshold method with a non-stationary threshold, Global and Planetary Change 72:55–68

Laurent C, Parey S (2007) Estimation of 100-year-return-period temperatures in France in a non-stationary climate: results from observations and IPCC scenarios. Glob Planet Change 57(1–2): 177–188

Lechner JA, Leigh SD, Simiu E (1992) Recent approaches to extreme value estimation with application to wind speeds. Part I: the Pickands Method. J Wind Eng Ind Aerodyn 41–44:509–519

Lin XG (2003) Statistical modeling of severe wind gust. In: International congress on modelling and simulation, Townsville, 14–17 July, vol 2, pp. 620–625

Mazas F, Hamm L (2011) A multi-distribution approach to POT methods for determining extreme wave heights. Coast Eng 58: 385–394

Naveau P, Nogaj M, Ammann C, Yiou P, Cooley D, Jomelli V (2005) Statistical methods for the analysis of climate extremes. C R Geosci 337:1013–1022

Nemeth C (2011) Extreme value theory project. Statistics and Operational Research Doctoral Training Centre (STOR-i DTC), Lancaster University, Lancaster

Parey S, Malek F, Laurent C, Dacunha-Castelle D (2007) Trend and climate evaluation: statistical approach for very high temperatures in France. Clim Change 81:331–352

Payer T, Kuchenhoff H (2004) Modelling extreme wind speeds at a German weather station as basic input for a subsequent risk analysis for high-speed trains. J Wind Eng Ind Aerodyn 92:241–261

Rao AR, Hamed KH (2000) Flood frequency analysis. CRC Press, Boca Raton

Reiss R, Thomas M (2001) Statistical analysis of extreme values from insurance, finance, hydrology and other fields. Birkhauser, New York

Renald B, Lang M, Bois P (2006) Statistical analysis of extreme events in a non-stationary context via a Bayesian framework: case study with peak-over-threshold data. Stoch Environ Res Risk Assess 21(2):97–112

Sanabria LA, Cechet RP (2007) A statistical model of severe winds. Geoscience Australia, Record 2007/12. ISSN 1448-2177, ISBN 978 1 921236 43 3

Sanabria LA, Cechet RP (2010) Extreme value analysis for gridded data. In: Swayne DA, Yang W, Voinov AA, Rizzoli A, Filatova T (eds) International Environmental Modelling and Software Society (iEMSs) 2010 international congress on environmental modelling and software modelling for environment's sake, fifth biennial meeting, Ottawa, Canada. http://www.iemss.org/iemss2010/papers/S28/S.28.02.Extreme%20value%20analysis%20for%20gridded%20data%20-%20L.A.%20SANABRIA.pdf

Smith RL (2003) Statistics of extremes, with applications in environment, insurance and finance. In: Finkenstadt B, Rootzen H (eds) Chapter 1 of extreme values in finance, telecommunications and the environment, Chapman, Hall/CRC Press, London, pp 1–78. http://www.stat.unc.edu/postscript/rs/semstatrls.pdf

Taghavi F, Mohammadi H (2007) Study the return period of extreme climate events for reduction of environmental impacts. J Environ Stud 33(43):11–20

Unkasevic M, Tosic I (2009) Changes in extreme daily winter and summer temperatures in Belgrade. Theor Appl Climatol 95:27–38. doi:10.1007/s00704-007-0364-7

Van Den Brink HW, Konnen GP, Opsteegh JD (2004) Statistics of extreme synoptic-scale wind speeds in ensemble simulations of current and future climate. J Clim 17:4564–4574