Al-Rafidain Journal of Computer Sciences and Mathematics (RJCSM)

www.csmj.uomosul.edu.iq

Hierarchical Bayesian Extreme Value Modeling of Heatwave Intensities: Application to ERA5-Land in a Semi-Arid Region

Saifuldeen Dheyauldeen Alrefaee

Department of Operations Research and Intelligent Technologies, College of Computer Science and Mathematics, University of Mosul, Mosul, Iraq
Email: saifldeen.alrefaee@uomosul.edu.iq

Article information

Article history:
Received 8 March, 2026
Revised 12 April, 2026
Accepted 2 May, 2026
Published 25 June, 2026

Keywords:
Extreme Temperature
Non-stationary GEV
Bayesian Extreme-Value Modeling
Spatial Extremes
Climate Risk Assessment
ERA5-Land

Correspondence:
Saifuldeen Dheyauldeen Alrefaee
Email: saifldeen.alrefaee@uomosul.edu.iq

Abstract

In arid and semi-arid areas, especially in Iraq, extreme heat has become one of the most significant impacts of climate change, as the long duration of summer heat is becoming a serious threat to human health, water supply and infrastructure. This paper explores the issue of summer extreme temperatures in Iraq based on ERA5-Land daily maximum temperature data of 1980-2024. Descriptive heatwave indicators were then computed after which annual summer maxima were modeled in a Bayesian non-stationary generalized extreme value (GEV) framework in three climatic regions: North, Central, and South Iraq. A linear time effect in the location parameter was used to introduce non-stationarity, whilst the scale and shape parameters were fixed to maintain inferential stability because of the small number of annual maxima to use in each region. The results reveal a clear north–south gradient in baseline extreme heat intensity and a statistically credible warming trend of approximately $0.32\,°C$ per decade. Return-level analysis further indicates that 20-year and 50-year extreme temperature levels increased by about $0.7 - 0.8\,°C$ between the early period (1980–1999) and the recent period (2000–2024), with the most severe extremes concentrated in southern Iraq. These findings provide probabilistic evidence that extreme heat risk in Iraq is intensifying under a non-stationary climate and demonstrate the value of Bayesian non-stationary extreme-value modeling for regional climate-risk assessment.

DOI: 10.33899/rjcsm.v20i1.60668,

1. Introduction

Climate change is intensifying the frequency, duration, and severity of extreme heat events worldwide, with particularly severe consequences in arid and semi-arid regions. Among the most vulnerable areas, the Middle East and North Africa (MENA) region has emerged as a major hotspot of climate-related heat risk, experiencing warming rates that exceed the global average and facing increasing pressure on water resources, infrastructure, energy systems, and public health [2,22].

Iraq is particularly prone to such risks due to its already hot climate in summer, frequent instances of droughts and heavy reliance on climate-dependent industries. Recent evaluations have revealed that the nation has undergone significant warming throughout the past several decades and that the number of very hot days and lengthy heat waves has significantly risen [1,11,18].

Nevertheless, a lot of the available literature about Iraq has centred on the alteration in the mean temperature or descriptive trend analysis of temperature indexes [1,18]. These studies have been useful in recording large-scale patterns of warming, but they tell us less about the behaviors of rare and high-impact extremes that are found in the upper part of the temperature distribution. Practically, most heat-risk measurements are based on predetermined thresholds or summary indices, which are convenient to monitor changes over time but do not describe the entire change in the probability structure of rare extreme events [12,21]. This is especially crucial in a warming climate where the stationarity assumption is becoming harder and harder to defend [16].

Extreme Value Theory (EVT) is a natural probabilistic model of rare events and the level of returns that they exhibit. The generalized extreme value (GEV) distribution is a popular type of distribution applied to model block maxima in climate applications, such as annual maximum temperature [3,12]. Nevertheless, when records are few, when spatial heterogeneity is high and when climatic non-stationarity is necessary to be modeled explicitly, reliable estimation of GEV parameters may be difficult. Such problems are especially pertinent to Iraq, where long, homogeneous records of the station are restricted, and where the variability of the climate across regions can be sharp. Bayesian methods have a number of benefits in this context, such as, coherent quantification of uncertainty, partial pooling between regions, and a flexible model of non-stationarity via time-varying model components [4,7,10,15].

The recent developments in Bayesian modeling of environmental extremes have shown the usefulness of the Bayesian non-stationary models to stabilize inference and enhance risk prediction in data constrained contexts [4,7,10,19]. Simultaneously, the regional studies of the larger MENA region have estimated more frequent and intense heatwaves with further warming [22]. In the specific case of Iraq, the past reports found increasing tendencies in the temperature extremes, yet there are relative fewer examples of probabilistic models of the behavior of annual maximum temperature in non-stationary extreme value models [1,18]. Consequently, there is still a disjuncture between descriptive evidence of warming and formal statistical approximations of the way there has been a risk of infrequent summer heat events to have changed across Iraqi climatic regions.

The present study fills this gap by using a Bayesian non-stationary GEV model on annual summer temperature extremes based on ERA5-Land reanalysis data of 1980-2024 [17]. To represent broad spatial variation in extreme heat behavior, Iraq is subdivided into three major climatic regions namely North, Central and South. A temporal trend in the location parameter is used to achieve non-stationarity and the scale and shape parameters remain unchanged to maintain the inferential stability due to the small number of annual maxima that can be measured per region. Besides descriptive heatwave indicators, the proposed framework offers region-specific estimates of the baseline extreme heat intensity, warming trends over time, and time-varying return levels of rare summer temperature events.

The primary value of the study is thus not that extreme summer heat in Iraq is becoming more intense, but rather that it is becoming more intense and that such increase is changing the probability structure of rare events in various climatic regions. The research, which connects descriptive evidence, Bayesian non-stationary inference, and analysis at the level of returns, offers a probabilistic foundation of climate-risk assessment in one of the hottest regions of the MENA region. The rest of the paper is structured in the following way. Section 2 is a description of data, model structure, prior specification and inferential framework. Section 3 gives the empirical findings such as regional parameter estimates, warming trend, and variations in the level of returns. Section 4 ends with the primary implications, limitations and future work directions.

2. Data and Methodology

2.1 Study Area, Data Source, and Construction of Annual Summer Maxima

The research is centered on Iraq, which is a climatically varied nation with strong summer heat, and strong thermal gradient between the north and the south. Iraq was divided into three large climatic areas North ($35.1 - 37.5°N$), Central ($32.6 - 35.0°N$), and South ($29.0 - 32.5°N$) to represent the wide regional difference in extreme weather of summer. This regional division provides a climatologically meaningful and computationally practical framework for regional modeling of spatial differences in extreme heat.

Daily maximum 2-m air temperature (T_{max}) was obtained from the ERA5-Land reanalysis dataset for the period 1980–2024 [17]. ERA5-Land has been widely used in environmental and climatological analyses because it provides spatially continuous temperature fields suitable for regional-scale inference, particularly in settings where long and homogeneous station records are limited [5,17]. Because the most severe heat stress in Iraq occurs during summer, the analysis was restricted to the warm-season months from June through September.

Let X_{rtd} denote the regional daily maximum temperature for region r, year t, and summer day $d \in \mathcal{S}_t$, where $\mathcal{S}_t$ is the set of all summer days in year t. For each climatic region, grid-cell values were spatially aggregated to obtain a regional daily series, and the annual summer block maximum was then defined as

$$Y_{rt} = \max_{d \in \mathcal{S}_t} X_{rtd}, r = 1,2,3, \quad t = 1980, \ldots, 2024. \qquad (1)$$

Thus, Y_{rt} represents the annual summer maximum temperature for region r in year t. Annual block maxima are the common form of extreme value analysis and a natural step to generalized extreme value distribution modeling [3,12].

2.2 Descriptive Heatwave Indicators

A descriptive analysis summarizing the temporal evolution of heatwave frequency, persistence and intensity across Iraq was performed before fitting the extreme-value model. This preliminary step provides an empirical underpinning for evaluating whether the summer heat

extremes distribution has changed through time, and as such whether a non-stationary modeling framework is appropriate.

Three complementary heatwave definitions were considered in order to represent different dimensions of thermal stress [21-22]:

1. Absolute threshold: any day with temperature exceeding $40°C$.

2. Extreme-intensity threshold: any day on which daily maximum temperature exceeded $45°C$, representing severe heat conditions relevant to public health and infrastructure stress.

3. Percentile-based threshold: any day on which temperature exceeded the 95th percentile of the local 1981–2010 reference period [21].

Using these definitions, annual heatwave metrics were computed, including Heatwave Number (HWN), Heatwave Duration (HWD), and Heatwave Amplitude (HWA). These indicators are not part of the likelihood used in the extreme-value model; rather, they provide descriptive evidence on whether extreme summer heat in Iraq has become more frequent, more persistent, or more intense. In this sense, they complement the formal probabilistic analysis by linking observed heatwave behavior to the subsequent modeling of annual maxima.

2.3 Generalized Extreme Value Model for Annual Maxima

To model rare summer heat extremes, the annual regional maxima Y_{rt} were assumed to follow a generalized extreme value (GEV) distribution:

$$Y_{rt} \sim \text{GEV}(\mu_{rt}, \sigma_r, \xi) \tag{2}$$

where μ_{rt} is the location parameter, $\sigma_r > 0$ is the region-specific scale parameter, and ξ is the common shape parameter. The GEV distribution is the asymptotic limiting distribution of normalized block maxima and is therefore the standard model for annual extreme temperatures [3,12].

Its cumulative distribution function is

$$F(y \mid \mu, \sigma, \xi) = exp\left\{-\left[1+\xi\left(\frac{y-\mu}{\sigma}\right)\right]^{-1/\xi}\right\} \tag{3}$$

defined for values satisfying

$$1 + \xi\left(\frac{y-\mu}{\sigma}\right) > 0,\ \sigma > 0. \tag{4}$$

The corresponding density function is

$$f(y \mid \mu, \sigma, \xi) = \frac{1}{\sigma}\left[1+\xi\left(\frac{y-\mu}{\sigma}\right)\right]^{-1/\xi-1} exp\left\{-\left[1+\xi\left(\frac{y-\mu}{\sigma}\right)\right]^{-1/\xi}\right\} \tag{5}$$

For annual summer temperature extremes, the location parameter μ_{rt} characterizes the baseline magnitude of extreme heat, whereas the scale parameter σ_r quantifies year-to-year variability. The common shape parameter ξ determines the tail behavior of the distribution [3,12]. A negative estimate of ξ therefore implies a bounded upper tail, which is physically plausible for temperature extremes.

2.4 Non-Stationary Structure in the Location Parameter

Because the main scientific objective is to determine whether extreme summer heat in Iraq has intensified over time, non-stationarity was introduced through the location parameter only. Specifically, for region r and year t, the location parameter was modeled as

$$\mu_{rt} = \alpha_r + \beta(t - t_0) \tag{6}$$

where α_r is a region-specific intercept, β is a common temporal trend shared across regions, and $t_0 = 2002$ is a centered reference year introduced to improve numerical stability and reduce posterior dependence between the intercepts and the slope. This formulation also induces a hierarchical Bayesian structure, as region-specific parameters are estimated jointly with shared temporal and distributional components across the three climatic regions

Under this formulation, α_r represents the baseline magnitude of annual summer temperature extremes in region r, whereas β quantifies the average long-term rate of change in extreme summer heat across Iraq as a whole. A positive value of β indicates that annual summer extremes are intensifying over time.

To preserve inferential stability, non-stationarity was restricted to the location parameter. The scale parameter was allowed to vary by region but was assumed constant over time, while the shape parameter was taken to be common across regions and constant over time:

$$\sigma_{rt} = \sigma_r,\ \xi_t = \xi \tag{7}$$

This specification constitutes a workable compromise between flexibility and identifiability. Conversely, allowing all GEV parameters to vary both by region and time would add unnecessary complexity (and potentially unstable estimation) since each of our regions produces only one annual summer max per year. In contrast, the adopted formulation preserves the dominant climate signal we are interested in, namely a shift of extremes to higher levels over time while preserving regional differences in baseline

intensity and variability. This strategy is also consistent with common practice in non-stationary extreme-value analysis when sample size is moderate and the main interest lies in changes in the location of the extreme-value distribution rather than in fully time-varying tail behavior.

2.5 Prior Specification

A Bayesian specification was used to regularize estimation and to ensure stable inference under a relatively short annual-maxima record. Let

$$\Theta = \alpha_r, \sigma_r : r = 1,2,3;\ \beta, \xi \tag{8}$$

denote the full parameter vector.

Weakly informative priors were assigned as follows. The region-specific intercepts were given

$$\alpha_r \sim \mathcal{N}(49, 3^2),\ r = 1,2,3, \tag{9}$$

reflecting plausible baseline values for annual summer temperature extremes in Iraq without being overly restrictive. The common temporal trend was assigned

$$\beta \sim \mathcal{N}(0.03, 0.02^2) \tag{10}$$

which centers prior belief on a modest positive warming trend while still allowing the data to support weaker or stronger rates of change.

Because the scale parameter must be positive, each regional scale parameter was assigned

$$\sigma_r \sim \text{Half-Normal}(0,2),\ r = 1,2,3. \tag{11}$$

Finally, the common shape parameter was assigned

$$\xi \sim \mathcal{N}(0, 0.3^2), \tag{12}$$

which mildly regularizes tail behavior toward the Gumbel neighborhood while still allowing bounded-tail or moderately heavy-tail solutions if supported by the data.

These priors were chosen to provide mild regularization, avoid exploration of physically implausible parameter regions, and stabilize inference without imposing overly strong assumptions on the extreme-temperature process. In particular, the prior on ξ is important because the shape parameter is often weakly identified in finite samples and can otherwise lead to unstable return-level estimates.

2.6 Posterior Inference

Bayesian inference proceeds by combining the likelihood implied by the non-stationary GEV model with the prior specification above. Conditional on the observed annual maxima $y = \{y_{rt}\}$, the joint posterior distribution is

$$p(\Theta \mid y) \propto \left[\prod_{r=1}^{3}\prod_{t=1980}^{2024} f\left(y_{rt} \mid \mu_{rt}, \sigma_r, \xi\right)\right] p(\Theta), \tag{13}$$

where $f(\cdot)$ denotes the GEV density and $\mu_{rt} = \alpha_r + \beta(t - t_0)$.

Because this posterior can be analyzed in no way using analytical methods to obtain an approximation of it, we used a method called MCMC that uses sampling for estimation by implementing the method through No-U-Turn Sampler (NUTS), which is part of the Stan program. We ran four parallel chains and each of them had 4000 iterations. The first 2000 iterations were considered as "warm up" or burn-in. Finally, we computed posterior summary statistics on the remaining iterations after warm-up across all chains.

Sampling quality and convergence were assessed using standard Bayesian diagnostics, including trace plots, the potential scale reduction factor $\hat{R}$, and effective sample size measures. Values of $\hat{R}$ close to 1 together with satisfactory effective sample sizes were taken as evidence of adequate convergence and posterior exploration.

To evaluate whether allowing for temporal non-stationarity improved model fit, the non-stationary formulation was compared with a stationary alternative using the Widely Applicable Information Criterion (WAIC). Lower WAIC values indicate better expected predictive performance. In the empirical analysis, the non-stationary model was preferred, supporting the inclusion of a temporal trend in the location parameter.

2.7 Non-Stationary Return Levels

A principal quantity of interest in extreme climate analysis is the return level, defined as the magnitude of an event expected to be exceeded on average once every T years [3,12]. Under stationarity, the return level associated with a given return period remains constant over time. In a non-stationary setting, however, return levels become time-dependent because the underlying distribution evolves over time through the location parameter.

For region r, year t, and return period T, the return level $z_{rt}(T)$ is defined implicitly by

$$P_r(Y_{rt} \le z_{rt}(T)) = 1 - \frac{1}{T} \tag{14}$$

Under the GEV model, when $\xi \neq 0$, the corresponding return level is given by

$$z_{rt}(T) = \mu_{rt} + \frac{\sigma_r}{\xi}\left[\left\{-log\ \left(1-\frac{1}{T}\right)\right\}^{-\xi} - 1\right] \tag{15}$$

When the shape parameter approaches zero, the Gumbel

limiting form is used:

$$z_{rt}(T) = \mu_{rt} - \sigma_r log\left[-log\left(1-\frac{1}{T}\right)\right] \quad (16)$$

Substituting the non-stationary location model,

$$\mu_{rt} = \alpha_r + \beta(t - t_0) \quad (17)$$

Gives

$$z_{rt}(T) = \alpha_r + \beta(t - t_0) + \frac{\sigma_r}{\xi}\left[\left\{-log\left(1-\frac{1}{T}\right)\right\}^{-\xi} - 1\right],\ \xi \neq 0. \quad (18)$$

This formulation implies that return levels vary over time according to the common temporal trend $\beta(t - t_0)$, whereas regional heterogeneity is captured through the intercepts α_r and the regional scale parameters σ_r. As a result, an event corresponding to a fixed historical return period may no longer have the same rarity in a warming climate.

To propagate parameter uncertainty into return-level estimation, $z_{rt}(T)$ was computed at each posterior draw of the parameter vector. This yielded a full posterior distribution for each return level of interest, from which posterior medians and 95% highest density intervals (HDIs) were obtained. In the empirical analysis, particular emphasis was placed on the 20-year and 50-year return levels because they provide practically interpretable summaries of moderate-to-severe and rarer extreme summer heat risk across the three climatic regions of Iraq. This formulation is especially useful for climate-risk assessment because it allows a given return period to be interpreted dynamically over time, thereby quantifying how the severity of rare summer heat events has shifted under ongoing warming.

2.8 Software and Computational Environment

All analyses were conducted in R, and the Bayesian non-stationary GEV model was estimated in Stan using the No-U-Turn Sampler (NUTS) [15,20].

3. Results

3.1 Regional Heatwave Characteristics

Marked spatial contrasts in extreme summer heat are clear across Iraq, as reflected in both the descriptive indicators and the fitted model parameters. **Table 1** summarizes the three climatic regions, their spatial extent, and their ERA5-Land grid-point coverage. The South encompasses the largest grid-point coverage, while the North and Central regions also cover spatially extensive areas.

Table 1. Spatial definition of the three climatic regions used in the analysis, including latitudinal and longitudinal extent, centroids, and the number of ERA5-Land grid-points in each region.

Region	*Grid Points (Count)*	*Min Latitude*	*Max Latitude*	*Centroid Lat*	*Min Longitude*	*Max Longitude*	*Centroid Lon*
Central	2900	32.6	35	33.8	38	49.5	43.75
North	2900	35.1	37.5	36.3	38	49.5	43.75
South	4176	29	32.5	30.75	38	49.5	43.75

A distinct trend for increased extreme heat during the summers can be seen through descriptive statistics. Both average temperatures throughout each season as well as the yearly maxima have significantly increased from 1980 – 2024 in all three climatic regions see **Figure 1**. An increasing frequency of exceptionally large values of maxima were also observed after 2000 which suggests that the distribution of summer extreme values has shifted upward, as opposed to simply increasing in variance. Thus, a non-stationary extreme value model with a varying location parameter is justified.

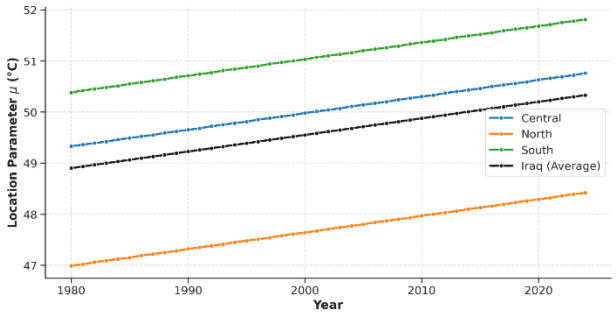


Figure 1. The trajectory of seasonal mean temperature and annual block maxima shows an increasing trend across the North, Central, and South regions of Iraq and the temporal evolution of summer temperature extremes. It is being used in Iraq from 1980 to 2024.

The long-term implications of this increased temperature can be illustrated with an example using data from **Table 2**, which presents yearly heat wave data to show how many extreme-heat days occur annually (using a high heat intensity threshold $T_{max} > 45\,°C$). This particular threshold is useful since it represents the highest danger zone for public health and infrastructure. The table indicates that while the number of extreme heat days, their length and their peak temperature vary year by year, there have been more occurrences of extreme and lengthy heat waves during the last few decades than in previous ones. As indicated in **Table 2**, maximum heat-wave intensities exceed $50\,°C$ regularly; as

such, they are now rare events rather than extreme events.

Table 2. Annual statistics of hot heatwaves ($T_{max} > 45\ °C$) occurring in Iraq (1980–2024) include HWN: HW Total Days, HWD: Max Duration, HWA: Max Intensity, and HWM: Mean.

Year	*HWN (Frequency)*	*Total Heat Days*	*HWD (Max Duration)*	*HWA (Max Intensity)*	*HWM (Mean Intensity)*
1980	5	81	56	51.27	46.79
1981	6	84	59	51.58	46.46
1982	7	78	20	51.08	46.76
1983	6	77	28	52.16	47.1
1984	5	63	37	49.53	46.49
1985	7	85	36	50.6	46.52
1986	3	89	79	51.21	46.55
1987	6	84	57	51.52	46.76
1988	3	75	59	51.86	47.39
1989	6	92	40	50.46	47.12
1990	9	90	26	51.11	46.99
1991	4	73	33	50.8	46.9
1992	8	64	22	50.17	46.33
1993	6	83	29	50.67	46.69
1994	5	77	30	50.73	46.75
1995	8	74	37	50.3	46.5
1996	5	92	72	50.9	46.88
1997	9	79	28	49.99	46.6
1998	5	101	56	52.01	47.04
1999	3	107	74	51.23	47.68
2000	3	77	63	51.92	47.72
2001	5	86	56	51.26	46.77
2002	6	81	33	50.52	47
2003	5	87	42	52.38	47
2004	5	94	75	50.44	46.88
2005	9	85	32	51.82	47.09
2006	4	104	37	50.37	47.53
2007	4	97	53	51.01	46.94
2008	5	88	61	51.04	46.99
2009	8	93	24	50.89	47.06
2010	3	97	74	52.55	47.81
2011	5	87	43	52.51	47.46
2012	6	95	53	51.39	46.68
2013	6	73	25	51.76	46.92
2014	6	88	33	51.66	47.68
2015	5	106	64	51.34	46.77
2016	4	90	80	52.87	47.28
2017	3	110	89	52.69	48.2
2018	3	112	92	51.52	47.53
2019	4	117	78	51.43	47.3
2020	7	114	50	53.1	46.92
2021	3	112	82	51.89	47.98
2022	3	105	79	52.55	47.82
2023	4	91	69	52.11	47.45

3.2 Posterior GEV Parameter Estimates

The results of the fitted non-stationary Bayesian GEV model show significant regional differences in the baseline level and variability of extreme summer temperature in Iraq. The posterior estimates are summarized graphically in **Figure 2** and **Table 3** reports posterior means and 95% highest density interval (HDI) for region-specific location and scale parameters along with the common shape parameter.

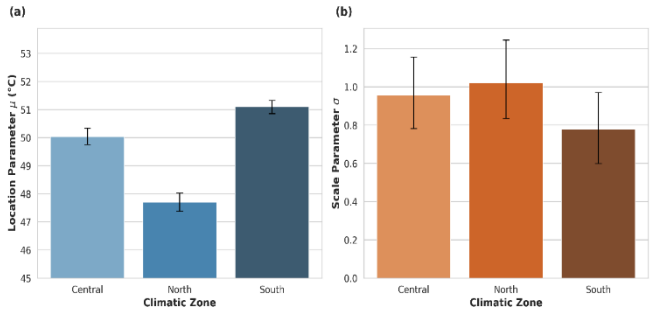


Figure 2. The parameters are illustrated in the two panels of this figure. Panel (a) shows the location parameter (μ), representing baseline extreme heat intensity. Panel (b) shows the scale parameter (σ), representing interannual variability in annual summer maxima. The error bars on the posterior estimates of the regional GEV parameters are 95% highest density intervals.

The location parameter μ, which gives the baseline magnitude of the annual summer extremes, has the highest posterior mean in the South ($51.10\ °C$; 95% HDI: $50.85 - 51.33$) followed by the Central ($50.04°C$; 95% HDI: $49.75 - 50.33$) and the North ($47.71\ °C$; 95% HDI: $47.38 - 48.02$). Consequently, the baseline degree of extreme heat is significantly.

The estimate of the scale parameter reflecting interannual variability of annual summer maxima was largest in the North ($1.02\ °C$, 95% HDI: $0.83 - 1.24$), intermediate in the Central ($0.96\ °C$, 95% HDI: $0.78 - 1.15$), and smallest in the South ($0.78\ °C$, 95% HDI: $0.60 - 0.97$) region. Thus, although southern Iraq shows the highest baseline extreme heat, northern Iraq exhibits greater year-to-year variability in annual summer maxima.

The posterior mean of the common shape parameter was negative, $\xi = -0.32$ (95% HDI: -0.45 to -0.18), indicating

a bounded upper tail. This is physically plausible for temperature extremes and supports the fitted GEV specification.

Table 3. Posterior means and 95% highest density intervals (HDIs) for the region-specific location and scale parameters, together with the common shape parameter of the non-stationary Bayesian GEV model.

Region	*Location Parameter (μ)*	*Scale Parameter (σ)*	*Shape Parameter (ξ)*
Central	50.04 (49.75, 50.33)	0.96 (0.78, 1.15)	-0.32 (-0.45, -0.18)
North	47.71 (47.38, 48.02)	1.02 (0.83, 1.24)	-0.32 (-0.45, -0.18)
South	51.10 (50.85, 51.33)	0.78 (0.60, 0.97)	-0.32 (-0.45, -0.18)

3.3 Temporal Trend in Extreme Heat

The estimated temporal trend provides strong evidence that extreme summer heat in Iraq has intensified over the study period. **Table 4** summarizes the posterior distribution of the common trend coefficient β, expressed both as an annual rate and as a decadal rate of change.

The estimated warming trend is $0.0324\,°C$/year (95% HDI: $0.0208 - 0.0441$), equivalent to roughly $0.3237\,°C$/decade (95% HDI: $0.2084 - 0.4409$), according to our calculations. Since zero is excluded from the 95% HDI, the trend is credibly positive. The increase in extreme summer temperature is therefore not just due to short-term internal variability.

As earlier evidence pointed to, the GEV coefficient results can provide formal confirmation of the non-stationary specification that the authors have adopted. In practical terms, the positive value of β implies that the distribution of annual summer temperature extremes has shifted upward over time, thereby increasing the likelihood of unusually severe heat events across Iraq.

Table 4. Posterior summary of the common temporal trend coefficient β, reported as annual and decadal rates of change with 95% highest density intervals (HDIs).

Trend Metric	*Mean Estimate*	*95% HDI Lower Bound*	*95% HDI Upper Bound*
Warming Trend per Year (°C)	0.0324	0.0208	0.0441
Warming Trend per Decade (°C)	0.3237	0.2084	0.4409

3.4 Non-Stationary Return Levels and Changes in Risk

The return-level analysis provides the clearest probabilistic interpretation of how extreme summer heat risk has evolved across Iraq under the fitted non-stationary GEV model. **Table 5** presents the numerical posterior summaries of the 20-year and 50-year return levels for the early and recent periods, **Figure 3** provides an overview of period-based differences in return levels across the three climatic regions, and **Figure 4** highlights the same contrast with explicit uncertainty intervals. Taken together, these results allow both a numerical and visual interpretation of how the severity of rare summer heat events has shifted over time.

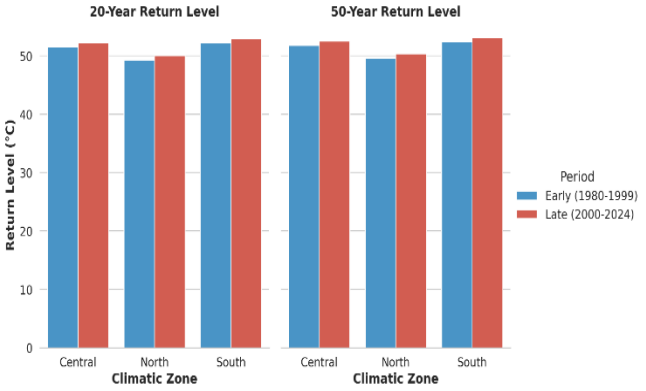


Figure 3. Comparison of posterior median non-stationary return levels between the early (1980–1999) and recent (2000–2024) periods for the 20-year and 50-year return periods across the North, Central, and South regions of Iraq.

Across all three climatic regions, both the 20-year and 50-year return levels are higher in the recent period than in the earlier period, indicating a systematic upward shift in the severity of rare summer heat events. The estimated increase is broadly consistent across regions, with return levels rising by approximately $0.7 - 0.8\,°C$ between the two reference periods. This pattern is fully consistent with the positive common temporal trend identified in the fitted model and shows that the warming signal detected in the annual maxima translates directly into higher levels of rare-event risk.

A clear regional gradient is also evident in the return-level estimates. The South consistently exhibits the highest return levels, followed by the Central region and then the North. This order agrees with the regional pattern observed in the posterior location estimates and further confirms that Southern Iraq continues to be the most vulnerable area to extreme heat during summer. Thus, the model shows that the southern area is not only the area with the highest extreme heat level but also the one exposed to the maximum extreme heat risk.

The uncertainty ranges on the estimated return levels continue to stay quite small, suggesting that the increased likelihood of very high temperatures is a feature that is strongly embedded within the posterior distribution model used, as opposed to simply being a function of certain years characterized by abnormal behavior. From an applied perspective, the findings above suggest that temperature levels associated with fairly rare occurrences during the summer season have become far more probable in recent times.

Table 5. Posterior median non-stationary return levels and 95% HDIs for the 20-year and 50-year return periods in the early (1980–1999) and recent (2000–2024) periods across the three climatic regions of Iraq.

Region	*Period Label*	*Center Year*	*Return Period (Years)*	*Return Level median (°C)*	*95% HDI Lower Bound (°C)*	*95% HDI Upper Bound (°C)*
Central	early_1980_1999	1989.5	20	51.47	51.13	51.92
Central	early_1980_1999	1989.5	50	51.77	51.39	52.32
Central	late_2000_2024	2012	20	52.2	51.85	52.69
Central	late_2000_2024	2012	50	52.5	52.09	53.11
North	early_1980_1999	1989.5	20	49.26	48.89	49.75
North	early_1980_1999	1989.5	50	49.59	49.18	50.2
North	late_2000_2024	2012	20	49.99	49.63	50.48
North	late_2000_2024	2012	50	50.32	49.91	50.94
South	early_1980_1999	1989.5	20	52.18	51.89	52.59
South	early_1980_1999	1989.5	50	52.43	52.11	52.94
South	late_2000_2024	2012	20	52.92	52.6	53.36
South	late_2000_2024	2012	50	53.16	52.79	53.73

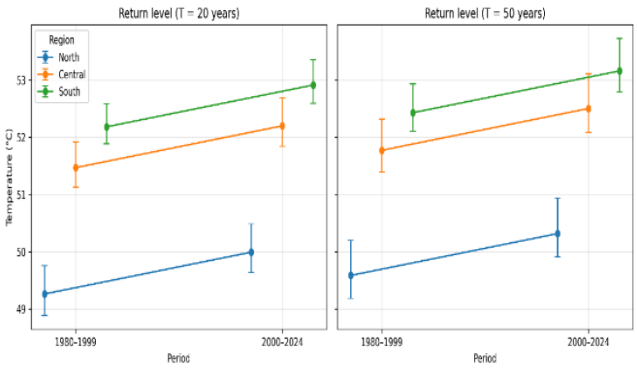
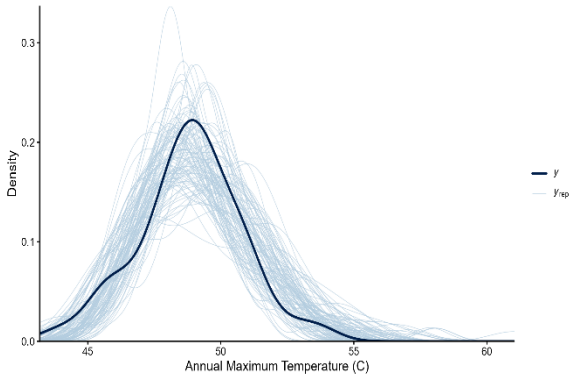


Figure 4. Regional comparison of posterior median return levels for the 20-year and 50-year return periods in the early (1980–1999) and recent (2000–2024) periods. Error bars represent 95% HDIs.

Overall, the return-level results provide consistent evidence that rare summer heat events in Iraq have become more severe over time across all three climatic regions. In this respect, **Table 5** supplies the quantitative magnitude of the change, **Figure 3** provides a visual comparison of return-level differences between the earlier and more recent periods, and **Figure 4** offers the same contrast with explicit uncertainty intervals across regions. Together, these results reinforce the conclusion that extreme summer heat in Iraq is no longer adequately characterized by a stationary framework and that climate-risk assessment must account for the ongoing upward shift in the distribution of extreme temperatures.

3.5 Model Adequacy: Posterior Predictive Check

To assess the adequacy of the fitted Bayesian non-stationary GEV model, posterior predictive checks (PPCs) were performed by generating replicated datasets from the posterior predictive distribution and comparing them with the observed ERA5-Land annual maximum temperature series. As shown in **Figure 5**, the empirical density of the observed data is closely aligned with the densities of the posterior predictive replicates, indicating that the model reproduces the main distributional features of the observed extremes reasonably well. This result provides supportive



evidence that the fitted model offers an adequate representation of the annual summer temperature maxima across the studied climatic regions.

Figure 5. Posterior predictive check (PPC) density overlay for the fitted Bayesian non-stationary GEV model. The dark curve represents the empirical density of the observed ERA5-Land annual maximum temperatures, whereas the light blue curves correspond to 100 replicated datasets simulated from the posterior predictive distribution.

Although the posterior predictive check supports the adequacy of the fitted model, robustness of the substantive conclusions must also be examined under alternative assumptions. For this reason, a sensitivity analysis was conducted with respect to model structure, prior specification, and temporal sample definition.

3.6 Sensitivity Analysis

To assess the robustness of the proposed common-trend Bayesian GEV model, three complementary sensitivity analyses were conducted. First, model-level sensitivity was evaluated by comparing the stationary and non-stationary specifications using information criteria. As shown in **Table 6**, the common-trend model yielded substantially lower values of both LOOIC and WAIC than the stationary alternative, indicating improved expected predictive performance and supporting the inclusion of temporal non-stationarity in the location parameter.

For sensitivity analysis, 90% credible intervals were reported to emphasize directional robustness across

alternative specifications, whereas 95% intervals were used in the main inferential results.

Second, prior sensitivity was evaluated by refitting the common-trend model under wider and tighter prior settings. The posterior estimate of the temporal trend parameter β remained highly stable across all prior scenarios. Under the baseline, wider, and tighter priors, the posterior mean of β was 0.0317, 0.0320, and $0.0296\ °C$ per year, respectively, and the corresponding 90% credible intervals remained entirely above zero in every case (**Table 6**; **Figure 6**). Thus, the evidence for a positive long-term intensification of extreme summer temperature is not attributable to a particular prior choice. Likewise, the posterior estimate of the shape parameter ξ remained consistently negative and showed only negligible variation across prior settings, indicating that the bounded-tail interpretation of the fitted GEV model is robust.

Third, temporal sensitivity was examined by refitting the common-trend model after excluding the most recent five years of observations, yielding a reduced sample covering 1980–2019. According to the improved model, the posterior expectation for β was still positively estimated, standing at $0.0279\ °C$ per year, with the 90% credible set ranging from 0.0170 to 0.0387 (**Table 6**; **Figure 6**). Despite being slightly lower than the corresponding figure obtained through the full dataset, the results were essentially unaffected. The posterior estimate of ξ also remained negative and stable, further supporting the robustness of the fitted extreme-value structure.

Overall, these results show that the main substantive conclusions of the study are stable to alternative model specifications, prior settings, and temporal sample choices. In particular, the positive warming trend in annual summer extremes and the negative shape parameter remain robust across all sensitivity scenarios considered

Table 6. Sensitivity analysis summary for alternative model, prior, and temporal specifications.

Analysis component	*Scenario*	*Parameter / Criterion*	*Estimate*	*90% CrI / Value*
Model comparison	Stationary model	LOOIC	385.36	—
Model comparison	Stationary model	WAIC	385.10	—
Model comparison	Common-trend model	LOOIC	355.26	—
Model comparison	Common-trend model	WAIC	355.06	—
Prior sensitivity	Baseline prior	$\boldsymbol{\beta}$	0.0317	[0.0228, 0.0409]
Prior sensitivity	Wider prior	$\boldsymbol{\beta}$	0.0320	[0.0228, 0.0415]
Prior sensitivity	Tighter prior	$\boldsymbol{\beta}$	0.0296	[0.0216, 0.0378]
Temporal sensitivity	Full period (1980–2024)	$\boldsymbol{\beta}$	0.0317	[0.0228, 0.0409]
Temporal sensitivity	Trimmed period (1980–2019)	$\boldsymbol{\beta}$	0.0279	[0.0170, 0.0387]
Prior sensitivity	Baseline prior	$\boldsymbol{\xi}$	-0.3160	[-0.4211, -0.2077]
Prior sensitivity	Wider prior	$\boldsymbol{\xi}$	-0.3191	[-0.4341, -0.1995]
Prior sensitivity	Tighter prior	$\boldsymbol{\xi}$	-0.3206	[-0.4225, -0.2142]
Temporal sensitivity	Full period (1980–2024)	$\boldsymbol{\xi}$	-0.3160	[-0.4211, -0.2077]
Temporal sensitivity	Trimmed period (1980–2019)	$\boldsymbol{\xi}$	-0.3017	[-0.4120, -0.1845]

Note: For the sensitivity analysis, intervals reported for β and ξ are 90% posterior credible intervals based on the 5th and 95th posterior quantiles, whereas the main inferential results are reported using 95% highest density intervals (HDIs).

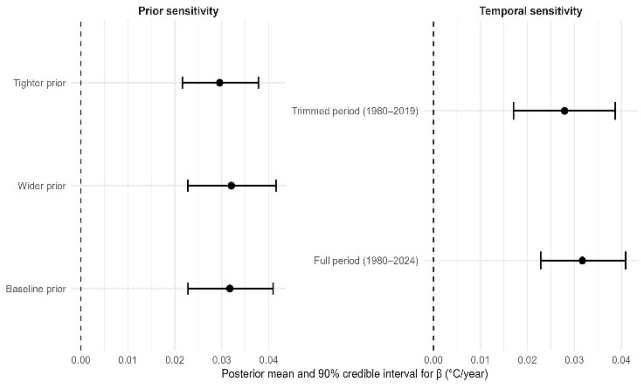


Figure 6. Sensitivity analysis for the posterior estimate of the temporal trend parameter β under alternative prior specifications and temporal sample definitions.

3.7 Interpretation in the Context of Earlier Studies

Findings from this study are generally consistent with earlier findings pointing to an increase in temperature extremes in Iraq as well as in the larger MENA region. The warming rate of approximately $0.32\ °C$ per decade is supportive of the belief that Iraq is part of the climate hotspots where increased heat is prevalent. The present study has contributed to knowledge by moving beyond merely describing trends and counting number of exceedances. Through the Bayesian non-stationary GEV framework, the analysis quantifies not only the temporal intensification of extreme heat, but also the spatial north–south gradient in

baseline risk and the upward shift in 20-year and 50-year return levels. This evidence suggests that very rare high summer temperatures during the historical period have become realistic within the current climate setting, and their impacts are particularly evident in the southern region of Iraq. The outcome of the sensitivity test corroborates this analysis since it shows that the major inferences from the study remain valid regardless of how the models are specified.

Conclusion

This research paper investigates the trend of extreme heat risks in summer seasons in Iraq based on ERA5-Land datasets from 1980 through 2024 using Bayesian non-stationary generalized extreme value theory. The results show that extreme heat risks increase, with significant north-south differences for baseline extreme heat and warming rates of about $0.32\ °C$ per decade. Southern Iraq experiences the highest baseline extreme heat conditions compared to northern Iraq, which exhibits higher variations from year to year. Results from the return level analysis demonstrate a significant increase in the magnitude of the extreme summer temperatures at 20- and 50-year return levels in recent decades, by an approximate rise $of\ 0.7 - 0.8\ °C$ in the three climatic zones studied. The results imply that extreme summer temperatures which were once rare occurrences in the past now seem more probable in current climate conditions. The overall study reveals the potential importance of using non-stationary extreme value modeling with a Bayesian framework in climate risk assessment, especially when faced with limited data sets in a region like Iraq. Sensitivity tests demonstrate that the key findings remain robust under different model formulations and priors. In total, the credible increasing trend in time, the strong geographical pattern in baseline extremes, and the increase in return level provide a consistent statistical framework for using non-stationary extreme value models in risk assessment and climate adaptation strategies in heat-vulnerable regions. The drawback of the present study is that the model performance was mainly based on convergence checks, comparison of information criteria, and posterior predictive checks, whereas future studies need to focus on posterior predictive checks and sensitivity analyses.

Conflict of interest

None.

References

[1] Al-Sudani, Z. A.: Long-term trends of temperature extremes in Iraq during 1982–2020. Iraqi Journal of Science 61(11), 2835–2850 (2020).

[2] Calvin, K. et al.: IPCC, 2023: Climate Change 2023: Synthesis Report. Contribution of Working Groups I II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, H. Lee and J. Romero (eds.)]. IPCC, Geneva, Switzerland (2023).

[3] Coles, S.: An introduction to statistical modeling of extreme values. Springer, London (2001).

[4] Cooley, D., Nychka, D., Naveau, P.: Bayesian spatial modeling of extreme precipitation return levels. Journal of the American Statistical Association 102(479), 824–840 (2007).

[5] Cornes, R. C., Van Der Schrier, G., Van Den Besselaar, E. J., Jones, P. D.: An ensemble version of the E-OBS temperature and precipitation data sets. Journal of Geophysical Research: Atmospheres 123(17), 9391–9409 (2018).

[6] Croce, P., Formichi, P., Landi, F.: A Bayesian hierarchical model for climatic loads under climate change. In UNCECOMP 2019: 3rd International Conference on Uncertainty Quantification in Computational Sciences and Engineering: proceedings (pp. 299–308). National Technical University of Athens (NTUA) (2019).

[7] Dyrrdal, A. V., Lenkoski, A., Thorarinsdottir, T. L., Stordal, F.: Bayesian hierarchical modeling of extreme hourly precipitation in Norway. Environmetrics 26(2), 89–106 (2015).

[8] Gelman, A., Shalizi, C. R.: Philosophy and the practice of Bayesian statistics. British Journal of Mathematical and Statistical Psychology 66(1), 8–38 (2013).

[9] Haylock, M. R., Hofstra, N., Klein Tank, A. M. G., Klok, E. J., Jones, P. D., New, M.: A European daily high-resolution gridded data set of surface temperature and precipitation for 1950–2006. Journal of Geophysical Research: Atmospheres 113(D20) (2008).

[10] Huser, R., Wadsworth, J. L.: Advances in statistical modeling of spatial extremes. Computational Statistics 14(1), e1537, Wiley Interdisciplinary Reviews (2022).

[11] International Energy Agency: National climate resilience assessment for Iraq. IEA, Paris (2025).

[12] Katz, R. W.: Statistics of extremes in climate change. Climatic change 100(1), 71–76 (2010).

[13] Lin, J.: An integrated procedure for Bayesian reliability inference using MCMC. Journal of Quality and Reliability Engineering 2014(1), 264920 (2014).

[14] Martins, E. S., Stedinger, J. R.: Generalized maximum-likelihood generalized extreme-value quantile estimators for hydrologic data. Water Resources Research 36(3), 737–744 (2000).

[15] McElreath, R.: Statistical rethinking: A Bayesian course with examples in R and Stan. Chapman and Hall/CRC (2018).

[16] Milly, P. C., Betancourt, J., Falkenmark, M., Hirsch, R. M., Kundzewicz, Z. W., Lettenmaier, D. P., Stouffer, R. J.: Stationarity is dead: Whither water management? Science 319(5863), 573–574 (2008).

[17] Muñoz-Sabater, J. et al.: ERA5-Land: A state-of-the-art global reanalysis dataset for land applications. Earth system science data 13(9), 4349–4383 (2021).

[18] Salman, S. A., Shahid, S., Ismail, T., Chung, E. S., Al-Abadi, A. M.: Long-term trends in daily temperature extremes in Iraq. Atmospheric research 198, 97–107 (2017).

[19] Sampaio, J., Costa, V.: Bayesian regional flood frequency analysis with GEV hierarchical models under spatial dependency structures. Hydrological Sciences Journal 66(3), 422–433 (2021).

[20] Stan Development Team: RStan: the R interface to Stan (R package version 2.32.6) (2024).

[21] Perkins, S. E., Alexander, L. V.: On the measurement of heat waves. Journal of climate 26(13), 4500–4517 (2013).

[22] Zittis, G. et al.: Business-as-usual will lead to super and ultra-extreme heatwaves in the Middle East and North Africa. Npj Climate and Atmospheric Science 4(1), 20 (2021).

[23] Vehtari, A., Gelman, A., Gabry, J.: Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. Statistics and computing 27(5), 1413–1432 (2017).