MAUSAM, **72**, 3 (July 2021), 607-618

551.524.36 : 551.524 : 519.226

A bayesian analysis of the annual maximum temperature using generalized extreme value distribution

CHERAITIA HASSEN

Department of Mathematics, University of Mohamed Seddik Ben Yahia, Jijel, Algeria

(*Received 6 August 2020, Accepted 15 June 2021*)

e mail : hassen.stat@hotmail.fr

सार – जिजेल मौसम स्टेशन के लिए सामान्यीकृत चरम मान (GEV) वितरण का उपयोग करके वार्षिक अधिकतम तापमान का मॉडल तैयार किया गया था। मान-केंडल (MK) और Kwiatkowski Philips, Schmidt and Shin (KPSS) के परीक्षण स्थानीय पैरामीटर में रैखिक प्रवृत्ति के बिना एक स्थिर मॉडल का सुझाव देते हैं। कर्टोसिस और वैषम्य आंकड़ों से पता चला कि सामान्यता की धारणा को खारिज कर दिया गया था। सर्वोत्तम मॉडल निर्धारित करने के लिए संभावना अनुपात परीक्षण का उपयोग किया गया था और गुडनेस-ऑफ-फिट परीक्षणों से पता चला है कि हमारा डेटा स्थिर गंबेल वितरण के साथ उपयुक्त है। मार्कोव चेन्स (MCMC) द्वारा मोंटे कार्लो पद्धति का उपयोग करते हुए अधिकतम संभावना अनुमान पद्धति और बायेसियन दृष्टिकोण का उपयोग गंबेल वितरण के मापदंडों को खोजने के लिए किया गया और विभिन्न अवधियों के लिए वापसी स्तर प्राप्त किए गए ।

जेईएल वर्गीकरण: C1, C13, C46, C490.

ABSTRACT. The annual maximum temperature was modeled using the Generalized Extreme Value (GEV) distribution to Jijel weather station. The Mann-Kendall (MK) and Kwiatkowski Phillips, Schmidt and Shin (KPSS) tests suggest a stationary model without linear trend in the location parameter. The Kurtosis and the Skewness statistics indicated that the normality assumption was rejected. The Likelihood Ratio test was used to determine the best model and the goodness-of-fit tests showed that our data is suitable with a stationary Gumbel distribution. The Maximum Likelihood estimation method and the Bayesian approach using the Monte Carlo method by Markov Chains (MCMC) were used to find the parameters of the Gumbel distribution and the return levels were obtained for different periods.

JEL Classification: C1, C13, C46, C490.

Key words – Generalized Extreme Value (GEV) distribution, Gumbel distribution, Maximum Likelihood estimate (ML), Markov Chains Monte-Carlo method (MCMC), Maximum temperature, Return level.

1. Introduction

The Extreme Value Theory (EVT) provides a statistical framework to make inferences about the probability of very rare and extreme events. It is based on the analysis of the maximum (or minimum) value in a selected period. The modern theory of extreme value was developed between 1920 and 1940 by Fréchet (1927), Fisher and Tippet (1928); Gnedenko (1943) and Gumbel (1958). It finds application in many fields; for instance, it has been used to estimate the extreme levels of a river causing floods in hydrology: Coles and Tawn (2005); Renard *et al*. (2006) in France. Concerning precipitations and the risk of floods caused by this phenomenon, a considerable number of studies aiming to model such events can be found in the literature: Coles (2001); Friederichs (2010) in Germany, Kim *et al*. (2009) in the south of Korea, Boudrissa *et al*. (2017) in Algeria, Delson and Retius (2015) in Zimbabwe and Deka *et al*. (2011) in India. Other studies related to the risk of high temperature: Tadele *et al*. (2015) in South Africa and Husna Hasan *et al*. (2013) in Malaysia.

Different methods can be used to estimate the parameters of the extreme laws such as the Maximum Likelihood method [Coles (2001); Smith (1999)], the Weighted Moments method [(Hosking (1990); Greenwood *et al*. (1979)] and semi-parametric methods [(Pickands (1975); Hill (1975)]. The Bayesian approach offers another estimation methodology. This method is based on the processing of the unknown parameters as random variables and then obtains appropriate estimators using Markov Chains Monte Carlo methods (MCMC). Bayesian analysis is becoming increasingly popular in many fields including meteorology and has shown its practical benefits in several applied problems. As an

example, Coles and Tawn (1996) adopted a Bayesian methodology to model the oceanographic flood events on the UK east coast, Khodja *et al.* (1998) applied a multivariate test for a shift in the mean of rainfall data of different stations in West Africa with a Bayesian argument, Lu and Berliner (1999) considered a class of Bayesian dynamic models that involve switching among various regimes for a runoff time series, Perreault *et al.* (1999) presented a Bayesian approach which can be used to study a change in the mean level of a set of independent normal random variables, Thyer and Kuczera (2000) developed a hidden state Markov (HSM) model as a new approach for generating hydro-climatic time series, Wang (2001) used a Bayesian approach to estimate the parameter of a bivariate generalized extreme value distribution, Thyer *et al.* (2002) used a Metropolis Hasting algorithm to simulate the posterior distribution of a Box-Cox transformation, Coles *et al.* (2003) adopted a Bayesian approach to model a set of daily rainfall extremes data on the central coast of Venezuela, Parent and Bernier (2003) developed an inference procedure for the peak over threshold (POT) model using semi-conjugate informative priors, Renard *et al.* (2006) presented a Bayesian application to the regional frequency analysis of extremes in a nonstationary context. Coles and Powell (1996) reviewed the literature linking the themes of Bayesian and extreme value analysis, Smith (1999) compared the Bayesian and the frequentist approaches to study parametric predictive inference. The zyp (Bronaugh and Werner, 2013), evd (Stephenson, 2002), extRemes (Gilleland and Katz, 2011) and ismev (Stephenson, 2014) packages of R (R core Team, 2015) were used for the data analysis.

Jijel City (Fig. 1) is located in the northeastern region of Algeria with an area of 2396.63 km^2. It is bordered from the east by Skikda city, from the West by Bejaia city, from the North by the Mediterranean Sea and from the South by Setif and Mila cities. Jijel has a Mediterranean climate characterized by mild and rainy winters with annual precipitations of 814 mm, hot and sunny summers. The annual average temperature in Jijel is around 18 °C.

This study aims to use the block maxima approach by GEV distribution to model annual maximum temperature in Northern Algeria and tries to understand the changes in temperature caused by global warming which affect Algeria's weather specifically Jijel city. Our objective is to investigate the temperature changes to forecast its extremely high levels that can occur in Jijel city and that authorities consider it to prepare a precautionary system against the catastrophes owing to high temperatures such as forest fires.

In addition to this introduction, the paper is organized as follows: the GEV distribution, the Maximum Likelihood estimates of its parameters, the Bayesian estimate and the return level are presented in Section 2. The theoretical model is applied to data in Section 3. Finally, some conclusions are given in Section 4.

2. Methodology

2.1. *Generalized Extreme Value (GEV) distribution*

The extreme value theorem provides a theoretical framework to model the distribution of extreme events and the three-parameter GEV is recommended for meteorology frequency analysis. The three parameters are location, scale and shape. The GEV distribution is a family of continuous probability distributions developed within the extreme value theorem. The GEV distribution arises from the limit theorem of Fisher and Tippet (1928) and Gnedenko (1943).

Supposing that $X_1, X_2,\ldots, X_N$ is a sequence of independent random variables from a common distribution function F. The order statistic M_n = Max $(X_1, X_2,\ldots, X_N)$ is the maximal value of the independent identically distributed (IID) random variables. The external type theorem states that if there exist normalizing constants ($a_n > 0$) and $b_n \in R$ such that:

$$P\left(\frac{M_n - b_n}{a_n} \le x\right) = P\left(M_n \le a_n x + b_n\right) \rightarrow F^n\left(a_n x + b_n\right) \rightarrow G(x) \text{ as } n \rightarrow \infty$$

where, G is a non-degenerate distribution function. The distribution of G belongs to the Gumbel, Frechet, or Weibull distribution. The combination of these three distribution families into a single-family model forms the GEV distribution as proposed by Jenkinson (1955) with cumulative function:

$$G(x,\mu,\sigma,\xi) = \begin{cases} \exp\left[-\left\{1+\xi\left(\frac{x-\mu}{\sigma}\right)\right\}^{-1/\xi}\right], & \text{for } \xi \neq 0,\ 1+\xi(x-\mu)/\sigma > 0 \\ \exp\left[-\exp\left(-\frac{x-\mu}{\sigma}\right)\right], & \text{for } \xi = 0 \end{cases} \quad (1)$$

$\mu, \sigma > 0$ and ξ are location, scale and shape parameters, respectively.

By derivation of the distribution function specified in expression (1), the density function is given by:

$$g(x,\mu,\sigma,\xi)=\begin{cases}\left[1+\xi\left(\frac{x-\mu}{\sigma}\right)^{-\frac{1}{\xi}-1}\right]\exp\left[-\left\{1+\xi\left(\frac{x-\mu}{\sigma}\right)\right\}^{-\frac{1}{\xi}}\right], & si\ \xi\neq 0\\ \exp\left[-\frac{1}{\sigma}\left(\frac{x-\mu}{\sigma}\right)+\exp\left(-\frac{x-\mu}{\sigma}\right)\right], & si\ \xi=0\end{cases}$$

The classical GEV; $G(x, \mu, \sigma, \xi)$ model assumes that the three parameters of location, scale and shape are time-independent. However, if trends are detected in the data sample, the non-stationary case where parameters are expressed as covariates (*e.g*., time) should be considered.

2.2. *Selection Period, model choice and Stationary Test*

The data of extreme value are grouped into blocks of equal length *n*. The maximum of each block forms a series of block maxima M_1, M_2, M_3…, M_n to be fitted with the GEV distribution. The selection of block size is critical as a too-small block can lead to a bias and if too large leading to a large estimation variance (Husna *et al*., 2018). In this study, data of maximum temperature are divided into annual blocks. Three models are considered : a stationary model $M_1 = GEV_1(\mu, \sigma, \xi)$ with constant parameters μ, σ and ξ, a non-stationary model $M_2 = GEV_2[\mu(t), \sigma, \xi)]$ with time as covariates in the location parameter : $\mu_t = \mu_0 + \mu_1 t$, ξ and σ are constants and stationary Gumbel model $M_0 = GEV_0(\mu, \sigma, 0)$ with $\xi = 0$ and μ, σ are constants. The stationarity assumption must be checked before the GEV model is fitted. The following tests are proposed:

2.2.1. *Mann Kendall test*

The non-parametric Mann-Kendall test is used to determine if the values of a random variable follow a monotonic trend (Ryden, 2011). The null hypothesis states that no trend exists. This test does not conform to any particular distribution and is particularly useful if datasets have missing values (Husna *et al*., 2018).

2.2.2. *Kwiatkowski Phillips, Schmidt and Shin (KPSS) test*

The KPSS test is applied to the data to check the stationarity. The null hypothesis of the KPSS test is that there is no trend, while the alternative hypothesis is that there is a trend in the data.

2.3. *Model diagnostic and validation tools*

The goodness-of-fit of the model is assessed by the following methods:

2.3.1. *The Likelihood Ratio test*

The Likelihood Ratio (LR) test is used to compare the fit of the null model M_0 and the alternative model M_1 (M_0 is a special case of M_1). The LR test statistic defined as: $\lambda = -2\ln\left(\frac{L_0}{L_1}\right)$.

where, L_0 and L_1 represent the Likelihood of the model M_0 and M_1 respectively. λ has a Chi-Square distribution with one degree of freedom (Boudrissa *et al*., 2017).

The null model (M_0) is preferred if $\lambda \leq \chi^2_{1,0.95} = 3.8415$. Otherwise, the alternative model (M_1) is more suitable.

2.3.2. *Quantile-Quantile plot (Q-Q)*

The (Q-Q) plot assesses the adequacy of a fitted distribution by comparing the $(1/n + 1)^{th}$ quantiles deriving from the theoretical and empirical distributions. If the two distributions being compared are similar, the points in the Q-Q plot will approximately lie on the diagonal line.

2.3.3. *The goodness of fit tests*

These tests [Kolmogorov Smirnov (KS), Anderson Darling (AD) and Cramer Von Mises (CVM)] are often used for comparing an empirical distribution determined from a sample $F_n(x)$ to a theoretical distribution $F(x)$. The null hypothesis is $F_n(x) = F(x)$ and the alternative hypothesis is $F_n(x) \neq F(x)$.

The distributions of their statistics are the subject of the statistical tables (KS, AD and CVM tables), which take into account the sample size and the accepted error α. It is then sufficient to compare the observed value with the appropriate value in the table.

2.4. *Parameters estimates*

Several methods have been introduced in the literature to estimate the parameters of the GEV distribution; for example, the method of Moments by Christopeit (1994), the L-Moments method (Hosking, 1990; Hosking and Wallis, 1997), for more details about this method see Hosking, 1990), the Bayesian method by

Smith and Naylor (1987); Lye *et al*. (1993); Coles and Tawn (2005) and the Maximum Likelihood method (Smith and Naylor, 1987) which is the most popular and has the advantage of allowing the addition to the fitting of co-variables such as trends, cycles, or physical variables (Katz and *et al*., 2002).

2.4.1. *Maximum Likelihood method (ML)*

Under the assumption that $X_1, \ldots, X_n$ are independent random samples having a GEV distribution, the log-Likelihood for the GEV parameters when $\xi \neq 0$ is:

$$l(x,\mu_t,\sigma,\xi) = -n\log(\sigma) - \sum_{t=1}^{n}\left[1-\xi\left(\frac{x_t-\mu_t}{\sigma}\right)\right]^{\frac{1}{\xi}} - \left(1+\frac{1}{\xi}\right)\sum_{t=1}^{n}\log\left[1+\xi\left(\frac{x_t-\mu_t}{\sigma}\right)\right]$$

where, $\mu_t = \mu$ for stationary three parameters model M_1 = GEV (μ, σ, ξ) and $\mu_t = \mu_0 + \mu_1 t$ for non-stationary four parameters model M_2 = GEV [μ(*t*), σ, ξ].

We differentiated the log-likelihood of GEV to find a set of equations which we solved using numerical optimization algorithms. In the case ξ > -0.5, the usual properties of consistency, asymptotic efficiency and asymptotic normality hold (Delson and Retius, 2015).

In the same way, for ξ = 0, the Logarithm of the Likelihood function for stationary Gumbel model M_0 = GEV [μ, σ, 0] is given by:

$$l(x,\mu_t,\sigma) = -n\ln\sigma - \sum_{i=1}^{n}\exp\left(-\frac{x_i-\mu}{\sigma}\right) - \sum_{i=1}^{n}\frac{x_i-\mu}{\sigma}$$

Differentiating this function for the two parameters, the following system of equations is obtained

$$\begin{cases} n - \sum_{i=1}^{n}\exp\left(-\frac{x_i-\mu}{\sigma}\right) = 0 \\ n + \sum_{i=1}^{n}\frac{x_i-\mu}{\sigma}\left[\exp\left(-\frac{x_i-\mu}{\sigma}\right)-1\right] = 0 \end{cases}$$

2.4.2. *Bayesian analysis of extreme value for GEVD*

The Bayesian approach improves the estimation accuracy by assuming the parameter θ = (μ, σ, ξ) as a random variable. Bayesian inference, based on the central idea of Bayes' theorem which as follows (Choeng and Gabda, 2017):

$$\pi(\theta/x) = \frac{L(\theta/x)\pi(\theta)}{\int L(\theta/x)\pi(\theta)d\theta} \quad (2)$$

where, x is the given observation. $L(\theta/x)$ denotes the likelihood function and $\pi(\theta)$ is the normal prior distribution. The denominator in equation (5) is treated as a normalizing constant so that the posterior distribution is integrated into one. Thereby, results $\pi(\theta/x) \alpha \mathrm{L}(\theta/x)\pi(\theta)$. Markov chain Monte-Carlo is employed to solve the complex computational of the posterior distribution in equation (2). The Markov chain Monte Carlo techniques are applied in this paper to give Bayesian analyses of the annual maximum temperature data for Jijel weather station. Markov chain Monte-Carlo is a simulation technique that can be used to find the posterior distribution and to sample from it by constructing a Markov chain that has the target distribution as its stationary distribution. In this paper, the prior is constructed by assuming that there is no information available about the process (temperature) apart from the data. The annual temperature data have a GEV, *i.e*, $X_i \sim \mathrm{GEV}_{(\mu,\sigma,\xi)}$ and the parameters μ, σ and ξ are treated as random variables for which we specify prior distributions. For specification of the prior, the parameterization $\phi = \log\sigma$ is easier to work with because σ is restricted to be positive. The specification of priors enables us to supplement the information provided by the data. The prior density is (Delson and Retius, 2015) :

$$\pi(\theta) = \pi(\mu,\phi,\xi) = \pi_\mu(\mu)\pi_\phi(\phi)\pi_\xi(\xi)$$

where, each marginal prior is normally distributed with large variances. The posterior distribution is given as:

$$\pi(\theta/x) = \pi(\mu,\phi,\xi/x)\,\alpha\,\mathrm{L}(\mu,\phi,\xi/x)\,\pi(\mu,\phi,\xi)$$

where,

$$L(\mu,\phi,\xi/x) = \frac{1}{\sigma^n}\exp\left\{-\left[1+\xi\left(\frac{x_i-\mu}{\sigma}\right)\right]^{\frac{1}{\xi}}\right\} \prod_{i=1}^{n}\left[1+\xi\left(\frac{x_i-\mu}{\sigma}\right)\right]^{\frac{1}{\xi}}$$

is the likelihood with σ replaced by e^{ϕ} .

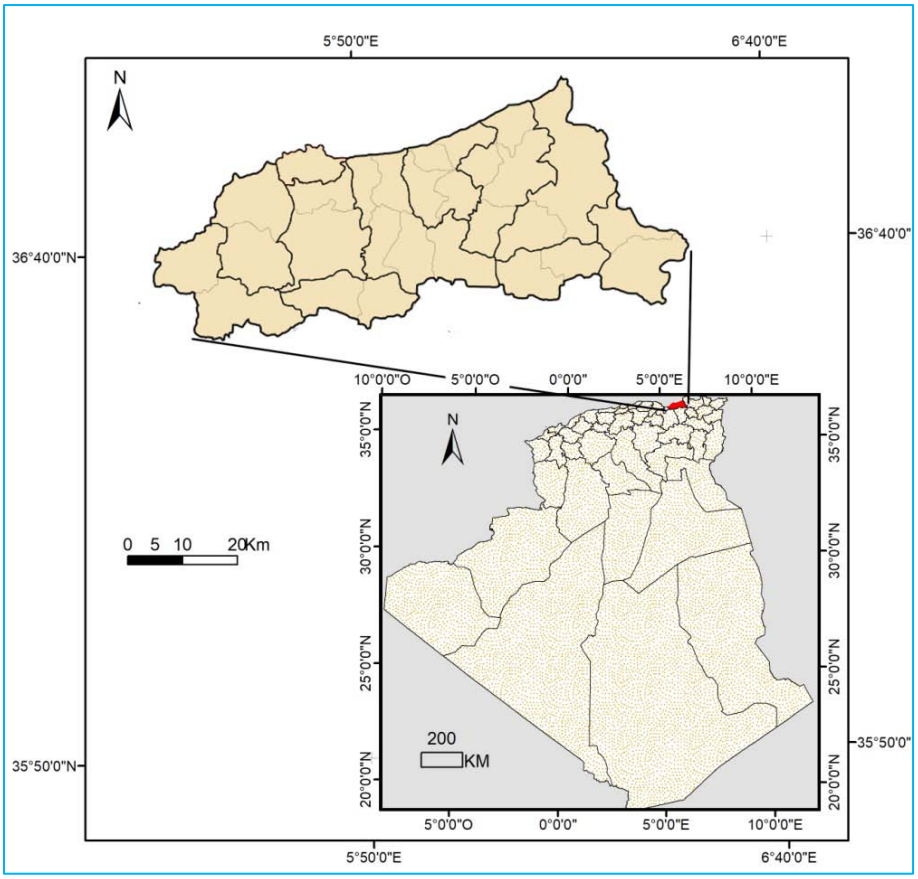


Fig. 1. Location of Jijel city

2.5. *Metropolis-Hastings algorithm*

The Metropolis Hasting algorithm is a form of generalized rejection sampling. The proposal value θ^* for θ_{i+1} is generated from arbitrary rule $q(./\theta_i)$. The Markov chain moves to θ^* with a specified acceptance probability. Specifically, let

$$\alpha_i = \min\left\{1, \frac{\pi(\theta^*)q(\theta_i/\theta^*)}{\pi(\theta_i)q(\theta^*/\theta_i)}\right\}$$

where, $q(\theta^*/\theta_i)$ is denoted as proposal distribution. The candidate is accepted if the probability is equal to α_i. Otherwise, the Markov chain remains in the current state θ_i. The steps involved can be illustrated into the following algorithm (Choeng and Gabda, 2017).

(*i*) Initialize θ_0

(*ii*) In *i* iteration

- Draw a candidate θ^* from proposal distribution $q(\theta^*/\theta_i)$
- Calculate the acceptance probability
- Draw u ~ Uniforme (0, 1)

If $\alpha_i < u$ then Set $\theta_{i+1} = \theta^*$ Else $\theta_{i+1} = \theta_i$

(*iii*) Increment 1 and return to step 2

2.6. *Return level estimate*

The return level is defined as a level that is expected to be equaled or exceeded on average once every interval of time (T) with a probability of *P*. The return level x_p for the return period $\frac{1}{P}$ is the quantile of order 1 – P of the GEV distribution for $0 < P < 1$.

TABLE 1

Statistical properties of the annual maximum temperature at Jijel weather station

N	Min	1^{st}Qu	Mediane	Mean	3^{rd}Qu	Max	Skewness	Kurtosis
31	24.9	25.70	26.10	26.35	26.75	28.60	0.75	2.75

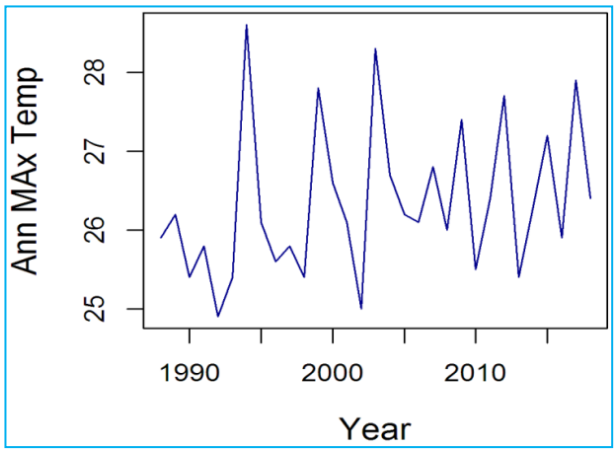


Fig. 2. Plot of annual maximum temperature at Jijel weather station from 1988 to 2018

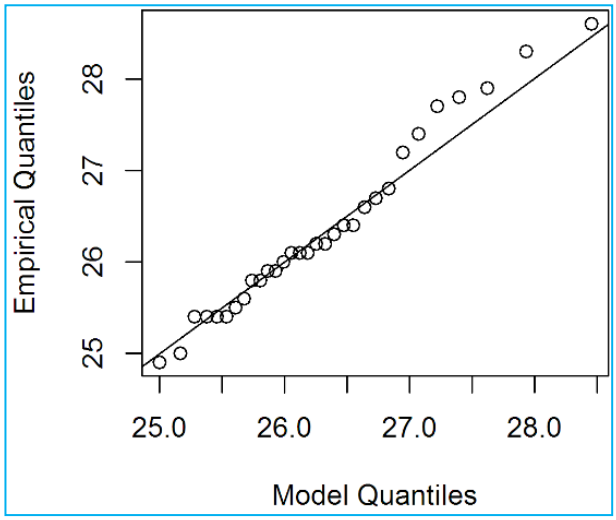


Fig. 3. The QQ plot for the Gumbel distribution of annual maximum temperature at Jijel weather station

The return level x_p is given by:

$$x_P = \begin{cases} \mu - \frac{\sigma}{\xi}\left\{1 - \left[\log(1-P)\right]^{-\xi}\right\}, \xi \neq 0 \\ \mu - \sigma \log\left[-\log(1-P)\right], \xi = 0 \end{cases} \quad (10)$$

TABLE 2

Stationarity checking tests

Sen's slope	KPSS p-value	Mann-Kendall p-value
0.03	0.13	P = 0.068

The Maximum Likelihood estimates of the return level x_p can be obtained using the ML estimates of μ, σ and ξ.

3. Results and discussion

3.1. *Data description*

The temperature data used in our study correspond to 31 value of the annual maximum temperature from 1988 to 2018, measured at Jijel weather station. To analyze the extreme temperature statistically in Jijel city (Fig. 1) with the GEV method, the monthly temperatures are grouped into blocks of size one year and the maximum of each block forms a series of block-maxima.

3.2. *Descriptive analysis*

The analysis was based on the annual maximum temperature at Jijel weather station from 1988 to 2018. The series is given in Fig. 2 and its statistical characteristics are given in Table 1. According to Fig. 2, it seems that the series is not affected by a trend, it appears stationary. From Table 1, it is found that our data is characterized by a mean of 26.35 °C. We note that 50% of the data are between 25.70 °C and 26.75 °C. The Normality assumption is rejected for the series of the annual maximum temperature because the Skewness is greater than zero and the Kurtosis is different from 3.

3.3. *Stationarity checking*

It is necessary to test whether the trend exists in our data. We have two types of stationarity tests : The Mann-Kendall (MK) and the KPSS test. Those tests aim to check if data (annual maximum temperature) is stationary. Estimation of the trend by the least-squares method may not be appropriate, therefore Kendall's tau test and the

TABLE 3

ML estimation of the parameters and theirs confidence intervals

Models	ML estimate	μ_0	μ_1	σ	ξ
Stationary GEV model	Estimates	-	25.90	0.73	0.024
$M_1: GEV_1(\mu, \sigma, \xi)$	Std.err	-	0.15	0.11	0.16
	CI	-	(25.6, 26.2)	(0.5, 0.95)	(-0.29, 0.33)
Non stationary GEV model	Estimates	25.49	0.03	0.68	0.04
	Std.err	0.25	0.01	0.1	0.15
$M_2: GEV_2(\mu(t), \sigma, \xi)$	CI	(24.99, 25.98)	(00, 0.05)	(0.48, 0.89)	(-0.25, 0.34)
Gumbel model	Estimates	-	25.91	0.73	-
$M_0: GEV_0(\mu, \sigma, 0)$	Std.err	-	0.14	0.10	-
	CI	-	(25.64, 26.18)	(0.53, 0.94)	-

TABLE 4

The likelihood ratio test

Models	Observed Chi-Squared	Table Chi-Squared	P-value
$M_1 - M_2$	3.74	3.84	0.053
$M_0 - M_1$	0.02	3.84	0.87

Sen's slope estimator are proposed to compute trend (Sen, 1968). The results of this analysis are presented in Table 2. The results show that the trend is not significant for the annual maximum temperature (KPSS p-value = 0.13 and Mann-Kendall *p*-value = 0.068).

3.4. *Maximum Likelihood (ML) estimates of the parameters*

The Maximum Likelihood method was used to estimate the three parameters of the GEV distribution. The ML estimates of the location, scale and shape parameters and the associated 95% confidence intervals are given in Table 3.

From Table 3, it is noted that the shape parameter ξ is positive in both models M_1 and M_2 $\left(\xi_{M_1} = 0.024; \xi_{M_2} = 0.04\right)$, this is the case of Frechet distribution. Its value is near to zero implying that Gumbel distribution is candidate, the confidence interval of ξ confirms this conclusion $0 \in CI_{\xi}$.

3.5. *Model selection : the Likelihood Ratio (LR) test*

The LR test aims to compare the stationary GEV model M_1 with the non-stationary one M_2 at the first time and with stationary Gumbel model M_0 at the second time. The results are presented in Table 4.

As we compare the result of model M_1 versus model M_2 and Model M_0 versus Model M_1, it can be seen that the Gumbel Model M_0 is preferred for Jijel weather station because of lowest observed Chi-Squared $\left[LR_{(M_1, M_2)} = 3.74 \text{ and } LR_{(M_0, M_1)} = 0.02\right]$.

3.6. *Model validation*

This section aims to validate the selected Gumbel model M_0. In this step, the goodness of fit criteria is taken into account by using a graphical method (the Quantile-Quantile plot) and analytical methods (the non-parametric tests).

TABLE 5

The Goodness fitting tests for the Gumbel distribution M_0

Tests	Kolmogorov Smirnov test (KS)	Anderson Darling test (AD)	Cramer Von Mises test (CVM)
Obs. Value	0.078	0.27	0.037
P-Value	0.99	0.67	0.72
Crit. Value	0.23	0.75	0.21

TABLE 6

Bayesian estimation of the parameters and its confidence intervals

Model	Bayesian estimate	μ_1	σ	ξ
Gumbel model	Estimates	25.90	0.83	-
$M_0: GEV_0(\mu, \sigma, 0)$	Std.err	0.203	0.18	-
	CI	(25.53, 26.33)	(0.55, 1.25)	-

3.6.1. *The Quantile-Quantile (QQ) plot*

To validate the chosen model, the Quantile-Quantile (QQ) plot technique (Fig. 3) was used. It can be seen that for the annual maximum temperature the QQ plot is approximately linear; showing that the stationary Gumbel model M_0 is adequate for the annual maximum temperature at Jijel weather station.

3.6.2. *The Goodness of fit tests*

The results of the non-parametric tests (Kolmogorov Smirnov (KS), Anderson Darling (AD) and Cramer Von Mises (CVM)) are resumed in Table 5.

For the 5% significant level, the calculated values of the 03 tests are less than all the critical ones. This leads to the decision of non-rejection of the null hypothesis. We conclude that the annual maximum temperature follow the specified stationary Gumbel model M_0.

3.7. *Bayesian estimates of the parameters for the stationary Gumbel model M_0*

The MCMC method was applied to the annual maximum temperature data. The GEV scale parameter was re-parameterised as $\sigma = \log\phi$ to retain the positivity of this parameter. The prior density was chosen to be:

$$\pi_\mu(\mu) \sim N(0,10000)$$

$$\pi_\phi(\phi) \sim N(0,10000)$$

$$\pi(\mu,\phi) \sim \pi_\mu(\mu)\pi_\phi(\phi)$$

The proposal distribution function is:

$$q(./\theta_i) \sim N(0,0.1)$$

A MCMC Metropolis Hastings random walk algorithm was applied such as:

$$\mu' = \mu_i + \varepsilon_\mu, \quad \varepsilon_\mu \sim N(0,0.02)$$

$$\phi' = \phi_i + \varepsilon_\phi, \quad \varepsilon_\phi \sim N(0,0.1)$$

To check that the chain will converge to the correct place, the ML estimate of the three parameters was used as starting points $(\mu_0 = 25.91, \phi_0 = 0.73)$. 10000 iterations

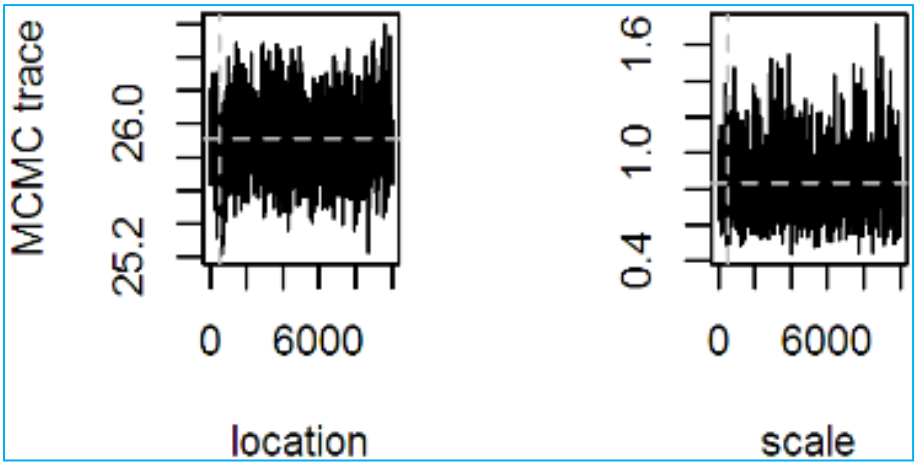


Fig. 4. Trace plots of the Gumbel distribution parameters using non-informative priors for annual maximum temperature at Jijel weather station (1988-2018)

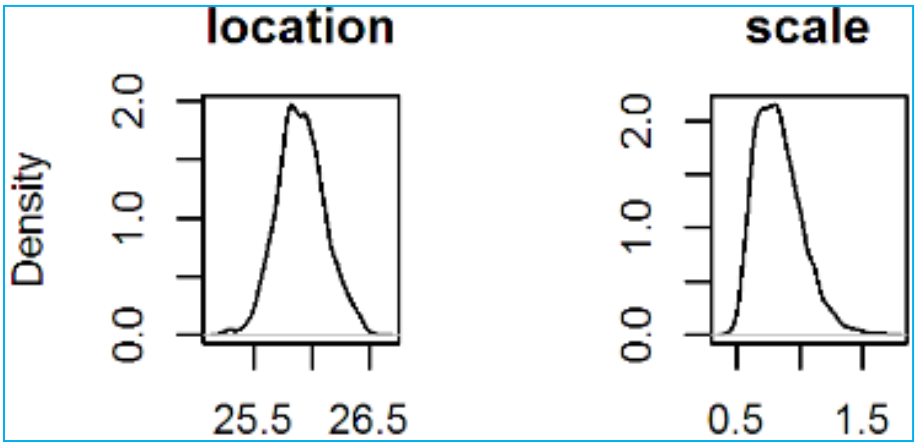
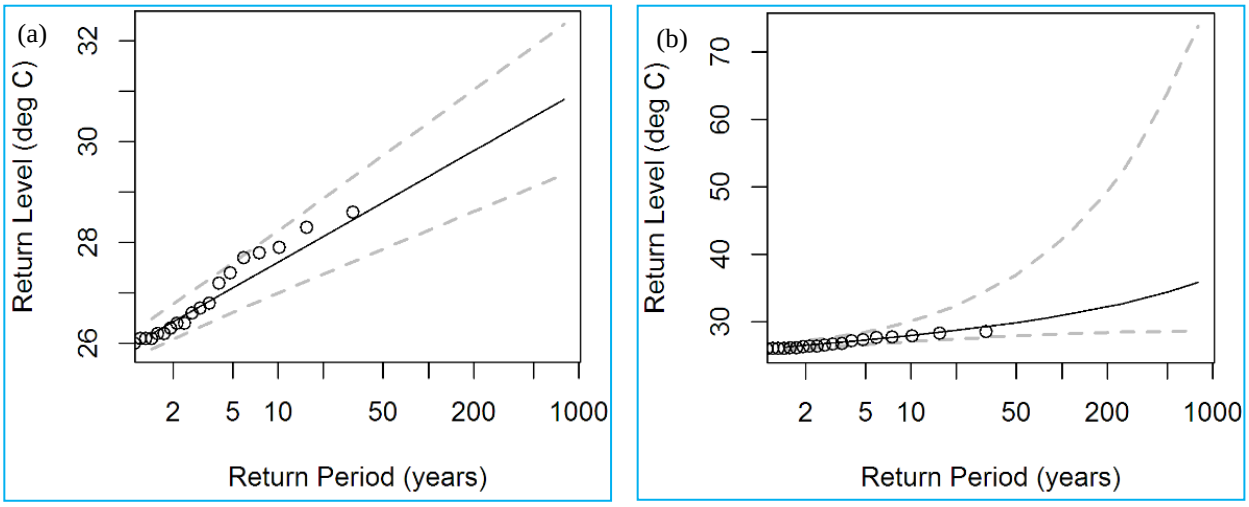


Fig. 5. Posterior densities of the Gumbel distribution parameters using non-informative priors for annual maximum temperature at the Jijel weather station



Figs. 6(a&b). Return levels of the annual maximum temperature : (a) ML return levels estimates and (b) Bayesian return levels estimates

TABLE 7

Return levels estimates

Method / Period	ML return levels		Bayesian Return levels	
	x_P (in °C)	CI	x_P (in °C)	CI
2-year	26.18	(25.87, 26.48)	26.21	(25.80, 26.72)
20-year	28.10	(27.35, 28.84)	28.70	(27.50, 32.23)
50-year	28.78	(27.86, 29.71)	29.87	(27.94, 36.87)
100-year	29.30	(28.23, 30.36)	30.93	(28.20, 42.17)
200-year	29.81	(28.60, 31.02)	32.20	(28.20, 49.27)

of the algorithm were carried out with a time of burn-in = 500. The results are resumed in Table 6. Fig. 4 shows the Markov chain Monte Carlo trace plots and the estimated posterior densities for the Gumbel parameters are given in Fig. 5.

3.8. *Return level estimates*

The estimation of the T-year return levels for T = 2, 20, 50, 100 and 200 are estimated using both ML and Bayesian methods with 95% confidence intervals (CI) as shown in Table 7 and Figs. 6(a&b). It can be seen from this Table that the return levels for maximum annual temperature and theirs confidence intervals increase slowly for higher return periods in both estimation methods.

Note also that estimated return levels for Jijel weather station are slightly differ for the two estimation methods; the Bayesian return level estimates are consistently greater than ML ones for all return periods.

4. Conclusions

In this study, the annual maximum temperature from 1988 to 2018 was modeled using stationary Gumbel distribution. The stationarity of our data was approved by the Mann-Kendall and KPSS tests which showed a non-significant trend in the data. This result was confirmed by applying the likelihood ratio test to compare the stationary GEV versus the non-stationary one. Model diagnostics that included QQ plot and goodness fitting tests (KS, AD and CVM) showed that the annual maximum temperature follows a stationary Gumbel distribution. We rely on an alternative method of Bayesian MCMC based on the Metropolis-Hastings algorithm to estimate the parameters in order to predict different levels return and theirs coming periods. The Gumbel parameter estimates using the Bayesian approach were close to the maximum likelihood estimates, with larger standard deviations. Bayesian return level estimates show that it will be a high annual temperature in the next years, it is approximately 30,93°C for 100 years-return level. The Algerian government should take the measure of prevention of the risk that can be caused by this expected pick of temperature by setting up an alert and vigilance plan. The use of an informative prior or using a pick over high threshold approach POT based on the generalized Pareto distribution GPD may improve the work.

Acknowledgement

We would like to thank everyone who helped us. We thank the editor and the anonymous reviewers for their valuable comments which improved the earlier version of the manuscript. We would also like to thank the Algerian Office of Meteorology for providing the data.

Disclaimer

The contents and views expressed in this research paper/article are the views of the authors and do not necessarily reflect the views of the organizations they belong to.

References

Boudrissa, N., Cheraitia, H. and Halimi, L., 2017, "Modelling maximum daily yearly rainfall in northern Algeria using generalized extreme value distributions from 1936 to 2009", *Meteorol. Appl*., **24**, 114-119. doi : 10.1002/met.1610.

Bronaugh, D. and Werner, A., 2013, "zyp : Zhang & Yue-Pilon trends package", R package version 0.10-1. URL: http://CRAN.R-project.org/package=zyp.

Choeng, R. Y. and Gabda, D., 2017, "Modelling maximum river flow by using Bayesian Markov Chain Monte Carlo", *J. Phys., Conf* : Ser 890 012146. https://doi.org/10.1088/1742-6596/890/1/012146.

Christopeit, N., 1999, "Estimating parameters of an extreme value distribution by the method of moments", *J. Stat.*, **41**, 2, 173-186. https://doi.org/10.1016/0378-3758(94)90162-7.

Coles, S. and Pericchi, L., 2003, "Anticipating catastrophes through extreme value modeling", *J. R. Stat. Soc*., **52**, 405-416. https://doi.org/10.1111/1467-9876.00413.

Coles, S., Pericchi, L. R. and Sisson, S., 2003, "A fully probabilistic approach to extreme rainfall modeling", *J. Hydrol*., **273**, 35-50. https://doi.org/10.1016/S0022-1694(02)00353-0.

Coles, S. G. and Powell, E. A., 1996, "Bayesian methods in extreme value modeling : A review and new developments", *International Statistical Review*, **64**, 119-136. https://doi.org/ 10.2307/1403426.

Coles, S. G. and Tawn, J., 2005, "Bayesian modeling of extreme surges on UK east coast", *Philos. Trans. R. Soc. A.*, **363**, 1387-1406. https://doi.org/10.1098/rsta.2005.1574.

Coles, S. G. and Tawn, J. A., 1996, "A Bayesian analysis of extreme rainfall data", *Applied Statistic*, **45**, 463-478. https://doi.org/ 10.2307/2986068.

Coles, S. G., 2001, "An Introduction to Statistical Modeling of Extreme Values", Springer-Verlag, New York. https://doi.org/10.1007/ 978-1-4471-3675-0.

Deka, S., Borah, M. and Kakaty, S. C., 2011, "Statistical analysis of annual maximum rainfall in North-East India: an application of L-moments", *Theor. Appl. Climatol*., **104**, 111-122. https://doi.org/10.1007/s00704-010-0330-7.

Delson, C. and Retius, C., 2015, "Modelling of extreme minimum rainfall using generalised extreme value distribution for Zimbabwe", *South African Journal of Science*, **111**, 1-8. https://doi.org/10.17159/sajs.2015/20140271.

Fisher, R. A. and Tippet, L. H. C., 1928, "Limiting forms of the frequency distribution of the largest or smallest member of a sample", *Proc. Cambridge Philos. Soc*., **24**, 180-190. https://doi.org/10.1017/S0305004100015681.

Fréchet, M., 1927, "Sur la loi de probabilité de l'écart maximum", *Ann. Soc. Pol. Math*., **6**, 3, 93-116.

Friederichs, P., 2010, "Statistical downscaling of extreme precipitation events using extreme value theory", *Extremes*, **13**, 109-132. https://doi.org/10.1007/s10687-010-0107-5.

Gilleland, E. and Katz, R. W., 2011, "New software to analyze how extremes change over time", *Eos., Transactions of the American Geophysical Union*, **92**, 2, 13-14. https://doi.org/10.1029/ 2011EO020001.

Gnedenko, B. V., 1943, "Sur la distribution limite du terme maximum d'une série aléatoire", *Ann. Math*., **44**, 423-453. https://doi.org/10.2307/1968974.

Greenwood, A., Maciunas, L. J. and Matalas, N. C., 1979, "Probability Weighted Moments : Definition and Relation to Parameters of Several Distributions Expressable in Inverse Form", *Water Resources Research*, **15**, 2, 1049-1054. doi : 10.1029/ WR015i005p01049.

Gumbel, E. J., 1958, "Statistics of Extremes", Columbia University Press, New York. https://doi.org/10.7312/gumb92958.

Hill, B. M., 1975, "A simple general approach to inference about the tail of a Loi", *Annals of Statistics*, **3**, 5, 1163-1174. doi : 10.1214/ aos/1176343247.

Hosking, J. R. M. and Wallis, J. R., 1997, "Regional Frequency Analysis : An approach Based on L-moments", Cambridge University Press, New York.

Hosking, J. R. M., 1990, "L-moments: analysis and estimation of distributions using linear combinations of order statistics", *J. R. Stat. Soc. B.*, **52**, 105-124. https://doi.org/10.1111/j.2517-6161.1990.tb01775.x.

Husna, H., Norfatine, S. and Mohd., B. A., 2013, "Modelling Extreme Temperature in Malaysia Using Generalized Extreme Value Distribution", *World Academy of Science, Engineering and Technology : International Journal of Mathematical and Computational Sciences*, **7**, 6, 983-989. waset.org/Publication/ 12983.

Jenkinson, A. F., 1955, "The frequency distribution of the annual maximum (or minimum) values of meteorological elements", Quarterly Journal of the Royal Meteorological Society, 81, 348, 158-171.https://doi.org/10.1002/qj.49708134804.

Katz, R. W., Parlange, M. B. and Naveau, P., 2002, "Statistics of extremes in hydrology", *Adv. Water Resour*., **25**, 1287-1304. https://doi.org/10.1016/S0309-1708(02)00056-8.

Khodja, H., Lubès-Niel, H., Sabatier, J. M., Servat, E. and Paturel, J. E., 1998, "Analyse spatio-temporelle de donneés pluviométriques en Afrique de l'Ouest, Recherche d'une rupture en moyenne, Une alternative intéressante: Les tests de permutations", *Rev. Stat. Appl*., **46**, 95-110. http://www.numdam.org/item?id= RSA_1998461950.

Kim, C., Suh, M. S. and Hang, K. O., 2009, "Bayesian change-point analysis of the annual maximum of daily and sub-daily precipitation over South Korea", *J. Clim*., **22**, 6741-6757. https://doi.org/10.1175/2009JCLI2800.1.

Lu, Z. Q. and Berliner, L. M., 1999, "Markov switching time series models with application to a daily runoff series", *Water Resour. Res*., **35**, 523-534. https://doi.org/10.1029/98WR02686.

Lye, L. M., Hapuarachichi, K. P. and Ryan, S., 1993, "Bayes estimation of the extreme value reliability function", *IEEE Trans. Reliab*., **42**, 4, 641-644. https://doi.org/10.1109/24.273598.

Parent, E. and Bernier, J., 2003, "Encoding prior experts judgments to improve risk analysis of extreme hydrological events via POT modeling", *J. Hydrol*., **283**, 1-18. https://doi.org/10.1016/S0022 -1694(03)00080-5.

Perreault, L., Hache, M., Slivitzky, M. and Bobee, B., 1999, "Detection of changes in precipitation and runoff over eastern Canada and US using a Bayesian approach", *Stochastic Environ Res Risk Assess*, **13**, 201-216. https://doi.org/10.1007/s004770050039.

Pickands, J., 1975, "Statistical inference using extreme order statistics", *Ann. Statist*., **3**, 1, 119-131. doi : 10.1214/aos/1176343003.

R. Core team, 2015, "R. Development Core Team R : A Language and Environment for Statistical Computing", R. Foundation for Statistical Computing, Vienna.

Renard, B., Garreta, V. and Lang, M., 2006, "An application of Bayesian analysis and Markov chain Monte Carlo methods to the estimation of a regional trend in annual maxima", *Water Resources Researchs*, **42**, W12422. https://doi.org/ 10.1029/2005WR004591.

Renard, B., Lang, M. and Bois, P., 2006, "Statistical analysis of extreme events in a non-stationary context via a Bayesian framework", *Stochastic Environ Res. Risk Assess*. doi : 10.1007/s00477-006-0047-4.

Rydén, J., 2011, "Statistical Techniques for Exploring Possibly Increasing Trend of Hurricane Activity, Recent Hurricane Research - Climate, Dynamics and Societal Impacts", Anthony Lupo, IntechOpen. doi : 10.5772/14980.

Sen, P. K., 1968, "Estimates of the regression coefficient based on Kendall's tau", *Journal of the American Statistical Association*, **63**, 324, 1379-1389. https://doi.org/10.1080/01621459.1968.10480934.

Smith, R. L. and Naylor, J. C., 1987, "A comparison of maximum likelihood and Bayesian estimators for the three parameter Weibull distribution", *J. R. Soc. C*., **36**, 3, 358-369. https://doi.org/10.2307/2347795.

Smith, R. L., 1999, "Bayesian and frequentist approches to parametric predectiv inference (with discussion), in Bayesian Statistics 6 (J. M. Bernardo, J. O. Berger, A. P. Dawid and A. F. M. Smith)", Oxford University press, London, 589-612.

Stephenson, A. and Ribatet, M., 2014, "evdbayes : Bayesian Analysis in Extreme Value Theory", Rpackage version 1.1-1. URLhttps://CRAN.R-project.org/package=evdbayes.

Stephenson, A. G., 2002, "evd : extreme value distributions", *R. News*, **2**, 31-32. URL http://CRAN.R-project.org/doc/Rnews/.

Tadele, A. D., Ondego, J. B. and Legesse, K. D., 2015, "Modeling extreme daily temperature using generalized pareto distribution at port Elizabeth", Proceedings of the 57th Annual Conference of SASA : 41-48. https://www.researchgate.net/publication/286346328.

Thyer, M. and Kuczera, G., 2000, "Modeling long-term persistence in hydro-climatic time series using a hidden state Markov model", *Water Resour Res.*, **36**, 3301-3310. https://doi.org/10.1029/2000WR900157.

Thyer, M., Kuczera, G. and Wang, Q. J., 2002, "Quantifying parameter uncertainty in stochastic models using the Box-Cox transformation", *J. Hydrol.*, **265**, 246-257. https://doi.org/10.1016/S0022-1694(02)00113-0.

Wang, Q. J., 2001, "A Bayesian joint probability approach for flood record augmentation" *Water Resour. Res.*, **37**, 1707-1712. https://doi.org/10.1029/2000WR900401.