Environmental Research Communications

PAPER

Statistical modeling of monthly maximum temperature in Senegal

Kouakou Innocent N’dri¹ and Saralees Nadarajah²∗

¹ AIMS Senegal, Mbour-Thies, Senegal
² University of Manchester, Manchester M13 9PL, United Kingdom

∗ Author to whom any correspondence should be addressed.

E-mail: kouakou.i.ndri@aims-senegal.org and mbbsssn2@manchester.ac.uk

Keywords: generalized extreme value distribution, maximum likelihood, return level

Abstract

We provide the first statistical analysis of maximum temperature in Senegal. The data are from twelve stations spread across Senegal. The generalized extreme value distribution was fitted to maximum temperature by the method of maximum likelihood. Probability and quantile plots showed that the generalized extreme value distribution provided an adequate fit for all stations. The vast majority of stations did not exhibit significant trends in temperature. Three of the stations exhibited positive trends in temperature. Estimates of return levels are given.

1. Introduction

Nearly 75 percent of the population in Senegal works in the agricultural sector, which is regularly threatened by inclement weather such as droughts and other climate changes. Droughts in Senegal have been occurring every two or three years since the 1970s. Along the Senegalese coastline between Saint Louis in the northwest and Ziguinchor in the southwest (passing through Dakar, Mbour and Fatick), studies have shown that each year is warmer than the previous one and that all the months are affected by warming. Hence, it is important that an assessment is made of the extreme values of temperature.

The generalized extreme value (GEV) distribution is a traditional and a most popular model for extreme values. The GEV distribution has been applied to data from many countries. But there have been only three papers on the application of the GEV distribution to climate data from Senegal. Sarr et al (2015) fitted the GEV distribution to extreme precipitation data from six stations in Senegal. They found that projected changes in extreme precipitations are not consistent across stations and return periods. Sane et al (2018) computed intensity-duration-frequency rainfall curves by fitting the GEV distribution to data from fourteen stations in Senegal. The fitted curves were extrapolated to give a spatial map for Senegal. Wilcox et al (2018) fitted the GEV distribution to extreme floods in the period 1950–2015 for seven tributaries in the Sudan-Guinean part of the Senegal River basin and four data sets in the Sahelian part of the Niger River basin. For the Senegal basin, stations switched from a decreasing streamflow trend to an increasing streamflow trend in the early 1980s. In the Niger basin the trend was generally positive since the 1970s. None of the stated papers are about extreme temperature.

But there have been several papers on temperature in Senegal. Vukovich et al (1987) derived surface-temperature and albedo relationships in Senegal using satellite data. Thiam and Singh (2002) performed space-time-frequency analysis of rainfall, runoff and temperature in the Casamance River basin, southern Senegal. Fall et al (2006) presented a geographical information systems-based analysis of monthly rainfall (twenty stations) and mean temperature (twelve stations) for 1971 to 1998 in Senegal. Stisen et al (2007) estimated diurnal air temperature using data in West Africa during the 2005 rainy season. Aifa and Dabo (2015) studied microstructures and temperature variability during the Eburnean deformations in the Dalema area, eastern Senegal. Djaman et al (2017) investigated trends in annual precipitation, sunshine duration, wind speed, annual mean minimum temperature, monthly mean minimum temperature, annual mean maximum temperature, monthly mean maximum temperature and relative humidity for six locations in Senegal for 1950–2000. Manzanas (2017) tested the suitability of statistical downscaling approaches to generate seasonal forecasts of daily maximum temperature and daily maximum precipitation for stations in Senegal during 1979–2000. Brotem and Brooks (2018) examined
obstacles to rural livelihood adaptations to hotter 21st century temperatures in eastern Senegal. Sambou et al. (2020) studied springtime heat wave occurrences over Senegal using data from 12 stations. Sambou et al. (2021) studied long-term (1950–2100) observed and projected changes in springtime heat waves in Sahel, Senegal, and three thermally-coherent zones within Senegal. Yet again none of these papers apply statistical models for extreme temperature in Senegal.

The aim of this paper is to provide the first statistical analysis of extreme values of temperature in Senegal.

We will be able to answer the following questions and more: What are the hottest areas with respect to extreme temperature? What are the coolest areas with respect to extreme temperature? The answers to these questions and more could lead to actions (for example, increased agricultural production in coolest areas and planting of crops withstanding droughts in hottest areas) which may be of help to improve the economy of Senegal.

The contents of the paper are organized as follows. Section 2 describes the data from twelve locations in Senegal: Dakar, CapSkiring, Diourbel, Kolda, Kedougou, Zinguinchor, Saintlouis, Matam, Tambacounda, Linguere, Podor and Kaolack. Section 3 describes the method used to analyze the data. Section 4 presents the results of the method and their discussion. The paper is concluded in section 5.

2. Data

The data are monthly temperature in centigrade for twelve stations in Senegal. The station names and years of record are given in Table 1. The location of the stations are shown in Figure 1. We see that the stations give a good representation of the geography of Senegal. The data were obtained from the Department of Meteorology in Dakar.

We take monthly maximum temperature for each year as the extreme value. It was computed as the maximum of the twelve monthly values. Some summary statistics (mean, median, skewness, kurtosis, standard deviation, range, minimum and maximum) of the monthly maximum temperature are also shown in Table 1.

The largest monthly maximum temperature varies between 29.30 Celsius and 37.32 Celsius. The largest temperature of 37.32 Celsius was observed in Matan. The smallest temperature was observed in Dakar. The mean values for Kedougou, Kolda and Tambacounda are smaller than their median values, which indicates that the monthly maximum temperature are positively skewed for these locations. The mean values for CapSkirring, Dakar, Diourbel, Kaolack, Linguere, Matam, Podor, Saintlouis and Ziguinchor are larger than their median values, which indicates that the monthly maximum temperature are negatively skewed for these locations. The kurtosis values for all but two of the locations are less than 3, which indicates that their distributions are lighter than the normal distribution. The kurtosis values for Diourbel and Kaolack are larger than 3, which indicates that their distributions are heavier than the normal distribution. Matam has the largest standard deviation with a value of 3.331 881. Zinguinchor has the smallest standard deviation.

3. Method

Let $X$ denote a random variable representing the monthly maximum temperature. According to extreme value theory (see Leadbetter et al. 1983, Resnick 1987 and Embrechts et al. 1997), the cumulative distribution function of $X$ can be approximated by

| Station        | Count | Mean | Min  | Max  | Median | SD       | Coefficient of variation | Skewness | Kurtosis |
|----------------|-------|------|------|------|--------|---------|-------------------------|----------|----------|
| CapSkiring     | 239   | 26.18| 22.45| 29.30| 26.90  | 1.848   | 0.070                   | 0.608    | 0.2865648|
| Dakar          | 239   | 24.75| 19.11| 29.32| 24.84  | 2.812   | 0.113                   | 0.632    | 1.596    |
| Diourbel       | 213   | 28.77| 23.50| 33.80| 29.15  | 1.794   | 0.062                   | 0.362    | 1.576496 |
| Kaolack        | 239   | 28.90| 24.04| 32.98| 29.11  | 1.498   | 0.051                   | 0.846    | 1.251    |
| Kedougou       | 224   | 29.48| 23.30| 35.57| 28.76  | 2.700   | 0.091                   | 1.614    | 2.234    |
| Kolda          | 220   | 28.76| 23.18| 34.10| 28.41  | 2.380   | 0.082                   | 2.787    | 2.437    |
| Linguere       | 239   | 29.70| 23.08| 34.50| 29.88  | 2.421   | 0.081                   | 5.437    | 2.615    |
| Matam          | 239   | 30.69| 23.42| 37.32| 30.89  | 3.331   | 0.108                   | 5.552    | 2.150    |
| Podor          | 239   | 29.64| 21.76| 35.20| 30.72  | 3.187   | 0.107                   | 5.553    | 2.700    |
| Saintlouis     | 239   | 26.02| 21.32| 30.73| 25.79  | 2.351   | 0.090                   | 3.864    | 1.723    |
| Tambacounda    | 239   | 30.23| 25.44| 35.45| 29.63  | 2.508   | 0.083                   | 7.455    | 2.024    |
| Zinguinchor    | 239   | 27.50| 23.84| 30.35| 27.64  | 1.262   | 0.045                   | 2.785    | 2.985    |

Table 1. Descriptive statistics of monthly maximum temperature data.
\[ F_X(x) = \exp \left[ - \left(1 + \frac{x - \mu}{\xi \sigma} \right)^{-1/\xi} \right] \]  

(1)

for \( \mu - \sigma/\xi \leq x < \infty \) if \( \xi > 0 \), \(-\infty < x < \infty \) if \( \xi = 0 \) and \(-\infty < x \leq \mu - \sigma/\xi \) if \( \xi < 0 \), where \(-\infty < \mu < \infty \) denotes a location parameter, \( \sigma > 0 \) denotes a scale parameter and \(-\infty < \xi < \infty \) denotes a shape parameter. Note that if \( \xi > 0 \) then \( X \) has a heavy tail bounded below by \( \mu - \sigma/\xi \). If \( \xi < 0 \) then \( X \) has a short tail bounded above by \( \mu - \sigma/\xi \).

The distribution in (1) is the GEV distribution. The GEV distribution was fitted to the data in section 2 by the method of maximum likelihood. Suppose \( x_1, x_2, \ldots, x_n \) is an enumeration of the data in section 2. The maximum likelihood estimates of \( \mu, \sigma \) and \( \xi \) were obtained by maximizing

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

over all possible values of \( \mu, \sigma \) and \( \xi \). The maximum likelihood estimates are the values of \( \mu, \sigma \) and \( \xi \) corresponding to the maximum of \( L(\mu, \sigma, \xi) \). The maximization was performed using the command \texttt{fgev} in the R package \texttt{evd} (Stephenson 2018, R Core Team 2022). Other distributions (for example, the normal distribution) may provide better fits to the monthly maximum temperature. But the GEV distribution is theoretically justified.

Let \( \hat{\mu}, \hat{\sigma} \) and \( \hat{\xi} \) denote the maximum likelihood estimates of \( \mu, \sigma \) and \( \xi \), respectively. A quantity of interest based on (1) is the \( T \)-year return level loosely interpreted as the monthly maximum temperature expected on average once in every \( T \) years. Let \( x_T \) denote the \( T \)-year return level corresponding to (1). It must satisfy

\[ F_X(x_T) = 1 - \frac{1}{T}. \]  

(2)

Inverting (2), we obtain

\[ x_T = \hat{\mu} + \frac{\hat{\sigma}}{\hat{\xi}} \left\{ -\log \left(1 - \frac{1}{T} \right) \right\}^{-\hat{\xi}} - 1, \]  

(3)

which coincides with equation (3.4) in Coles (2001).
The GEV distribution was fitted to the monthly maximum temperature data from each of the twelve stations. The estimates, standard errors and 95 percent confidence intervals for the parameters of the GEV distribution are shown in Table 2.

The shape parameter estimate is not significantly different from zero for CapSkirring, Dakar, Diourbel, Kaolack, Matam, Saintlouis and Ziguinchor. The shape parameter estimate is significantly negative for Kedougou, Kolda, Linguere, Podor and Tambacounda. The monthly maximum temperature is bounded above by the probable maximum, \( \tilde{\mu} - \tilde{\sigma}\, / \tilde{\xi} \), for these five locations. \( \tilde{\mu} - \tilde{\sigma}\, / \tilde{\xi} \) takes the values 35.6, 34.2, 34.7, 35.8 and 35.7 for Kedougou, Kolda, Linguere, Podor and Tambacounda, respectively.

The largest of the probable maximum of monthly maximum temperature is for Podor, and the second largest of the probable maximum of monthly maximum temperature is for Tambacounda. The smallest of the probable maximum of monthly maximum temperature is for Koda. The second smallest of the probable maximum of monthly maximum temperature is for Linguere.

The standard errors appear smallest for Capskiring with respect to all three parameters. The standard errors appear largest for Dakar with respect to the shape parameter and largest for Kedougou with respect to the scale and location parameters.

The fit of the GEV distribution for each station was checked by probability plots and quantile plots. The plots are shown in figures 2 and 3 for the twelve stations.

Probability plots are plots of \( \hat{F}(x_i) \), the observed probabilities, versus \( i/(n + 1) \), the expected probabilities, where \( x_{(1)} \leq x_{(2)} \leq \cdots \leq x_{(n)} \) are the data arranged in increasing order and

4. Results and discussion

| Table 2. Estimates, standard errors and 95 percent confidence intervals for the parameters of the GEV distribution for monthly maximum temperature. |
|-------------------------------------------------|
| Parameter estimates | \( \xi \) | \( \sigma \) | \( \mu \) |
|---------------------|----------|----------|----------|
| Capskiring           | 0.1592   | 0.2275   | 28.5170  |
| Standard errors      | 0.109 09 | 0.036 78 | 0.055 32 |
| Confidence interval (95%) | (-0.373,0.055) | (0.155,0.300) | (28.409,28.625) |
| Dakar                | 0.1309   | 0.2606   | 28.0771  |
| Standard errors      | 0.190 56 | 0.051 28 | 0.066 69 |
| Confidence interval (95%) | (-0.243,0.504) | (0.160,0.361) | (27.946,28.208) |
| Diourbel             | -0.07226 | 0.775 27 | 30.534 0 |
| Standard errors      | 0.1136   | 0.1305   | 0.1938   |
| Confidence interval (95%) | (-0.295,0.150) | (0.519,1.031) | (30.155,30.914) |
| Kaolack              | -0.01117 | 0.612 15 | 30.592 43 |
| Standard errors      | 0.1581   | 0.1169   | 0.1939   |
| Confidence interval (95%) | (-0.391,0.369) | (0.383,0.841) | (30.283,30.902) |
| Kedougou             | -0.8116  | 1.3575   | 33.9169  |
| Standard errors      | 0.1581   | 0.2871   | 0.3255   |
| Confidence interval (95%) | (-1.121,0.501) | (0.795,1.920) | (33.279,34.555) |
| Kolda                | -0.5721  | 1.2371   | 32.0670  |
| Standard errors      | 0.1188   | 0.2173   | 0.2965   |
| Confidence interval (95%) | (-0.805,0.339) | (0.811,1.663) | (31.486,32.648) |
| Linguere             | -0.5436  | 0.9287   | 32.9435  |
| Standard errors      | 0.1990   | 0.1774   | 0.2267   |
| Confidence interval (95%) | (-0.855,0.232) | (0.581,1.276) | (32.499,33.388) |
| Matam                | -0.08702 | 0.524 84 | 35.621 19 |
| Standard errors      | 0.182 01 | 0.096 46 | 0.133 82 |
| Confidence interval (95%) | (-0.444,0.270) | (0.336,0.714) | (35.359,35.883) |
| Podor                | -0.3548  | 0.9097   | 33.2177  |
| Standard errors      | 0.1494   | 0.1996   | 0.2242   |
| Confidence interval (95%) | (-0.648,0.062) | (0.597,1.222) | (32.778,33.657) |
| Saintlouis           | 0.1548   | 0.3184   | 29.1167  |
| Standard errors      | 0.158 37 | 0.060 53 | 0.079 08 |
| Confidence interval (95%) | (-0.156,0.465) | (0.200,0.437) | (28.962,29.272) |
| Tambacounda          | -0.3878  | 0.5078   | 34.3898  |
| Standard errors      | 0.1530   | 0.0914   | 0.1252   |
| Confidence interval (95%) | (-0.688,0.088) | (0.329,0.687) | (34.144,34.635) |
| Ziguinchor           | 0.2322   | 0.2988   | 28.9990  |
| Standard errors      | 0.158 37 | 0.060 53 | 0.079 08 |
| Confidence interval (95%) | (-0.156,0.465) | (0.200,0.437) | (28.962,29.272) |
Figure 2. Observed probabilities versus expected probabilities (plotted as dots) for the fit of the GEV distribution for monthly maximum temperature from the locations. The diagonal lines correspond to the equality of probabilities. The dashed lines are 95 percent simulated confidence intervals.

Quantile plots are plots of $x_i$, the observed quantiles, versus $\hat{F}^{-1}(i/(n + 1))$, the expected quantiles, where $\hat{F}^{-1}(\cdot)$ denotes the inverse function of $\hat{F}(\cdot)$. The diagonal straight lines in figure 2 correspond to the observed and expected probabilities being equal. The diagonal straight lines in figure 3 correspond to the observed and expected quantiles being equal. The dashed lines in figures 2 and 3 are the 95 percent simulated confidence intervals. The confidence intervals in figure 2 were computed as follows:

$$\hat{F}(x) = \exp \left[ - \left(1 + \frac{\xi x - \hat{\mu}}{\hat{\sigma}}\right)^{-\frac{1}{\xi}} \right].$$
(i) simulate a random sample of size $n$ from $\hat{F}$;
(ii) refit the GEV distribution to the sample and let $\hat{\mu}$, $\hat{\sigma}$ and $\hat{\xi}$ denote the parameter estimates;
(iii) compute

$$\bar{F}(x_{(i)}) = \exp \left[ - \left( 1 + \frac{x_{(i)} - \hat{\mu}}{\hat{\sigma}} \right)^{-\frac{1}{\hat{\xi}}} \right]$$

(4)
for $i = 1, 2, \ldots, n$;

(iv) repeat steps i)–iii) 10 000 times, giving 10 000 values for $\hat{F}(x_{(i)})$;

(v) compute the empirical distribution function of the 10 000 values in step iv), denoting it by $\hat{F}$;

(vi) compute $\hat{F}^{-1}(0.025)$ and $\hat{F}^{-1}(0.725)$, where $\hat{F}^{-1}(\cdot)$ denotes the inverse function of $\hat{F}(\cdot)$;

(vii) plot $(\hat{F}^{-1}(0.025), \hat{F}^{-1}(0.725))$, the 95 percent simulated confidence interval, versus $x_{(i)}$ for $i = 1, 2, \ldots, n$.

The confidence intervals in figure 3 were computed as follows:

(i) simulate a random sample of size $n$ from $\hat{F}$;

(ii) refit the GEV distribution to the sample and let $\hat{\mu}$, $\hat{\sigma}$ and $\hat{\zeta}$ denote the parameter estimates;

(iii) compute $\hat{F}^{-1}(i/(n + 1))$ for $i = 1, 2, \ldots, n$, where $\hat{F}^{-1}(\cdot)$ denotes the inverse function of (4);

(iv) repeat steps i)–iii) 10 000 times, giving 10 000 values for $\hat{F}^{-1}(i/(n + 1))$;

(v) compute the empirical distribution function of the 10 000 values in step iv), denoting it by $\hat{F}$;

(vi) compute $\hat{F}^{-1}(0.025)$ and $\hat{F}^{-1}(0.725)$, where $\hat{F}^{-1}(\cdot)$ denotes the inverse function of $\hat{F}(\cdot)$;

(vii) plot $(\hat{F}^{-1}(0.025), \hat{F}^{-1}(0.725))$, the 95 percent simulated confidence interval, versus $x_{(i)}$ for $i = 1, 2, \ldots, n$.

The closer the plotted points (in figures 2 and 3) are to the diagonal lines the better the fit. The plotted points must lie within the simulated confidence intervals for the fit to be considered adequate. Hence, the fit of the GEV distribution for monthly maximum temperature from the twelve locations is adequate.

Having checked the goodness of fit, we computed (5) for every station and a range of values of $T$. Plots of $x_T$ versus $T = 2, 3, \ldots, 100$ are shown in figure 4.

As expected, the return level estimates increase with the return period. The return level estimates are largest for Matam, second largest for Kedougou, third largest for Tambacounda, fourth largest for Podor, fifth largest

![Figure 4. Estimates of $x_T$ versus $T = 2, 3, \ldots, 100$.](image-url)
The data that support the findings of this study are available upon reasonable request from the authors.
Authors contributions

KIN initiated, performed the analysis and designed the study. SN wrote and edited the draft.

Code availability

The code can be obtained by contacting the corresponding author.

Ethical approval

All authors kept the ‘Ethical Responsibilities of Authors’.

Consent to participate

All authors gave explicit consent to participate in this study.

Consent to publish

All authors gave explicit consent to publish this manuscript.

Competing interests

Authors declare no conflicts of interest.

ORCID iDs

Saralees Nadarajah https://orcid.org/0000-0002-0481-0372

References

Aifa T and Dabo M 2015 Microstructures and temperature variability during the Eburnean deformations in the Dalema area, Eastern Senegal Arabian J. Geosci. 8 677–89

Brottem L and Brooks B 2018 Crops and livestock under the sun: Obstacles to rural livelihood adaptations to hotter 21st century temperatures in eastern Senegal Land Degrad. Dev. 29 118–26

Coles S G 2001 An Introduction to Statistical Modeling of Extreme Values (New York, NY: Springer)

Djaman K, Balde A B, Rudnick D R, Ndaiya O and Irmak S 2017 Long-term trend analysis in climate variables and agricultural adaptation strategies to climate change in the Senegal River Basin Int. J. Climatol. 37 2873–88

Embrechts P, Klüppelberg C and Mikosch T 1997 Modelling Extremal Events for Insurance and Finance (Berlin: Springer)

Fall S, Niyogi D and Semazzi F HM 2006 Analysis of mean climate conditions in Senegal (1971–98) Earth Interact. 10 5

Leadbetter M R, Lindgren G and Rootzén H 1983 Extremes and Related Properties of Random Sequences and Processes (New York, NY: Springer)

Manzanas R 2017 Assessing the suitability of statistical downscaling approaches for seasonal forecasting in Senegal Atmos. Sci. Lett. 18 381–6

R Core Team 2022 R: A Language and Environment for Statistical Computing (Vienna, Austria: R Foundation for Statistical Computing)

Resnick S I 1987 Extreme Values, Regular Variation and Point Processes (New York, NY: Springer)

R Core Team 2022 R: A Language and Environment for Statistical Computing (Vienna, Austria: R Foundation for Statistical Computing)

Sambou M J G, Janicot S, Pohl B, Badiane D, Dieng A L and Gaye A 2020 Heat wave occurrences over Senegal during spring: regionalization and synoptic patterns Int. J. Climatol. 40 440–57

Sambou M J G, Pohl B, Janicot S, Famin A M L, Roucou P, Badiane D and Gaye A T 2021 Heat Waves in spring from Senegal to Sahel: evolution under climate change Int. J. Climatol. 41 6238–53

Sane Y et al 2018 Intensity-duration–frequency (IDF) rainfall curves in Senegal Natural Hazards and Earth System Sciences 18 1849–66

Sarr M A, Seidou O, Tramblay Y and EI Adlouni S 2015 Comparison of downscaling methods for mean and extreme precipitation in Senegal Journal of Hydrology-Regional Studies 4 369–85

Stephenson A 2018 evd: functions for extreme value distributions R package version 2.3-3

Stisen S, Sandholt I, Norgaard A, Fensholt R and Eklundh L 2007 Estimation of diurnal air temperature using MSG SEVIRI data in West Africa Remote Sens. Environ. 110 262–74

Thiam E H I and Singh V P 2002 Space-time–frequency analysis of rainfall, runoff and temperature in the Casamance River basin, southern Senegal, West Africa Water SA 28 259–70

Vukovich F M, Toll D L and Murphy R E 1987 Surface-temperature and albedo relationships in Senegal derived from NOAA-7 satellite data Remote Sens. Environ. 22 41–21

Wilcox C, Vischel T, Panthou G, Bodian A, Blanchet J, Descreix L, Quatrix G, Casse C, Tanimoune B and Kone S 2018 Trends in hydrological extremes in the Senegal and Niger Rivers J. Hydrol. 566 531–45