Journal of Renewable Energies 27 (2024) 213 – 227

DOI: https://doi.org/10.54966/jreen.v27i2.1189

Journal of Renewable Energies

Revue des Energies Renouvelables

journal home page: https://revue.cder.dz/index.php/rer

Research paper

Applying the Generalized Pareto Principle to Predict Peak Temperatures in Northeast Algeria

Aberrahmane Ziari [a,*], **Abderrahmane Medjerab** [a], **María Teresa Camacho Olmedo** [b]

[a] *University of Science and Technology Houari Boumediene, Algiers, Algeria. LREAU, FSTGAT, USTHB*

[b] *Department of Regional Geography, Faculty of Philosophy, University Campus, Granada, Spain*

ARTICLE INFO

Article history:

Received May 3, 2024

Accepted December 3, 2024

Keywords:

Climate change,
Generalized Pareto,
Distribution Global warming,
Northeastern Algeria.

ABSTRACT

This study employs the Generalized Pareto Distribution (GPD) to model high temperature events at the Batna station, strategically applying varying thresholds to enhance accuracy and mitigate tail distribution bias. This was done by analyzing historical rainfall data from 1981 to 202. Using maximum likelihood estimation and an innovative approach to fitting multi-threshold GPDs, stable thresholds were established, particularly identifying the Pareto II type as the most suitable model for temperatures exceeding 25°C. Analysis revealed a consistent average temperature of 27.52°C, with a negative skewness indicating a bias toward lower values within extreme temperatures. The study also predicts return levels for different periods, providing critical temperature thresholds expected to be exceeded every 2, 20, 50, and 100 years. These findings contribute to understanding extreme temperature dynamics in the region, supporting urban planning, infrastructure design, and climate resilience strategies.

1. INTRODUCTION

According to projections by the Intergovernmental Panel on Climate Change (IPCC, 2021), the Mediterranean region is expected to experience an increase in average annual temperatures from 1.5°C to 2.5°C by the end of the century, Surpassing global averages due to increased vulnerability to global warming. Concurrently, the region could experience a 35% reduction in summer rainfall on the south coast and 25% on the north coast, which would significantly reduce the frequency of rainy days (Lionello et al., 2012). This trend highlights the urgent need for targeted adaptation strategies, especially in countries like Algeria, where recent studies show that temperatures are rising faster than the global average, while rainfall patterns are becoming increasingly unpredictable (Beniston et al., 2018).

[*] *Corresponding author, E-mail address: ziariabderrahmane@gmail.com ; aziari@usthb.dz*
Tel : + 213 552032584

ISSN: 1112-2242 / EISSN: 2716-8247

These anticipated climate changes pose significant risks to ecosystems, agriculture, water resources and public health. Research suggests that Algeria may face increased frequency and intensity of droughts, heat waves and severe weather events, resulting in a significant decrease in water availability, Which would have a cascade effect on agriculture, drinking water sources and energy production in the region (Beniston et al., 2018; WMO, 2019). According to the World Health Organization (WHO, 2020), climate change can also promote the spread of diseases such as malaria because rising temperatures allow vector-borne diseases to reach areas as far away as not affected, adding additional health risks (WHO, 2020; Caminade et al., 2014).

In response to these climatic pressures, a comprehensive understanding of extreme weather events and their patterns in northeastern Algeria is essential. Developing robust adaptation strategies informed by this knowledge will be essential to mitigate the impacts of rising temperatures and decreasing rainfall, thereby supporting sustainable development and public health .

2. LITERATURE REVIEW AND METHODOLOGY

The literature on the Mediterranean climate and Algeria underscores an urgent need for comprehensive studies on climate extremes. While early studies provided foundational insights into rainfall patterns, recent research using statistical models like GPD has expanded understanding of temperature extremes and their long-term impacts. This body of research is critical for forming adaptation strategies to protect vulnerable regions, especially in northeastern Algeria, where climatic extremes pose serious risks to water security, agriculture, and health. Going forward, further research integrating local climate data with advanced modeling techniques will be essential for effective climate resilience planning across the Mediterranean basin (Giorgi & Lionello, 2008; Cramer et al., 2018).

2.1 Introduction to Mediterranean Basin Climate Challenges

The Mediterranean basin, situated between Africa, Asia, and Europe, is a region of immense ecological and socio-economic significance. With over 2.6 million km² of coastal watersheds and critical biodiversity, it supports 4% to 18% of known marine species while covering less than 1% of the world's ocean area (Macias et al., 2013; Coll et al., 2010). However, the region faces heightened vulnerability to climate change, marked by warming temperatures and declining rainfall (Lionello et al., 2012).

Fig 1. Mediterranean basin. (d-Maps, 2023)

According to the IPCC's projections, Mediterranean regions may witness a decrease in precipitation by as much as 50% by 2100, with temperatures rising up to 3.8°C under high-emission scenarios (IPCC, 2021). Such changes include more prolonged droughts and higher temperatures, particularly in southern and eastern Mediterranean areas, exacerbating the risk of water scarcity, agricultural loss, and ecosystem stress (Cramer et al., 2018; Guiot & Cramer, 2016). These projections emphasize an urgent need for mitigation, particularly in semi-arid and arid areas where water resources are limited and demand is high (García-Ruiz et al., 2011).

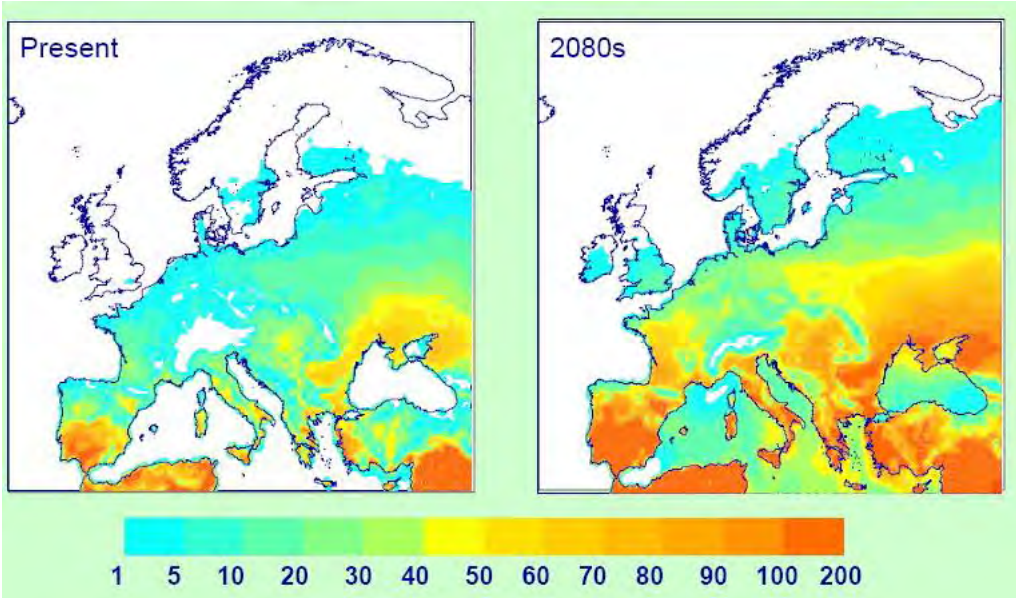

Fig 2. Daily maximum temperatures Number of days per year above 30C° (UK Met Office)

2.2 Climate Change Impacts in Algeria

Algeria, with a climate that spans from Mediterranean along the coast to desert in the south, illustrates the stark regional diversity within the Mediterranean basin (Toumi, 2015). Its northern coast experiences wet, mild winters and dry, hot summers, whereas inland areas are semi-arid, transitioning into the hyper-arid Sahara in the south (Giorgi & Lionello, 2008). This climatic diversity has significant implications, particularly as the north faces growing pressures from urbanization and agriculture, while the arid regions deal with extreme temperatures and minimal rainfall (Medjerab, 1984; Habibi, 2013).

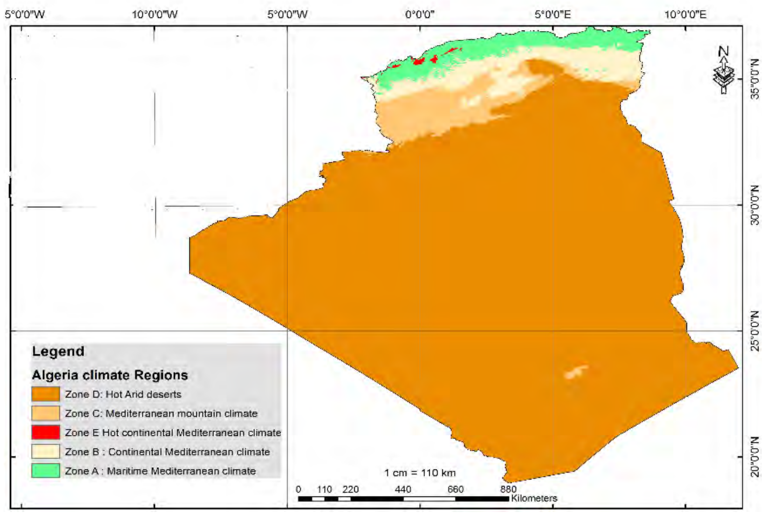

Fig 3. Climatic region map of the 5 zones (Map designed by A.Ziari using ArcGIS 9.2)

2.2.1 Zone D

This large area, covering most of southern Algeria, represents the Sahara Desert region. It is characterized by extremely hot and arid conditions with minimal precipitation. This is the most dominant climate zone in the country.

2.2.2 Zone C

This zone is located in some parts of northern Algeria, likely covering regions of higher elevation, such as the Atlas Mountains. The Mediterranean mountain climate brings cooler temperatures and more rainfall compared to the lowland regions, especially in the winter.

2.2.3 Zone E

These small, isolated regions, seen in red on the map, likely correspond to localized areas that experience a hotter version of the Mediterranean climate but with more extreme summer temperatures and cold winters.

2.2.4 Zone B

This climate zone extends across parts of northern Algeria, further inland from the coast. It experiences hot summers and cool to cold winters, with a moderate amount of precipitation, particularly in winter.

2.2.5 Zone A

This narrow strip along the northern coastline of Algeria represents the coastal regions. These areas experience mild, wet winters and hot, dry summers, typical of a Mediterranean climate heavily influenced by the nearby Mediterranean Sea.

Climate models predict that Algeria may experience more frequent droughts, intense heatwaves, and reduced water availability, which threaten both agriculture and public health (Beniston et al., 2018; WMO, 2019). For example, research by Diffenbaugh and Giorgi (2012) suggests that, by mid-century, temperatures in Algeria's northern regions may increase beyond the global average, with extreme temperatures affecting crop yields and water resources. Furthermore, these environmental changes facilitate the spread of diseases like malaria in areas previously unaffected, introducing new public health risks (WHO, 2020; Caminade et al., 2014).

2.3 Climate Extremes and Methodological Approaches

Early research on Algeria's climate predominantly focused on water resources due to their critical role in semi-arid regions. Studies by Medjerab (1984) and Habibi (2013) concentrated on rainfall variability and drought patterns. However, this emphasis left gaps in the understanding of other extreme weather events, such as heatwaves and severe temperature variations, which have significant implications for long-term resilience and infrastructure stability (Toumi et al., 2019; Cherif et al., 2017).

Recent studies address these gaps by applying advanced statistical methods like the Generalized Pareto Distribution (GPD) to model temperature extremes, particularly in northeastern Algeria. Cheratia (2021) uses GPD to analyze Batna's temperature data, providing a framework to predict rare events and extreme temperature trends effectively. Such methodologies are essential, given that extremes are expected to intensify under climate change, as indicated by research using GPD models to understand distribution tails (Coles et al., 2001; Davison & Smith, 1990). This approach enables a more precise understanding of risk associated with climate extremes, offering critical insights for decision-making (Gilleland & Katz, 2011).

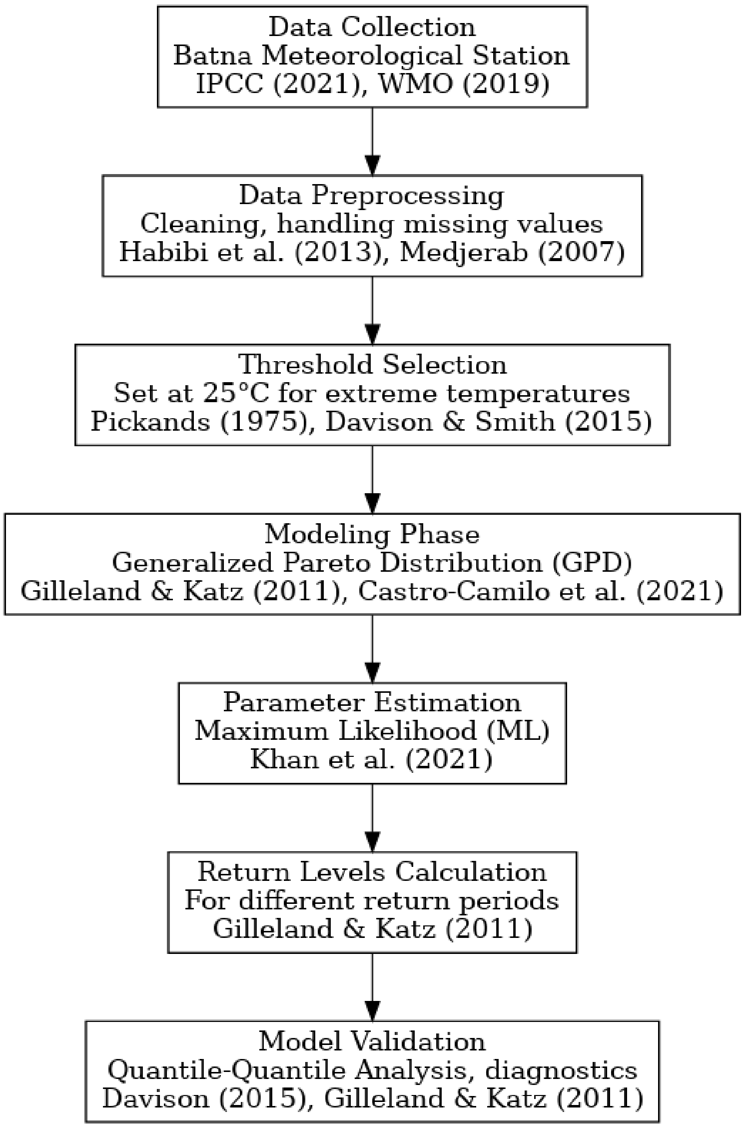

Fig 4. Workflow diagram of the methodological approach to modelling extremes temperature

2.4 Regional Focus: Batna's Semi-Arid Climate

The Batna region, located in northeastern Algeria, provides a microcosm of Algeria's broader climate challenges. Positioned at the intersection of the Tellian Atlas and Saharan Atlas ranges, Batna experiences semi-arid conditions with substantial seasonal variations due to its mountainous terrain and altitude (Toumi, 2015). This region is classified into three bioclimatic zones: a high-moisture zone (above 1,800 m with rainfall up to 1,200 mm annually), a middle zone (moderate rainfall, 400–800 mm), and an arid zone (low rainfall, 200–400 mm), illustrating the climatic diversity within a relatively small geographic area (Castro-Camilo, 2021).

The unique climate of Batna, especially in its arid zones like the Beni Imloul forest, experiences extreme temperatures, reaching up to 45°C in summer. This pattern underscores the importance of understanding the region's climate variability and extremes to mitigate future risks (Toumi et al., 2019; Cherif et al., 2017). Effective adaptation strategies, informed by local climate data, are essential for water resource management, agriculture, and public health, especially as extreme temperature days are projected to increase (Cheratia, 2021; Bronaugh & Werner, 2013).

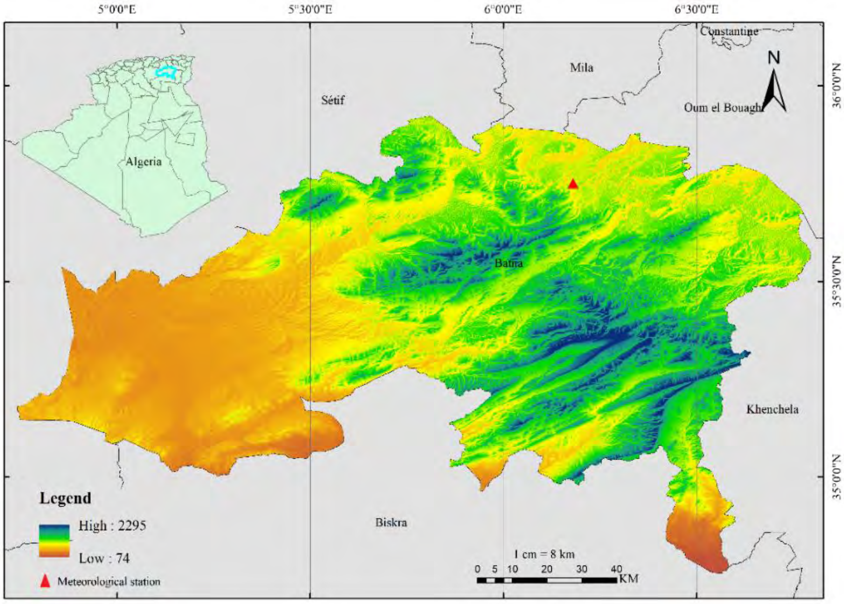

Fig 5. Location of study area (Map designed by A.Ziari using ArcGIS 9.2)

3. MODELING

The use of extreme value theory to establish thresholds in statistical analysis is a significant advancement, particularly influenced by foundational contributions from Pickands (1975) and others. This methodological framework, based on the Balkema-de-Haan-Pickands theorem, enables effective simulation of distribution tails, which is essential for analyzing extreme observations. Recent studies have underscored the importance of the Generalized Pareto Distribution (GPD) in temperature extreme modeling, highlighting its versatility across various applications.

Recent work, such as that by Maposa et al. (2021), explores a bivariate time-varying threshold approach to model extreme temperatures, demonstrating how the GPD can be adapted to capture the complex dynamics of maximum temperatures in South Africa. Similarly, Chen et al. (2024) propose an asymmetric GPD based on a peaks-over-threshold model, reinforcing the distribution's utility in handling asymmetries in climate extremes.

Studies by O'Sullivan et al. (2019) use a Bayesian approach to spatially analyze temperature extremes in Ireland, showing how the GPD can integrate into spatial models to better represent local climate extremes. Additionally, Raggad (2018) compares stationary and non-stationary approaches for modeling extreme temperatures in Riyadh, highlighting the GPD's value in capturing temporal variations in temperature extremes. Although focused on rainfall events, Martins et al. (2020) also illustrate the flexibility of the GPD in analyzing intense climatic events in Latin America.

These studies confirm the relevance of the GPD in analyzing temperature extremes and other climatic phenomena, providing a robust methodological foundation for future research and applications in climate risk management.

The data used in our analysis were obtained from the Batna meteorological station and include monthly average temperatures recorded from 1981 to 2021. This extended observation period strengthens the robustness of our GPD model, as it captures long-term climatic variations and offers a comprehensive understanding of regional climate trends and anomalies. By accurately identifying and modeling temperature extremes, our study makes a meaningful contribution to the scientific literature, providing enhanced strategies for managing the risks associated with extreme weather events and effectively supporting public policy decision-making.

3.1 Threshold

Employing extreme value theory to establish thresholds for statistical analysis is a crucial progression in the field, prominently influenced by seminal contributions from Pickands (1975), Davison (2015), Frid (2012), and Smith (1989). These researchers have built upon the Balkema-de-Haan-Pickands theorem, initially established in 1975, which is a cornerstone in the statistical modeling of extreme events. This theorem provides a methodological framework for effectively simulating the tails of probability distributions, which is particularly valuable when dealing with the most extreme observations in data sets. The technique derived from this theorem is designed specifically to handle the tail behavior of distributions, a crucial aspect when the most extreme values are of primary interest. This involves establishing a well-defined threshold, which serves as a demarcation point indicating where the tail begins and the central part of the distribution ends. This threshold is not arbitrary but is determined based on statistical rigor to ensure that the subsequent analysis of tail data is both valid and reliable. The process of setting this threshold is sensitive to the assumptions about the independence of data points and their perturbations. Independence here refers to the lack of autocorrelation within the data, meaning that the occurrence of one extreme value does not influence the occurrence of another. This assumption is vital as dependencies can significantly distort the extremity modeling. Similarly, understanding how data perturbations such as measurement errors or anomalies can affect the analysis is crucial for maintaining the integrity of the model's output.

After setting the threshold, the method aims to choose a group of data points, namely independent peaks that are higher than this threshold. These summits are then believed to adhere to the Generalised Pareto Distribution (GPD). The GPD is particularly suited for this kind of analysis due to its flexibility in modeling different shapes of distribution tails, making it a robust choice for extreme value theory applications. Figure 5 in the referenced material likely illustrates these autonomous peaks and the threshold level, providing a visual representation of how the GPD fits these extremes. This threshold technique not only isolates the most extreme events from the bulk of the data but also ensures that the statistical analysis concentrates on the behavior of the data in these extreme regions.

The assumption that the peak values exceeding the established threshold conform to a GPD allows statisticians to estimate the tail properties such as the frequency and magnitude of extreme occurrences. This is particularly useful in disciplines such as hydrology, finance, and environmental science, where predicting extreme events can be crucial for risk management and planning.

$$P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!} \tag{1}$$

Where: λ: is a positive constant called the law parameter.
X: symbolizes the number of occurrences in a random event.
k: Represents a specific value or realization of the random variable X

We employ the Generalized Pareto Distribution (GPD) to represent the characteristics of distribution tails, specifically by establishing a defined threshold referred to as u. Understanding the parameters of the GPD is essential because they offer valuable information about the extreme values of the original distribution. The "tail index" or shape parameter ξ, is a key parameter in this regard, as it directly informs us about the weight or 'heaviness' of the tails.

In the GPD model, a higher tail index indicates that the tails of the distribution are thicker, suggesting a higher probability of extreme values occurring as compared to what might be expected under a standard normal distribution. This is particularly important in fields like risk management and financial modeling, where understanding the likelihood of extreme outcomes is crucial for making informed decisions.

When the tail index ξ exceeds 0, it implies that the distribution has a "heavy tail." Heavy tails are characteristic of distributions, where the probabilities of realizing extremely high values are non-trivial, thereby indicating a higher likelihood of observing extreme events than would be anticipated under the Gaussian model. This property makes the GPD an excellent choice for modeling extreme values in diverse applications ranging from environmental data analysis to financial risk assessment.

Furthermore, to aid in the practical understanding and visualization of this concept, the Mean Excess Function (ME) Plot is utilized. The ME Plot provides a graphical representation of the expected excess over the threshold u as a function of the threshold itself. It is defined mathematically as:

$$E(X-u \mid X>u) \quad (2)$$

where X is a random variable exceeding the threshold u.

This plot is instrumental in verifying whether the GPD is an suitable model for the data under analysis. It helps in illustrating how the average size of the exceedances over the threshold changes as the threshold increases, thus offering valuable insights into the stability and extremity of the (Please complete the sentence).

The Mean Excess ME Plot is defined as follows:

$$\{(u, e_n(u)), X_{n:n} < u < X_{1:n}\} \quad (3)$$

where, $X_{n:n}$ et $X_{1:n}$ represent the highest and lowest values in the sample defined by.

$$e_n(u) = \frac{\sum_{i=1}^{n}(X_i-u)^+}{\sum_{i=1}^{n} I_{(Z_i>u)}} = \frac{1}{N_u}\sum_{i=1}^{n}(X_i-u)^+ \quad (4)$$

This is the sum of the excesses above the threshold *u* divided by the number N_u of data exceeding *u*. The sample mean excess function $e_n(u)$ is the empirical estimator of the mean excess function.

$$e(u) = \mathrm{E}[X - u\mathrm{I}\ x\ > u] \quad (5)$$

For GPD, the *ME* is as follows:

$$e(u) = \frac{\sigma+\xi u}{1-\xi} \quad (6)$$

If the ME-plot exhibits a linear pattern beyond a specific threshold value of u (25°C in our case), it indicates that the excesses above this threshold conform to a Generalized Pareto Distribution (GPD).

3.2 Results and discussion

If the ME-plot exhibits a linear pattern beyond a specific threshold value of *u* (25°C in our case), it indicates that the excesses above this threshold conform to a Generalized Pareto Distribution (GPD).

By strategically employing various threshold ranges in our analysis, we can significantly improve the accuracy of our model and mitigate potential biases associated with the tail distribution. The Generalized Pareto Distribution (GPD) is particularly well-suited for this purpose, as it allows flexible modeling across different thresholds. This adaptability is crucial for analyzing and determining the optimal threshold that provides the most accurate estimate suited to our specific scenario. In our approach, we further explore the exceptional values within our dataset by applying GPD across these different thresholds. This allows a comprehensive examination of the behavior of different tail segments under different conditions, offering a nuanced view of distribution characteristics. An innovative technique we employ involves fitting data to a multi-row GPD distribution, where scale and shape parameters are plotted against different thresholds. This method, known as “stable scale and shape parameters”, is designed to identify the most stable and appropriate thresholds for analysis. By systematically establishing a distinct threshold each time, we can observe and evaluate the stability of the parameters (shape and scale). This iterative process enables us to locate the exact point at which the parameters show stable behavior, indicating a reliable threshold setting. We implement this method using the Extremes package of the R statistical analysis system (Gilleland et al., 2004). This package is equipped with sophisticated tools specifically designed to determine optimal thresholds based on a detailed analysis of the stability of shape (ε) and scale (σ) parameters. These tools are invaluable for our analysis, as demonstrated in Figure 7, which illustrates the application of the method. Our objective through this analytical process is to determine a range of stability points for these two parameters where the graph displays approximately linear and stable behavior. This stability is crucial to ensure that the model’s predictions are both reliable and robust. Interestingly, we observe that, once the temperature exceeds 25°C, the scale and shape parameters maintain a relatively constant profile, suggesting a significant threshold that may be crucial for understanding temperature extremes in our dataset.

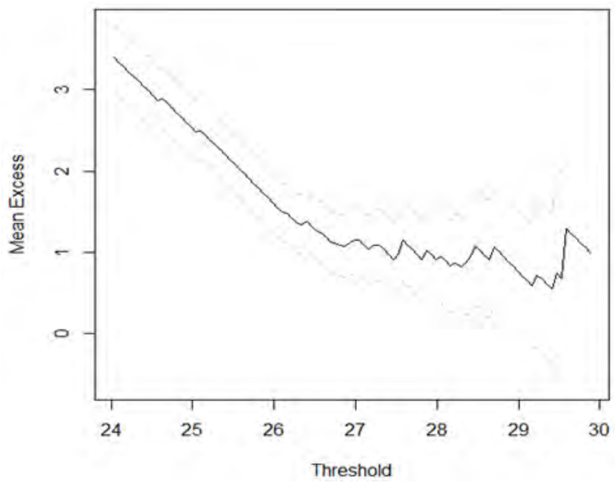

Fig 6. Average monthly maximum temperature exceedance curve at Batna weather station

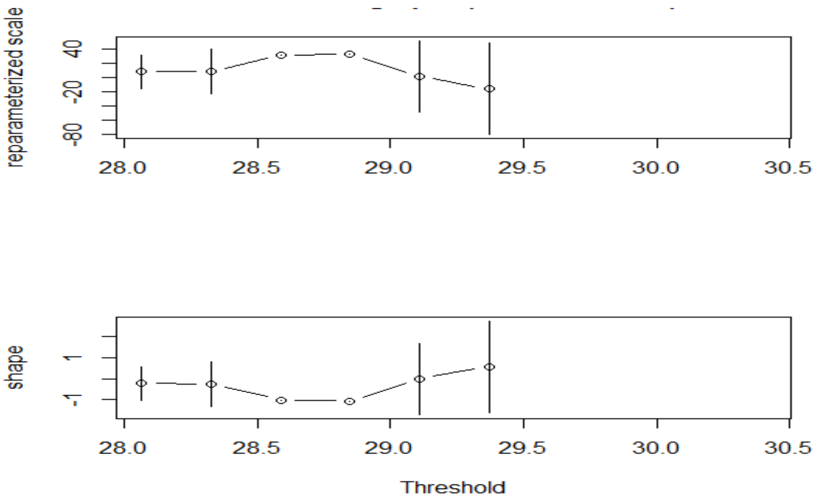

Fig 7. Plot of ML estimates of shape and corrected scaling parameters for different threshold ranges of monthly maximum temperature at Batna meteorological station.

3.3 Descriptive analysis

Figure 8 provides a clear visual depiction of temperature readings that exceed the set threshold, showing these extreme values in a vivid manner. In the meantime, Table 1 thoroughly describes the statistical characteristics linked to these temperatures. The data shows that the station's average temperature is 27.52°C. Additionally, half of the temperatures recorded fall between 25.06°C and 26.85°C, hinting at a consistent temperature pattern in this range. Furthermore, examining the shape of the data also shows a negative skewness, suggesting that the maximum temperatures distribution is biased towards lower values. Additionally, the kurtosis value varies from the standard of 3, resulting in the dismissal of the assumption of normality for this set of monthly maximum temperatures. This departure indicates a distribution with tails that vary from those of a standard distribution.

Figure 6 also confirms that the sequence of peak temperatures exhibits a stationary pattern, showing no significant trends over time and reinforcing the stability of these extreme temperature conditions.

Table 1. Descriptive statistics

N	N_u	Mean	3^{rd}Qu	Max	1^{st}Qu	Min	Kurtosis	Skewness
36	34	27.52°C	28.14°C	30.88°C	26.85°C	25.06°C	0.74	0.70

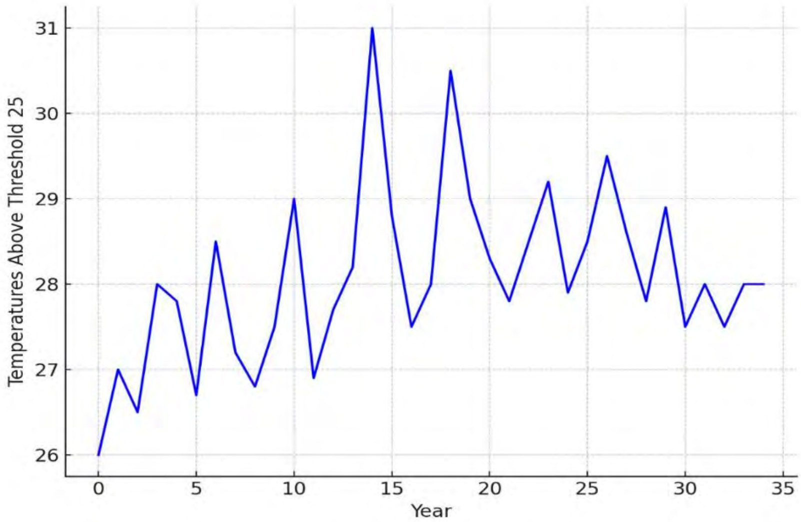

Fig 8. Maximum monthly temperatures above 25°C

4. PARAMETER ESTIMATION AND MODEL VALIDATION

After identifying the surplus values, GPD parameters can be estimated using this data. The parameters include a shape parameter that determines the form of the extreme value distribution and a scale parameter that measures how far the extreme values are from the specified threshold. ML maximum likelihood was used to estimate the parameters. Assuming that our sample of excesses is $X = \left(X_1, \dots, X_{N_u}\right)$, it follows the GPD distribution:

$$GPD_u(x) = 1 - \left[1 - \frac{\xi(x-u)}{\tilde{\sigma}}\right]^{\frac{-1}{\xi}} \quad \xi \neq 0 \tag{7}$$

If $x = y - u$, the density function g of $GPD_u(x)$ is then for:

$$\frac{d}{dx}GPD_u(x) = g(x) \tag{8}$$

$$g(x) = \frac{1}{\sigma}\left(1 + \xi\frac{x}{\sigma}\right)^{-\frac{1}{\xi}-1} \tag{9}$$

The probability is determined by: $\mathcal{L}(\xi, \mu, \sigma, X) = \prod_{i=1}^{n} g(X_i)$ (10)

The log-likelihood is determined as follows: $\iota(\xi, \sigma, X) = \ln \mathcal{L}(\xi, \sigma, X)$ (10)

$$\iota(\xi, \delta; X) = -N_u \ln \sigma - \left(\frac{1}{\xi} + 1\right)\sum_{1}^{N_u} \ln\left(1 + \frac{\xi}{\sigma}X_i\right) \tag{11}$$

Obtaining this function results in: $\frac{\partial \iota(\xi,\sigma;X)}{\partial \xi} = 0$ (12)

$$\frac{\partial \iota(\xi,\sigma;X)}{\partial \sigma} = 0 \tag{13}$$

We derive a system consisting of two equations and two unknowns, with the solution being the Maximum Likelihood estimators $\left(\hat{\xi}_{N_u}, \hat{\sigma}_{N_u}\right)$.
And for ξ=0, we have:

$$g(\chi) = \frac{1}{\sigma}\exp(-\frac{\chi}{\sigma}) \tag{14}$$

$$l(0, \sigma; X) = -N_u ln\sigma - \frac{1}{\sigma}\sum_{i=1}^{N_u} X_i \tag{15}$$

Next, we derive the estimator as the empirical average of the excess values:

$$\delta N_u = \sum_{i=1}^{N_u} {X_i}/{N_u} \tag{16}$$

Table 2 presents the scale and shape parameters estimates, as well as their 95% confidence intervals and covariance matrices.

Table 2. ML estimates, confidence intervals(CI) and covariance matrices for shape and scale parameters of the GPD model fitted for maximum monthly temperature

Scale	Shape
Estimates	3.88 -0.64
Standard error	0.74 0.14
CI (2.42, 5.34)	(-0.91, -0.36)
Estimated parameters covariance matrix	
Scale	0.55 -0.10
Shape	0.019

According to the information displayed in Table 2, it can be seen that the shape parameter is in a negative state, which classifies the Generalized Pareto Distribution (GPD) used in this analysis as of the Pareto II type. This parameter is notably close to zero, distinctly implying that the exponential distribution does not fit within this framework. To substantiate this conclusion, confidence intervals have been meticulously calculated and utilized, offering robust statistical support for the exclusion of the exponential model. Further validation of the chosen models was carried out using the Quantile-Quantile (QQ) plot technique. The QQ plot specifically tailored for the Batna weather station displays a nearly perfect linear correlation, underscoring the appropriateness of the GPD type II model. This model effectively captures the nuances of the highest monthly temperatures, particularly with a set threshold of 25°C. The alignment of the data points along the linear trend in the QQ plot convincingly confirms that the GPD type II model is well-suited for analyzing the extreme temperature events recorded at the Batna station, ensuring accurate representation and analysis of the temperature extremes.

Generalized Pareto distribution (scale=3.88, shape=-0.64)

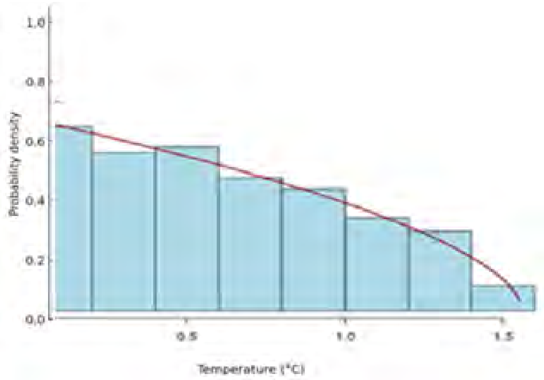

Fig 9. Estimated distribution

4.1 Model validation

The graph (Figure 10) clearly illustrates a positive correlation between statistical model quantiles and empirical quantiles, demonstrating that an increase in model quantiles leads to a corresponding increase in empirical quantiles. This relationship is underlined by the regression line, which indicates an approximate linear trend between these two variables. Furthermore, the 95% confidence intervals reinforce this observation, providing statistical assurance that the true value of the empirical quantities for a specific model quantile is likely to lie within the range specified by the grey bands.

This analysis is corroborated by a second graph, which also compares model quantities with empirical quantities, confirming the existence of a similar positive relationship. As the regression line shows, the relationship between these quantities remains approximately linear. The 95% confidence intervals, illustrated by the grey bands, add a further layer of certainty, indicating that the true value of the empirical quantities for a given quantile lies with 95% probability within the defined interval.

Together, these graphs suggest that the model provides a reliable prediction of empirical quantities, validating its effectiveness as a modeling tool.

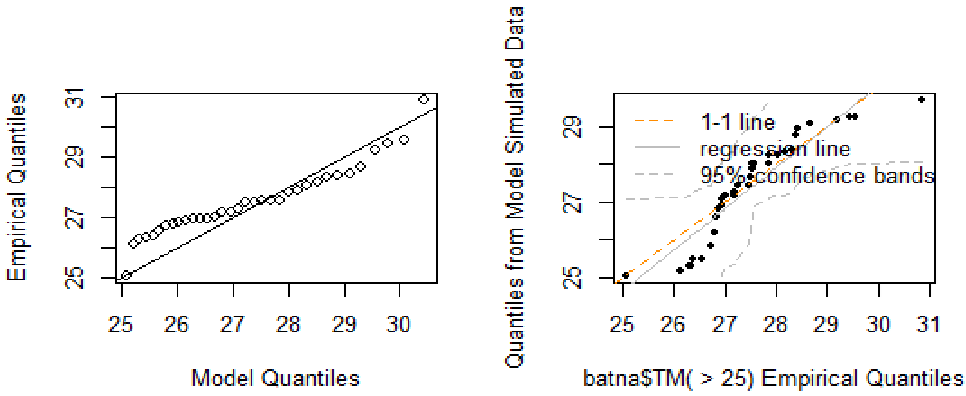

Fig 10. QQ plot for Batna station

4.2 GPD estimated yield

The Maximum Likelihood (ML) estimation method was rigorously applied to determine return levels for various return periods of monthly maximum temperatures, the results being carefully recorded in Table 3. These results are further supported by 95% confidence intervals, rigorously calculated according to the profile probability method, reinforcing the reliability of the estimates. The return level is a crucial statistical measure, representing the upper quartile of the distribution. This metric quantifies the temperature values statistically predicted to exceed a predetermined threshold at regular intervals, called "return periods". The concept of return period is inversely related to the probability of exceeding the temperature threshold in a given year. For example, a return period of 100 years suggests that, on average, temperatures should exceed this specific level once every hundred years. This statistical framework is essential for assessing and mitigating the risks associated with increasing temperature extremes due to climate change. For practical implications, the data predict that temperatures are expected to exceed 30.96°C every 2 years, 31.03°C every 20 years, 31.045°C every 50 years, and 31.049°C every 100 years. These projections are fundamental to urban planning, infrastructure development and emergency preparedness, providing a basis for designing strategies to cope with and adapt to these expected temperature extremes.

Table 3. Monthly peak temperatures reach 9% (°C) using GPD.

Return period	Projected value of temperature level (in °C)
2-years	30.96
20-years	31.03
50-years	31.045
100-years	31.049

5. CONCLUSION

This study highlights an advanced approach to modeling temperature extremes at the Batna station using the Generalized Pareto Distribution (GPD), strategically applying various thresholds to enhance model accuracy and reduce bias in tail distributions. The main findings of this study are as follows:

- By examining different threshold ranges, the analysis identifies the Pareto II type, particularly effective in capturing temperature dynamics above a 25°C threshold, thanks to its tail-limited behavior. Through iterative fitting of scale and shape parameters, stable thresholds were established, ensuring reliable model predictions.
- Key findings include the identification of stable temperature thresholds, with average temperatures around 27.52°C, while values frequently exceeded 25°C, indicating consistency in temperature patterns. The station's temperature distribution shows a negative skewness and non-normality, further supporting the use of GPD type II as validated by Quantile-Quantile (QQ) plots and stability analyses, which demonstrated the model's accuracy in capturing extreme events.
- The analysis of return levels for varying periods reveals a statistically supported framework to predict temperatures expected to exceed specific thresholds (e.g., 30.96°C every 2 years), critical for regional planning and infrastructure. These results provide valuable insights for managing climate-related risks, aiding urban planning and emergency preparedness efforts in response to increasing temperature extremes.

ACKNOWLEDGMENTS

We express our gratitude to the National Office of Meteorology and the National Water Resources Agency for supplying the data. We are also thankful to (power.larc.nasa.gov) for providing additional data. Additionally, we extend our heartfelt appreciation to Professor María Teresa Camacho Olmedo for welcoming me into her laboratory and for her invaluable guidance and support.

DECLARATION OF INTEREST

The authors declare no conflicts of interest and have no known competing financial interests or personal relationships that could influence the work presented in this article.

REFERENCES

Balkema, A. A., & de Haan, L. (1974). Residual life time at great age. *The Annals of Probability, 2*(5), 792–804.

Beniston, M., Stoffel, M., & Guillet, S. (2018). *Climate change and its impacts on glaciers and snow cover*. Cambridge University Press.

Bronaugh, D., & Werner, A. (2013). *zyp: Zhang + Yue-Pilon trends package* [Software]. R Project for Statistical Computing. https://www.r-project.org/

Caminade, C., Kovats, S., Rocklov, J., Tompkins, A. M., Morse, A. P., Colón-González, F. J., Stenlund, H., Martens, P., & Lloyd, S. J. (2014). Impact of climate change on global malaria distribution. *Proceedings of the National Academy of Sciences*, *111*(9), 3286-3291. https://doi.org/10.1073/pnas.1302089111

Castro-Camilo, D., Huser, R., & Rue, H. (2021). Practical strategies for GEV-based regression models for extremes [arXiv preprint arXiv:2106.13110].

Cheraitia, H. (2021). Modélisation de la température mensuelle maximale 1988 à 2018. *Revue d'Économie et de Statistique Appliquée, 18*(1), 45-55. https://www.asjp.cerist.dz/en/article/156834

Davison, A. C., & Smith, R. L. (2015). Models for exceedances over high thresholds. *Journal of the Royal Statistical Society, Series B*, 393–442.

Desvina, A. P., Safitri, E., & Rahma, A. N. (2019). Statistical modelling of extreme data of air pollution in Pekanbaru City. *InPrime: Indonesian Journal of Pure and Applied Mathematics, 1*(1), 48–56.

Frid A, Zamboni L. On automatic infinite permutations. RAIRO - Theoretical Informatics and Applications. 2012;46(1):77-85. doi:10.1051/ita/2011129

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

Gilleland, E., & Katz, R. W. (2016). extRemes 2.0: An Extreme Value Analysis Package in R. Journal of Statistical Software, 72(8), 1–39

Gnedenko, B. V. (1943). Sur la distribution limite du terme maximum d'une série aléatoire. *Annals of Mathematics, 44*, 423–453

Habibi, B., Meddi, M., & Boucefiane, A. (2013). Analyse fréquentielle des pluies journalières maximales Cas du Bassin Chott-Chergui. *Nature & Technology*, (8), 41B.

IPCC (2021). Sixth Assessment Report of Working Group I: The Scientific Basis of Climate Change. United Nations, Geneva, Switzerland.

Khan, M., Hussain, Z., & Ahmad, I. (2021). Effects of L-moments, maximum likelihood and maximum product of spacing estimation methods in using Pearson type-3 distribution for modeling extreme values. *Water Resources Management, 35*(5), 1415–1431. https://doi.org/10.1007/s11269-021-02767-w

Kotz, S., & Nadarajah, S. (2000). *Extreme Value Distributions: Theory and Applications*. Imperial College Press.

Makarov, M. A. (2006). On permutations generated by infinite binary words. *Sibirskie Èlektronnye Matematicheskie Izvestiya, 3*, 304-311.

Medjerab, A. (2007). Les situations pluviométriques extrêmes dans le nord-ouest de l'Algérie. In: Climat, tourisme, environnement, Actes du XXème colloque de l'Association Internationale de Climatologie (Carthage, Tunisie), H. Ben Boubaker (Ed.), p. 381-386.

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

R Core Team. (2015). *R: A Language and Environment for Statistical Computing*. R Foundation, for Statistical Computing, Vienna, Austria. http://www.R-project.org/

Stephenson, A. G. (2002). evd: Extreme value distributions. *R News*. Retrieved from http://CRAN.R-project.org/doc/Rnews/, pp. 641–644.

United Nations Environment Programme. (2023). Climate Change Impacts and Adaptation in the Mediterranean Region. Nairobi: UNEP.

World Health Organization. (2020). Climate Change and Health: Implications for the Mediterranean Region. Geneva: WHO.

World Meteorological Organization. (2019). *WMO Statement on the State of the Global Climate in 2018* (WMO-No. 1234). Geneva: WMO.