Statistical modeling of monthly rainfall variability in Soummam watershed of Algeria, between 1967 and 2018

Amir Aieb | Khalef Lefsih | Marco Scarpa | Brunella Bonaccorso | Nicola Cicero | Omar Mimeche | Khodir Madani

1Laboratoire de Biomathématiques, Biophysique, Biochimie, et Scientométrie, Université de Bejaia, Bejaia, Algérie
2Department of Engineering, Department BioMorf, University of Messina, Messina, Italy
3Laboratoire de recherche en Hydraulique Appliquée et Environnement, Université de Bejaia, Bejaia, Algérie

Correspondence
Amir Aieb, Laboratoire de Biomathématiques, Biophysique, Biochimie, et Scientométrie, Université de Bejaia, 06000 Bejaia, Algérie.
Email: amir18informatique@gmail.com

Abstract
The monthly precipitations obtained during 51 years of measurement in 24 stations of Soummam watershed in Algeria were analyzed to describe rainfall trends and aridity state of the area using statistical modeling. The choice of distribution laws was justified by comparing fitting results of different distributions laws used in literature reviews. Hence, the $p$ values proved that Generalized Extreme Value, Weibull (3) and Logistic the distribution law are more adequate to analyze rainfall frequencies in different part of the watershed. The diagnostic given by Q-Q plot, P-P plot and survival regression curve showed the period of wetness and dryness in the northeastern and the southwestern part of the watershed, respectively. Moreover, the study given by the De Martonne index explains the consequences of climate change by a new form of aridity in the watershed between 1994 and 2018.
Recommendations for Resource Managers

- The annual rainfall of Soummam watershed has a moderate and irregular rainfall distribution between 1967 and 2018.
- Using distribution function on monthly rainfall in each bioclimatic floor to analyze the trend of rainfall frequency gives a spatio-temporal description of climate in the area.
- Fitting by Kolmogorov-Smirnov test allows us to choose generalized extreme value, Weibull (3) and Logistic for modeling monthly rainfall variability in each part of the watershed.
- The diagnostic obtained by P-P plot, Q-Q plot and survival regression curve proved a change of aridity in the northeastern and southwestern part of the watershed between 1994 and 2018.

KEYWORDS
aridity index, bioclimatic floor, fitting test, monthly rainfall, statistical modeling, watershed

1  |  INTRODUCTION

Climate change has become a concrete study; it is measured by changes in temperature, precipitation, and other parameters. According to the scientific community reports, the global climate demonstrates a big change, and the intergovernmental panel on climate change (IPCC) showed that the global average of temperature trend from 1880 to 2012 is 0.85°C, which had a degree of uncertainty between 0.65°C and 1.06°C (GIEC et al., 2013). During the last decade (2003–2012), it was shown a rise on temperature average, which arrived to +0.78°C (for a variation range between 0.72°C and 0.85°C). However, the World of Meteorological Organization (WMO) considers the period of 2011–2015 as the hottest on record, especially in 2015, it was reported as the hottest year since modern observations began in the late of 19th century (WMO, 2016). The global precipitation evolution is a subject of measurement, which depends mainly on spatial and temporal climatic variability, this characteristic raises a problem on the acceleration and the shift of the hydrological cycle, such as a simple increase of precipitations followed by a substantial elevation of temperature can cause a very important evaporation especially in the regions of the world favorable to such a climatic evolution (Nouaceur & Murărescu, 2016). Different studies have been done to understand the concept of rainfall evolution in several sites of the world. They proved that climate change is more expressed in the wettest conditions (Alexander et al., 2006; Dore, 2005), where it was observed between 1991 and 2010 an increase in rainfall followed by a repetition of extremes (New et al., 2001; Planton et al., 2005). Climate change has a direct implication in natural disasters, it depends mainly on the large differences of temperature and rainfall variability, the climate anomalies have consequences on hurricanes and floods in different regions (Hailemariam, 1999). The drought that affects African countries is a good example of the consequences of climate change and its impacts on the availability of water resources, especially water reserves; this issue deserves great efforts to exploit water and improve food security in these
countries (Faramarzi et al., 2013). In the last years, Algeria has been also a victim of climate change problems. The Algerian authorities reported two natural catastrophes, which affected the country in 2001 and 2008, known by the floods of Bab El Oued and Ghardaia, respectively. It is caused by the big change of rainfall variability, but no work has shown the main variable that produced this phenomenon, it can be explained only by modeling rainfall records during the period that showed breaks of frequency and the high rainfall variability (Boudrissa et al., 2017).

The climate risk assessment depends mainly on the variability of rainfall extremes. One of the fundamental steps to avoid human and material damage is the need of a good knowledge and the mechanisms which can be taken to analyze climate change, then to make an in-depth study about a space and temporal distribution of climatic variables (rainfall, temperature); consequently, one can make a forecast about disasters occurrence and to take appropriate measures for mitigating the negative consequences and to reduce the impact of unforeseen events (Stern & Cooper, 2011). In literature, several works have been done to parameterize the adequate mathematical models which express variability of climatic parameters in several sites of the world. The first models are proposed by Fisher and Gambel in 1920 and 1940, respectively (Fisher & Tippett, 1928; Gnedenko, 1943; Gumbel, 1958); however, the choice of statistical law depends generally on several factors such as the time of observation, the period of the study, the seasonality and the aridity of study area. Recently, different studies showed that Generalized Extreme Value, Gamma, Normal and Weibull are the most applicable to model the monthly rainfall (Alam et al., 2018; Al-Suhili & Khanbilvardi, 2014). Nevertheless, the best choice is always given according to the p values results obtained by the fit test (Oseni & Ayoola, 2013).

The aim of this study was to study the behavior of climatic variability in Soummam watershed (Algeria), using an adequate the distribution law on a monthly rainfall measurements series of 51 years, to describe the spatio-temporal assessment of the climate over the entire surface of the watershed. This paper gives a statistical description of annual and inter-monthly rainfall followed by a presentation of the distribution models used in our study, according to the results obtained by the fit test. The De Martonne aridity index was also used to express the change of bioclimatic watershed levels during the period of study.

2 | STUDY AREA

Soummam watershed is one of the 17 watersheds in Algeria, its identification number is 15 according to the classification obtained by the National Water Resources Agency (Zouggaghe & Moali, 2009). It is located in the central part of Algerian northern and it extends over 9,125 km². The watershed has an irregular shape characterized by Latitude of 35.75°, 36.75° and a longitude of 3.60°, 5.5° (Figure 1). In the north, the Soummam is delimited mainly by the Djurdjura mountain range, whose highest peak is 2,308 m of altitude (Lalla Khedidja), while its limit in the south is located near to the Hodna Mountains, which have an attitude lower than Djurdjura (the maximum altitude is 1,862 m). From the east to the west, Soummam watershed is open and in this part located the plateau of Setif and Bouira. The climatic conditions of Soummam watershed are not uniform. It was noted that the climate is essentially semihumid in the height part of Soummam valley (Figure 1). However, some studies showed that in Bejaia, the climate is generally humid with slight changes in temperature. On the upland of Setif and Bouira the climate is Mediterranean with humid winters and hot summers. The annual rainfall of Soummam varies enormously between 400 mm on the Setif
upland and 1,000 mm near the Bejaia coast. This variation depends on the geographical parameters, it increases with the altitude under humid winds in the west direction and decreases away from the coastal zones (Lounaci, 2005). The rainfall frequency also decreases from north to south (Turki et al., 2016). Generally, the Soummam watershed climate varies from the north to the south and from the east to the west between humid and Mediterranean with an extension of semihumid conditions.

3 MATERIALS AND METHODS

3.1 Materials

3.1.1 Data description

The study of climatic variability in Soummam watershed (Algerian northeastern) were applied on a monthly time series of rainfall and temperature parameters from January 1967 to December 2018 over 24 climatic stations, which are positioned on the whole surface of the watershed. The data were obtained from National Water Resources Agency (ANRH) and National Centers for Environmental Information (NCEI-NOAA), https://www.ncdc.noaa.gov/. In fact, the data history shows gaps of the record due to difficulty of measurements in some years. The Table 1 showed that the missing rainfall data had an interval range of (1.1%, 14.2%), whereas the temperature values are less frequent, which had an interval range of (0.2%, 4.1%). Hence, the missing data have been processed by using the new approach proposed in our former work to get a continuous and reliable time series (Aieb et al., 2019).
| Code  | Station             | Latitude (°) | Longitude (°) | Altitude (m) | IA* rainfall (mm) | IA* temperature (mm) | Rainfall MD* (%) | Temperature MD* (%) |
|-------|---------------------|--------------|---------------|--------------|-------------------|----------------------|------------------|---------------------|
| 150101| Sour El ghazlen     | 36.15        | 3.69          | 882          | 535               | 14.26                | 4.6              | 0                   |
| 150103| El hachimia        | 36.24        | 3.84          | 730          | 542               | 14.38                | 5.7              | 0.8                 |
| 150201| Souk El khemis     | 36.08        | 3.99          | 1055         | 457               | 14.02                | 2.7              | 1.2                 |
| 150202| Bourdj Oukhriss    | 36.28        | 3.97          | 750          | 460               | 14.12                | 8.3              | 2.5                 |
| 150204| El Asnam           | 36.32        | 4.01          | 435          | 503               | 14.68                | 1.1              | 0.2                 |
| 150206| Bejaia, Airport    | 36.02        | 3.78          | 1            | 663               | 16.8                 | 5.2              | 0.2                 |
| 150301| Mechellah          | 36.37        | 4.27          | 450          | 508               | 15.61                | 8.5              | 1.7                 |
| 150401| Porte de fer       | 36.21        | 4.37          | 500          | 465               | 14.85                | 3.7              | 0.6                 |
| 150503| Ighil Ali          | 36.34        | 4.47          | 700          | 580               | 15.65                | 5.1              | 4.1                 |
| 150607| Ain amat           | 36.19        | 5.31          | 1,029        | 470               | 14.23                | 7.7              | 3.7                 |
| 150608| Bouira Coligny     | 36.2         | 5.27          | 1,010        | 507               | 14.41                | 5.1              | 2.1                 |
| 150610| Farmatou           | 36.23        | 5.39          | 1,030        | 512               | 14.35                | 6.4              | 1.2                 |
| 150611| Mahouane           | 36.26        | 5.35          | 1,150        | 518               | 14.28                | 9.2              | 0.0                 |
| 150613| Zairi              | 36.31        | 5.44          | 1,130        | 520               | 14.62                | 6.4              | 0.0                 |
| 150614| Tixter             | 36.03        | 4.88          | 960          | 420               | 14.12                | 4.0              | 0.4                 |
| 150703| Boubirek           | 36.47        | 4.81          | 500          | 582               | 15.79                | 7.0              | 1.2                 |
| 150706| Ain Abessa         | 36.3         | 5.29          | 1,100        | 549               | 14.73                | 2.5              | 2.2                 |
| 150707| Bir kasdali        | 36.14        | 5.03          | 980          | 450               | 14.25                | 2.0              | 3.7                 |
| 150801| Sidi Mebarak       | 36.1         | 4.91          | 1,011        | 425               | 14.22                | 3.8              | 0.9                 |
| 150802| Zammourah          | 36.27        | 4.86          | 950          | 510               | 15.12                | 9.6              | 1.4                 |

(Continues)
| Code   | Station        | Latitude (°) | Longitude (°) | Altitude (m) | IA* rainfall (mm) | IA* temperature (mm) | Rainfall MD* (%) | Temperature MD* (%) |
|--------|----------------|--------------|---------------|---------------|-------------------|----------------------|------------------|--------------------|
| 150904 | Sidi yahia     | 36.42        | 4.61          | 255           | 580               | 15.67                | 3.8              | 0.8                |
| 151001 | Sidi Aich      | 36.61        | 4.69          | 196           | 592               | 15.89                | 7.9              | 3.2                |
| 151004 | Beni Maouch   | 36.64        | 4.84          | 71            | 598               | 16.32                | 14.2             | 0.4                |
| 151006 | Sddouk        | 36.55        | 4.69          | 370           | 582               | 15.81                | 5.93             | 2.7                |

Abbreviations: IA*, inter annual; MD*, missing data.
3.1.2 | **Classification of Soummam bioclimatic floors**

To determine the bioclimatic floors in Soummam watershed of Algeria, which cover a big part of northeastern of the country, the principal component analysis (PCA) and agglomerative hierarchical clustering (AHC) shown in Figure 2 were applied on two climatic parameters for each individual (weather station) used in our study. The measurements were carried out at stations located all over the area for giving the best information about the bioclimatic floors of the watershed. The PCA graph showed two components PC1 and PC2, which are the interannual rainfall and the interannual temperature of the 24 stations over 51 years (from 1967 to 2018), respectively. It is represented from min to max (from 420 to 663 mm and from 14.02°C to 16.8°C; Table 1), reading from bottom to top and from left to right. PC1 component is the dominant parameter, which can cover 72.00% of the total information compared to PC2. The PCA and AHC results highlight three clusters that have very different characteristics. As shown in AHC plot, all clusters are well separated. The 1st cluster corresponds to the highest area of PCA graph, which represent just Bejaia airport station characterized by a very high annual

![Figure 2](image-url)  
**Figure 2** Principal component analysis followed by agglomerative hierarchical clustering plot showing the Soummam bioclimatic floors, using interannual rainfall and temperature measurements of 24 weather stations during 51 years
precipitation and temperature measurement. This area represents the humid bioclimatic floor of the watershed that overlooks on the Mediterranean Sea (Figure 1). On the other hand, the second cluster showed the largest proportion of the watershed, given by 16 stations, located on the red zone in the center of PCA graph (Figure 2). It is characterized by a very varied rainfall and temperature, between interval values $[-1.5, 2]$ and $[-0.5, 1]$ of PC1 and PC2, respectively. According to the similarity of stations showed in AHC graph, we determined two sets on the same cluster, which are (Beni Maouch, Sour Elghozlen, Sidi Aich, Ighil Ali, Seddouk, Boubirek and Sidi yahia) and (Al Asnam, Mahouane, Farmatou, El hachimia, Zammorah, Mechdellah, Bouira Coligny, and Ain abbess), respectively. This part of the watershed represents the semihumid bioclimatic floor (Figure 1). The third cluster showed all the meteorological stations which have a Mediterranean climate, corresponding to the lowest area in the PCA graph, which have a climate characterized by low values of rainfall and temperature compared to all stations. The AHC graph displayed a great similarity between all the stations of the same cluster.

### 3.2 Methods

According to the literature, the methodology used to study the variability extremes and to analyze the trend of frequency of climatic data series is based mainly on the good knowledge to choose the best distribution model describing the phenomenon. The Kolmogorov-Smirnov ($D_{KS}$) is a nonparametric test used to prove the good fitting of distribution models by evaluated the compatibility between empirical and theoretical cumulative distribution function, which are noted by $F(x)$ and $G(x)$, respectively (Justel et al., 1997; Massey, 1951). The $D_{KS}$ is also given by:

$$ D_{KS} = \max |F(x) - G(x)|. $$ (1)

The test based on the null and the alternative hypotheses, which are:

- $H_0$: the data follow the specified distribution;
- $H_A$: the data do not follow the specified distribution.

The proposed statistical distribution model is also rejected when the $p$ value given by the fit test is lower than the significance level alpha, which is given generally by 0.05 (Simard & L’Ecuyer, 2011). The $p$ value ($P$) of the $D_{KS}$ test is:

$$ P = P[D_{KS} \geq x] = 1 - F_n(x), $$ (2)

where $F_n$ denotes the cumulative distribution function of $D_{KS}$ under the null hypothesis $H_0$.

### 3.2.1 Probability distribution models

In this study, we have selected the generalized extreme value (GEV), Weibull (3) and logistics model according to the fitting tests results obtained in the third section of this study. On the other hands, the advantages of those probability distributions are simple and popular in minima and maxima monthly rainfall analysis (Alam et al., 2018; Uwimana & Joseph, 2018).
GEV distribution

The GEV is a family of continuous probability distributions developed within the extreme value theorem. It was used in many field of study as climatology to study maxima of temperatures, precipitation and river discharges (Hosking & Wallis, 1987). Three parameters of GEV were recommended for meteorology frequencies analysis, which are the shape $\kappa$, the scale $\alpha$ and the location $\xi$, where the parameters estimation is given by method of moment (Bhunya et al., 2007). The probability density function (PDF) of GEV is defined as:

$$f(x) = \alpha^{-1}\exp[-1(1-\kappa)y - \exp(-y)],$$

$$y = -\kappa^{-1}\log\left(1 - \frac{k(x-\xi)}{\alpha}\right), \text{ where } \kappa \neq 0,$$

$$y = \frac{x - \xi}{\alpha}, \text{ where } \kappa = 0.$$  

Weibull distribution

The Weibull probability function is the most widely used in fitting the distribution of rainfall (Udomboso et al., 2010). Its general probability density function is given by Equation (6), where $x \geq \mu$, $\gamma > 0$ and $\alpha > 0$ (Cousineau, 2009).

$$f(x) = \frac{\gamma}{\alpha}\left(\frac{x - \mu}{\alpha}\right)^{(\gamma-1)}\exp\left[-\left(\frac{x - \mu}{\alpha}\right)^{\gamma}\right].$$  

where $\gamma$ is the shape parameter, $\mu$ is the location parameter, and $\alpha$ is the scale parameter. The Equation (6) is called the standard Weibull distribution in the case where $\mu = 0$ and $\alpha = 1$. On the other hand, it is also called the 2-parameter Weibull distribution in the case where $\mu = 0$ (Wilks, 1989).

Logistic distribution

Logistic distribution is a continuous probability distribution. It defined by two parameters for all the real numbers, which are a scale $s$ and a location $\mu$ (Ye et al., 2018). Its PDF plot is always symmetric about the mean. The logistic model can also be used to analyze monthly rainfall occurrence, this technique based on assumption of the average Mean and the variance Var of the number of wet days (Buishand et al., 2004). It can be represented by:

$$\text{Mean}(M_w) = M\Pi,$$

$$\text{Var}(M_w) = \phi M\Pi(1 - \Pi),$$

where $M$ stands for the total number of days in the month of interest, $\phi$ is a dispersion parameter, and $\Pi$ is the wet-day probability. The PDF of logistic distribution used in the third section to fit monthly rainfall of Soummam watershed during 51 years of observation is given by:

$$f(x) = \frac{e^{-\left(\frac{x-\mu}{s}\right)}}{s\left[1 + e^{-\left(\frac{x-\mu}{s}\right)}\right]^2}, \text{ where } -\infty < x < \infty.$$
3.2.2 | Climatic statistical tests

**DE Martonne aridity index**

The De Martonne aridity index ($I_{DM}$) is one of the best known and widely used aridity/humidity indices in applied climatology (Croitoru et al., 2013; de Martonne, 1925). This index is very important in the arid/humid climate classification. Despite the fact that it is one of the oldest indices, it is still used with good results worldwide to identify dry/humid conditions of different regions (Adnan & Haider, 2012; Baltas, 2007; Caloiero, 2004; Shahid, 2008; Zarghami et al., 2011). This index is given by Equation (10):

$$I_{DM} = \frac{P}{T + 10},$$  \hspace{1cm} (10)

where $P$ and $T$ are the annual amount of precipitation and mean annual surface temperature in millimeter and in degree Celsius, respectively.

**Precipitation concentration index**

The Precipitation concentration index considered as a powerful indicator to control temporal precipitation distribution, it was proposed by Oliver (Oliver, 1980), then it was developed by De Luis, González-Hidalgo (De Luis et al., 1997). This index is also very useful for the assessment of annual and seasonal precipitation change, which noted by PCI and SPCI, given by Equations (11) and (12), respectively.

$$PCI = \frac{\sum_{i=1}^{12} P_i^2}{\left(\sum_{i=1}^{12} P_i\right)^2} \times 100,$$

$$SPCI = \frac{\sum_{i=3}^{12} P_i^2}{\left(\sum_{i=3}^{12} P_i\right)^2} \times 25,$$  \hspace{1cm} (11) \hspace{1cm} (12)

where $pi$ represents the monthly precipitation in month $i$, that is calculated for each year of the study.

### 4 | RESULTS

This section highlights the important findings of this study and describes rainfall variability in Soummam watershed in Algeria. Both of graphical and statistical approaches were applied to examine trends of monthly rainfall series during 51 years of the observations.

#### 4.1 | Trend analysis of rainfall and temperature parameters

##### 4.1.1 | Annual rainfall analysis

Figure 3 shows the histogram of difference from the annual rainfall, followed by annual PCI index curve, obtained from average measurements of 24 meteorological stations in Soummam
watershed during 51 years of observations (from 1967 to 2018). The results show a periodic change of annual rainfall variability between dry and wet. The rainfall was driven by an abrupt shift from wet years to successive dry years between 1976 and 2009. Then, an increase in rainfall measurement was given above the average between 2010 and 2018. The results of annual precipitation distribution were given by using the PCI test for further investigation of the climate characteristics (Figure 3). It was observed during all the period of study that PCI values varied between 10 and 18. The annual rainfall distribution showed that the most years have a uniform regime of precipitation. However, in 1973, 1988, 1999, 2002, 2004, 2006 and 2017; the annual rainfall showed a moderate distribution. During this period, the watershed had probably known some month of dryness or wetness, which changed the rainfall ratio from moderate to irregular distribution. These results allowed us to bring about the study of rainfall trends throughout this period.

4.1.2 Inter-monthly rainfall and temperature analysis

Figure 4 shows inter-monthly observations of rainfall and temperature during 51 years (from 1967 until 2018), which are given by a histogram and a curve, respectively. According to graphical results, one can see a great variability given by rainfall observations compared to temperature. Table 2 shows a big distribution variability of inter monthly rainfall, proved by a coefficient of variation and SD, which equal 3.6 and 12.6, respectively. The monthly rainfall average is given by 34 mm for a maximum of 57 mm and a minimum of 5.8 mm. One can conclude a variability range of (+40.4%, −82.9%).

On the other hand, the temperature showed a good distribution according to all the measurements obtained during the 51 years. The table shows that the temperature average equal 18.1°C, it is centralized between a maximum of 24.8°C and a minimum of 12.4°C, this good distribution was given by an interval of variability of (+27%, −31.4%) and a coefficient of variation of 2.5. The PCI index shows that the concentration of seasonal precipitation to be part...
of uniform and moderate distribution, knowing that autumn and summer season shows a uniform regime, given by 8.5 and 9.4, respectively. In contrary, Winter and Springer season are moderate, the PCI index given by 11.1 and 12.2, respectively. Table 3 shows statistical parameters to expresses the irregular distribution of the inter-monthly rainfall that was observed in Soummam watershed during the 51 years. The variations average which are between (5.7 and 57 mm) shows a maximum of precipitation recorded in January, given by 154.5 mm. The table shows that 25% of the rainfall measurements of January during the 51 years varied between (60.1 and 154.5 mm). The stormy periods was recorded on August with a high rainfall reached up to 65.1 mm. The results showed that the season of winter marks the big rainfall variability given by the values range of [25.9, 34.5].

**TABLE 2** Statistical description of monthly rainfall and temperature mean observed during 51 year in Soummam watershed

| Min     | 5.8 | 12.4 |
| Max     | 57.0 | 24.8 |
| Mean    | 34.0 | 18.1 |
| SD      | 12.6 | 6.4 |
| Coef. variation | 3.6 | 2.5 |
| Int. variation | −82.9%, +40.4% | −31.4%, +27% |
| PCI Autumn | 8.5 | – |
| PCI Winter | 11.1 | – |
| PCI Spring | 12.2 | – |
| PCI Summer | 9.4 | – |

**FIGURE 4** Histogram of inter-monthly rainfall followed by inter-monthly temperature curve during 51 years of observations in Soummam watershed
4.2 \hspace{1em} Modeling climate variability

4.2.1 \hspace{1em} Homogeneity analysis

The interannual rainfall measurements observed in each bioclimatic floor of Soummam watershed, which are humid, semihumid, and Mediterranean during 51 years (from 1967 to 2018) are shown as mobile averages curves in Figure 5. The graphs show that there is a change in trend of rainfall frequencies between 2000 and 2018 at humid and semihumid bioclimatic floor. The homogeneity analysis by Pettitt and standard normal homogeneity test (SNHT) are shown graphically in Figure 5a1–a3 and 5b1–b3, respectively. The H₀ hypothesis, which tries to prove that the statistical series are homogeneous, has been rejected in the case of precipitation measurements obtained by humid and semihumid bioclimatic floor. According to the results shown in Table 4, the p value values given by Pettitt test for each of the humid and the semihumid bioclimatic floor are <.0001. Hence, the breaks of frequencies were marked at the time 2001.69 and 2001.61, respectively. On the other hand, the SNHT test showed at 2011.10 and 2011.19 the presence of frequency breaks in humid and semihumid series, these breaks were also shown in graph of panels b1 and b2 with two different means μ1 and μ2 for each of the two previous cases. However, H₀ hypothesis was accepted, it states that rainfall is homogeneous in the case of Mediterranean climate, and we can see that in graph of panels a3 and b3, where it is shown only one mean (μ) for all the series during 51 years. In addition, Table 4 showed that Pettitt and SNHT tests gave a p values equal to .1231 and .1265, respectively.

4.2.2 \hspace{1em} Fitting test

The choice of climatic model depends almost on the conditions of the area and the parameter scale. The purpose of the studies shown in Figure 6 and Table 5 are to evaluate the best
distribution model of rainfall variability for each bioclimatic floor of Soummam watershed during 51 years of monthly observations by using $D_{KS}$ hypothesis. The $p$ values given by the table represents the $H_0$ hypothesis results of $D_{KS}$ test, which was applied on 17 distribution models. The method of moments was also used to estimate the parameters of distribution models selected for this study. Table 5 showed that monthly rainfall of humid bioclimatic floor follows GEV model, proved by a $p$ value of 0.123, which gives the best result compared to all models. This hypothesis is also verified for the log-normal distribution, according to $p$ value result, which equals 0.0474. In the case of the semihumid bioclimatic floor, $H_0$ hypothesis accepts Gamma (2), logistic and weibull (3) model. However, the fitting results favored weibull

**FIGURE 5** Moving average curves of monthly rainfall in humid, semihumid, and Mediterranean bioclimatic floors of Soummam watershed during 51 years of observations, followed by homogeneity graphs (a1, a2, and a3) and (b1, b2, and b3) obtained by Pettitt and SNHT test.
TABLE 4  Standard normal homogeneity test applied on monthly rainfall series of humid, semihumid, and Mediterranean bioclimatic floor during 51 years of measurement

|                | Humid       | Semihumid  | Mediterranean |
|----------------|-------------|------------|---------------|
| Pettitt K      | 38,312.00   | 39,312.00  | 9,552.0000    |
| t              | 2,001.69    | 2,001.61   | 1,983.78      |
| p value (two-tailed) | <.0001 | <.0001 | .1231         |
| alpha          | 0.05        | 0.05       | 0.05          |
| SNHT T₀        | 177.2739    | 193.1352   | 8.6251        |
| t              | 2,011.10833 | 2,011.19167| 1,974.85833   |
| p value (two-tailed) | <.0001 | <.0001 | .1265         |
| alpha          | 0.05        | 0.05       | 0.05          |

(3) with a p value of .1686 and $D_{KS}$ equal 0.0348. On the other hand, the selected model for the monthly rainfall distribution in Mediterranean climatic floor is logistic law, given by a p value equal .7529 and a $D_{KS}$ equal 0.0277. In this case, the fitting tests accept also the normal law according to a p value result, which equal .1501. Figure 6 showed a good fit of GEV, Weibull (3) and Logistic model used cumulative distribution curve and the density histogram. In the case of humid bioclimatic floor, the GEV theoretical model, which is defined by (0.0468, 11.1219 and 37.9512), it is found below the experimental model when the rainfall values are between 40 and 60 (Figure 6b1, b2); this fit is also similar with the rainfall distribution of semihumid climate, which is given by the weibull law (3). Inversely in case of Mediterranean climate, rainfall distribution showed a very good fit, superimposed on theoretical model (Figure 6c1, c2).

4.2.3  Time modeling of rainfall variability

The trends analysis of monthly precipitation in each bioclimatic floor was done temporally by Q-Q plot, P-P plot (which allowed us to cf. between empirical and theoretical distribution of monthly rainfall, as well as the similarity between empirical and theoretical quantiles of the same distribution model, respectively) and standard residual curve as a diagnostic tool, to check the distribution of rainfall extremes and to identify cases where there are breaks of frequency for each bioclimatic floor, compared to theoretical model. The P-P plot graph given by Figure 7 showed that the monthly rainfall distribution of Mediterranean bioclimatic floor has a very good adjustment with logistic model. No large deviations were observed. In this area, the values of the precipitation have a very good similarity with theoretical data given by the model, which is shown graphically by Q-Q plot. This similarity between the values has a very good correlation ($R^2$) of 0.9938. According to standard residual graph, logistic model does not display any break of frequency or trend in the rainfall series during 51 years of observation. On the other hand, Q-Q plot showed a big deviation of the curve from the bisecting line in the case of humid and semihumid precipitations, observed in the part of values that falls within the range of (70, 110) mm and (60, 80) mm, respectively. The differences between theoretical and experimental values are shown in the graph of standard residuals by trends between 400 and 600 months. The results of Figure 7 complete the information of the homogeneous test given in Figure 5 about the break of frequency which has been observed in the last years of analysis. The hypothesis that one can make from this comparison is that humid and semihumid climate exposed to a change in aridity, which can be temporal or spatial.
FIGURE 6  Histogram of monthly rainfall density, followed by cumulative distribution curves given by generalized extreme value, Weibull (3) and Logistic model at humid (a), semihumid (b) and mediterranean (c) bioclimatic floors.
Figure 8 represents spatial modeling results obtained from monthly precipitation distribution analysis in semihumid bioclimatic floor of Soummam watershed during 51 years of observations, given for each meteorological station. The purpose of this study was to specify the area showing break of rainfall frequency using Weibull (3) law, knowing that a big deviation of rainfall trends can cause also a change of aridity in the watershed. The results shown in Figure 8 are obtained by survival weibull regression test, applied on 16 weather stations, which expresses the rainfall distribution compared to theoretical model using the time scale. In addition, the statistical description of this diagnostics summarized in Table 6. The graphs noted by (1–16) show that monthly rainfall at the majority of stations have a good adjustment of weibull (3) distribution during 612 months of observation (from 1967 to 2018), given by a p value varying between .0858 and .5201 (Table 6). The best distribution is that of Ain Abessa station, given by a standard deviation and a p value parameter, which equal 26.615 and .5201, respectively. On the other hand,
Figure 8 shows that Sour Elghozlen, El Hachimia, and Beni Maouch stations have breaks of frequency compared to regression curve. However, Table 6 shows that Beni Maouch have the greatest rainfall variability in this part of the watershed, given by a SD of 46.958. The $D_{KS}$ test shows that the climatic performance of Sour Elghozlen and El Hachimia are similar, according to $p$ value and $D_{KS}$ parameters of (0.0684, 0.0511) and (0.0621, 0.0515), respectively. All the results presented in Figures 7 and 8 support the hypothesis that climate of the Soummam has experienced a shift from wet to dry periods during the last years of measurement.
4.3 | Aridity analysis

Rainfall variability analysis of stations that gave breaks of frequencies compared to GEV and Weibull (3) theoretical model are represented in Figures 7 and 8, respectively. It was given by the histogram of difference from the annual rainfall mean, followed by the annual De Martonne
curve for the study of the aridity. Figure 9a shows a big change of rainfall variability at Bejaia Airport station. The histogram demonstrated that the weather there has passed by two phases, from dry to wet. However, the cumulative rainfall was less than the average between 1967 and 1989; then an increase in precipitation was observed after 1990, where the annual rainfall gave values above the average. The annual De Martonne curve explained this variability by a change in aridity from humid to very humid during this period. On the other hand, Beni Maouch station has been also influenced by the same change of climate, as shown in Figure 9b. De Martonne curve show that the bioclimatic floor of this station changed from semihumid to humid after year 2000, as the weather there had recorded an enormous increase in precipitation between 2000 and 2018. Inversely, El Hachimia and Sour Elghozlen station showed that precipitation is above the average and climate aridity there is semihumid during the first period of the analysis. Then, they were affected by a drought of rainfall between 1991 and 2018, where it was observed a decrease of precipitation measurements and a change of climate aridity from semihumid to Mediterranean (Figure 9c,d). Figure 10 represents a description of Soummam climate aridity during different periods of study, according to interannual rainfall of 13 years, taken the following intervals: (1967, 1980), (1981, 1993), (1994, 2006) and (2007, 2018) referring to the maps a, b, c, and d, respectively. The figure shows a decrease of rainfall observations in the interior and south part of the watershed. It is more observed in map c and d in semihumid and mediterranean bioclimatic floor, which is characterized by a range of rainfall [400, 500 mm] and [500, 600 mm], respectively. However, the northern part of the watershed shows an increase of rainfall, varying between 600 and 800 mm. Figure 10 shows also that Soummam

| Number | Station       | Min | Max  | Mean | SD    | p value | $D_{KS}$ |
|--------|---------------|-----|------|------|-------|---------|----------|
| 01     | Sour Elghozlen| 0   | 204  | 34.496| 44.824|.0684    | 0.0511   |
| 02     | El Hachimia   | 0   | 204  | 33.891| 44.684|.0621    | 0.0515   |
| 03     | El Snam       | 0   | 204  | 29.972| 40.612|.0921    | 0.0452   |
| 04     | Ighil Ali     | 0   | 262.7| 46.705| 33.339|.3088    | 0.0385   |
| 05     | Sidi Aich     | 0   | 294.5| 35.066| 42.893|.0896    | 0.0496   |
| 06     | Bouira Coligny| 0   | 218.4| 32.853| 29.967|.4278    | 0.0349   |
| 07     | Farmatou      | 0   | 223.4| 41.218| 36.862|.2117    | 0.0422   |
| 08     | Mahouane      | 0   | 230.3| 31.991| 31.17 | .3948   | 0.0358   |
| 09     | Zairi         | 0   | 230.4| 33.149| 27.812|.5126    | 0.0331   |
| 10     | Boubirek      | 0   | 271.9| 43.581| 39.453|.1161    | 0.0425   |
| 11     | Ain Abessa    | 0   | 237.1| 47.389| 26.615|.5201    | 0.0325   |
| 12     | Mechedellah   | 0   | 204.1| 41.805| 37.842|.1812    | 0.0432   |
| 13     | Zammourah     | 0   | 204.84| 33.063| 31.226|.0858   | 0.0501   |
| 14     | Sidi Yahia    | 0   | 269  | 39.055| 34.946|.2145    | 0.0416   |
| 15     | Beni Maouch   | 0   | 304.7| 40.15 | 46.958|.0538   | 0.0541   |
| 16     | Sddouk        | 0   | 276.8| 47.395| 42.998|.3811   | 0.0364   |
watershed had a great change in aridity during the 51 years of observations. Between 2007 and 2018, the watershed was characterized by four bioclimatic floors: very humid, humid, semi-humid and Mediterranean. They were given by the following rainfall intervals: [750, 800 mm], [650, 700 mm], [650, 700 mm], and [400, 500 mm], respectively.

FIGURE 9  Histogram of differences from annual rainfall mean, followed by annual $I_{DM}$ curve of Bejaia, Beni Maouch, El hachimia, and Sour Elghozlen Station during 51 years
This study work aims to study rainfall variability of Soummam watershed between 1967 and 2018. The PCA test showed that the study area is characterized by humid climate in the northeastern part, semihumid in the interior part and mediterranean climate in the south part. The annual precipitation follows a moderate and irregular distribution. The adjustment analysis applied by the $D_{KS}$ test showed that the most suitable statistical laws for modeling of monthly rainfall in each bioclimatic floor is GEV, weibull (3) and logistic distributions according to the $p$ values given by .1234, .1686 and .7529, respectively. Moreover, the diagnosis given by P-P plot, Q-Q plot and survival regression curve showed a deviation of monthly rainfall compared to the theoretical data given by the previous models. This expresses a change of aridity proved by De Martonne index between 1994 and 2018, from humid to very humid in Bejaia Airport station, from semihumid to humid in Beni Maouch station. However; El Hachimia and Sour Elghozlen showed a decrease of rainfall, which changed its bioclimatic floor from semihumid to Mediterranean.

FIGURE 10  Spatial variability of inter annual rainfall in Soummam watershed. (a) 1967–1980; (b) 1981–1993; (c) 1994–2006; and (d) 2007–2018

5  CONCLUSION

This study work aims to study rainfall variability of Soummam watershed between 1967 and 2018. The PCA test showed that the study area is characterized by humid climate in the northeastern part, semihumid in the interior part and mediterranean climate in the south part. The annual precipitation follows a moderate and irregular distribution. The adjustment analysis applied by the $D_{KS}$ test showed that the most suitable statistical laws for modeling of monthly rainfall in each bioclimatic floor is GEV, weibull (3) and logistic distributions according to the $p$ values given by .1234, .1686 and .7529, respectively. Moreover, the diagnosis given by P-P plot, Q-Q plot and survival regression curve showed a deviation of monthly rainfall compared to the theoretical data given by the previous models. This expresses a change of aridity proved by De Martonne index between 1994 and 2018, from humid to very humid in Bejaia Airport station, from semihumid to humid in Beni Maouch station. However; El Hachimia and Sour Elghozlen showed a decrease of rainfall, which changed its bioclimatic floor from semihumid to Mediterranean.
6 | PROSPECT

Our future work aims to study the impact of different natural phenomena on the Mediterranean climate change, such as the change that can be caused by the pressures of North Atlantic Ocean to predict the future consequences that can alter the food diversity of those countries.

ACKNOWLEDGMENTS

We wish to thank the staff of biomathematics, biophysics biochemistry, and Scientometry Laboratory in Algeria (BBBS) for their precious help about climatology and for giving us opportunity to use their subset data. We also thank the staff of Mobile and Distributed Systems Laboratory in Italy (MDSLab) for their comments and advices on earlier drafts of this paper.

AUTHOR CONTRIBUTIONS

All authors of this study paper have directly participated in the planning, execution, and analysis of this study. A. A., K. L., and M. S. worked together on the analytical part, which is the collection of data, the fit tests and the choice of distribution models. B. B. and O. M. worked on tests of aridity and the cartography. N. C. and K. M. participated on validation of data and results, as well as the methodology followed in this study. Finally, it is important to point out that all authors wrote the manuscript and approved the final version.

ORCID

Amir Aieb [http://orcid.org/0000-0003-1731-0489]
Khalef Lefsih [https://orcid.org/0000-0002-7661-2121]
Brunella Bonaccorso [https://orcid.org/0000-0002-6973-6425]
Nicola Cicero [https://orcid.org/0000-0002-8055-1915]
Khodir Madani [https://orcid.org/0000-0001-5356-6890]

REFERENCES

Adnan, S., & Haider, S. (2012). Classification and assessment of aridity in Pakistan by using different aridity indices (ENV. Adnan_Pakistan_paper1. pdf). Retrieved from ftp://ftp.wmo.int/Documents/PublicWeb/arep/Weather_Mod_Bali/ENV%20bruntjes.chalon

Aieb, A., Madani, K., Scarpa, M., Bonaccorso, B., & Lefsih, K. (2019). A new approach for processing climate missing databases applied to daily rainfall data in Soummam watershed, Algeria. Heliyon, 5(2), e01247.

Alam, M., Emura, K., Farnham, C., & Yuan, J. (2018). Best-fit probability distributions and return periods for maximum monthly rainfall in Bangladesh. Climate, 6(1), 9.

Alexander, L. V., Zhang, X., Peterson, T. C., Caesar, J., Gleason, B., Klein Tank, A. M. G., ... Vazquez-Aguirre, J. L. (2006). Global observed changes in daily climate extremes of temperature and precipitation. Journal of Geophysical Research: Atmospheres, 111(D5), D05109.

Al-Suhili, R. H., & Khanbilvardi, R. (2014). Frequency analysis of the monthly rainfall data at Sulaimania region, Iraq. American Journal of Engineering Research, 3(5), 212–222.

Baltas, E. (2007). Spatial distribution of climatic indices in northern Greece. Meteorological Applications, 14(1), 69–78.

Bhunya, P. K., Jain, S. K., Ojha, C. S., & Agarwal, A. (2007). Simple parameter estimation technique for three-parameter generalized extreme value distribution. Journal of Hydrologic Engineering, 12(6), 682–689.

Boudrissa, N., Cheraitia, H., & Halimi, L. (2017). Modelling maximum daily yearly rainfall in northern Algeria using generalized extreme value distributions from 1936 to 2009. Meteorological Applications, 24(1), 114–119.

Buishand, T. A., Shabalova, M. V., & Brandsma, T. (2004). On the choice of the temporal aggregation level for statistical downscaling of precipitation. Journal of Climate, 17(9), 1816–1827.
Caloiiero, T. (2004). Climatic trends: An investigation for a Calabrian basin (southern Italy). *The Basis of Civilization—water Science?* 286, 255.

Cousineau, D. (2009). Fitting the three-parameter Weibull distribution: Review and evaluation of existing and new methods. *IEEE Transactions on Dielectrics and Electrical Insulation, 16*(1), 281–288.

Croitoru, A. -E., Piticar, A., Imbroane, A. M., & Burada, D. C. (2013). Spatiotemporal distribution of aridity indices based on temperature and precipitation in the extra-Carpathian regions of Romania. *Theoretical and Applied Climatology, 112*(3–4), 597–607.

Dore, M. H. (2005). Climate change and changes in global precipitation patterns: What do we know? *Environment International, 31*(8), 1167–1181.

Faramarzi, M., Abbaspour, K. C., Ashraf Vaghefi, S., Farzaneh, M. R., Zehnder, A. J. B., Srinivasan, R., & Yang, H. (2013). Modeling impacts of climate change on freshwater availability in Africa. *Journal of Hydrology, 480*, 85–101.

Fisher, R. A., & Tippett, L. H. C. (1928). Limiting forms of the frequency distribution of the largest or smallest member of a sample. In *Mathematical Proceedings of the Cambridge Philosophical Society*. Cambridge University Press.

Giec, T.F.S., et al., *Changements climatiques 2013: Les éléments scientifiques*, 2013, Cambridge University Press Royaume-Uni et New York, État-Unis.

Gnedenko, B. (1943). *Théorie des éléments* et de l’unicité et de la stabilité de la distribution limite de la série aléatoire. "Annales des Sciences Mathématiques*, 1, 55–105.

Gumbel, E. J. (1958). *Statistics of Extremes* (p. 201). New York, NY: Columbia University Press.

Hailemariam, K. (1999). Impact of climate change on the water resources of Awash River Basin, Ethiopia. *Climate Research, 12*(2–3), 91–96.

Hosking, J. R., & Wallis, J. R. (1987). Parameter and quantile estimation for the generalized Pareto distribution. *Technometrics, 29*(3), 339–349.

Justel, A., Peña, D., & Zamar, R. (1997). A multivariate Kolmogorov-Smirnov test of goodness of fit. *Statistics & Probability Letters, 35*(3), 251–259.

Lounaci, A. (2005). *Recherche sur la faunistique, l’écologie et la biogéographie des macroinvertébrés des cours d’eau de Kabylie (Tizi-Ouzou, Algérie).* Thèse de doctorat d’état en biologie. Université Mouloud Mammeri de Tizi.

De Luis, M., et al. (1997). Distribución espacial de la concentración y agresividad de la lluvia en el territorio de la comunidad valenciana. *Cuaternario y Geomorfología, 11*(3), 33–44.

de Martonne, E., *Traité de géographie physique: climat, hydrographie. Notions générales*. 1925: Armand Colin.

Massey, F. J., Jr (1951). The Kolmogorov-Smirnov test for goodness of fit. *Journal of the American Statistical Association, 46*(253), 68–78.

New, M., Todd, M., Hulme, M., & Jones, P. (2001). Precipitation measurements and trends in the twentieth century. *International Journal of Climatology: A Journal of the Royal Meteorological Society, 21*(15), 1889–1922.

Nouaceur, Z., & Murărescu, O. (2016). Rainfall variability and trend analysis of annual rainfall in North Africa. *International Journal of Atmospheric Sciences, 2016*, 1–12.

Oliver, J. E. (1980). Monthly precipitation distribution: A comparative index. *The Professional Geographer, 32*(3), 300–309.

Oseni, B. A., & Ayoola, F. J. (2013). Fitting the Statistical Distribution for Daily Rainfall in Ibadan, Based On Chi-Square and Kolmogorov-Smirnov Goodness-Of-Fit Tests. *West African Journal of Industrial and Academic Research, 7*(1), 93–100.

Planton, S., Déqué, M., Douville, H., & Spagnoli, B. (2005). Impact du réchauffement climatique sur le cycle hydrologique. *Comptes Rendus Geoscience, 337*(1-2), 193–202.

Shahid, S. (2008). Spatial and temporal characteristics of droughts in the western part of Bangladesh. *Hydrological Processes: An International Journal, 22*(13), 2235–2247.

Simard, R., & L’Ecuyer, P. (2011). Computing the two-sided Kolmogorov-Smirnov distribution. *Journal of Statistical Software, 39*(11), 1–18.

Stern, R., & Cooper, P. (2011). Assessing climate risk and climate change using rainfall data—A case study from Zambia. *Experimental Agriculture, 47*(2), 241–266.
Turki, I., Laignel, B., Massei, N., Nouaceur, Z., Benhamiche, N., & Madani, K. (2016). Hydrological variability of the Soummam watershed (Northeastern Algeria) and the possible links to climate fluctuations. *Arabian Journal of Geosciences, 9*(6), 477.

Udomboso, C., Chubwu, A., & Nja, M. (2010). The use of gamma and Weibull distributions in modeling rainfall data in Nigeria (a comparative analysis). *Journal of Science Research, 9*(6), 69–73.

Uwimana, O., & Joseph, M. K. (2018). Modelling extreme maximum rainfall using generalized extreme value distribution: Case study Kigali city *Journal of Science and Research (IJSR), 7*(6), 121–126.

Wilks, D. S. (1989). Rainfall intensity, the Weibull distribution, and estimation of daily surface runoff. *Journal of Applied Meteorology, 28*(1), 52–58.

WMO. (2016). *The Global Climate in 2011–2015*.

Ye, L., Hanson, L. S., Ding, P., Wang, D., & Vogel, R. M. (2018). The probability distribution of daily precipitation at the point and catchment scales in the United States. *Hydrology and Earth System Sciences, 22*(12), 6519–6531.

Zarghami, M., Abdi, A., Babaeian, I., Hassanzadeh, Y., & Kanani, R. (2011). Impacts of climate change on runoffs in East Azerbaijan, Iran. *Global and Planetary Change, 78*(3–4), 137–146.

Zouggaghe, F., & Moali, A., Variabilité structurelle des peuplements de macro-invertébrés benthiques dans le bassin versant de la Soummam (Algérie, Afrique du Nord). Revue d’écologie, 2009.

---

**How to cite this article:** Aieb A, Lefsih K, Scarpa M, et al. Statistical modeling of monthly rainfall variability in Soummam watershed of Algeria, between 1967 and 2018. *Natural Resource Modeling*. 2020;33:e12288. [https://doi.org/10.1111/nrm.12288](https://doi.org/10.1111/nrm.12288)