atmosphere

MDPI

Article

SPI-3 Analysis of Medjerda River Basin and Gamma Model Limits in Semi-Arid and Arid Contexts

Zoubeida Kebaili Bargaoui [1,*] and Sabrine Jemai [2]

[1] Department of Civil Engineering, Water and Environmental Research Laboratory, National School of Engineers of Tunis, University of Tunis, El Manar BP 37, Tunis 1002, Tunisia
[2] Water and Environmental Research Laboratory, National School of Engineers of Tunis, University of Tunis, El Manar BP 37, Tunis 1002, Tunisia
* Correspondence: zoubeida.bargaoui@enit.utm.tn

Abstract: The Standardized Precipitation Index *SPI-3*, associated with three months of rainfall accumulation, is a drought index for detecting immediate drought impacts. The two-parameter gamma distribution, recommended by the World Meteorological Organization as the underlying distribution for estimating *SPI*, has shown limits in semi-arid and arid conditions with respect to the normality test for the resulting *SPI* series. Our purpose was to evaluate its relevance for the Medjerda River Basin (Tunisia), a transboundary basin where the climate classes are temperate, dry, and hot summer, as well as arid hot desert and arid hot steppe. When analyzing the time series of 144 stations from 1950 to 2018, we found that the normality Shapiro–Wilk test was rejected in 17% of the cases, which agreed with the literature review results. The transition season (August, September, and October) had the highest rejection percentage. Three factors were identified to explain the deviation from normality. We first identified the rate of occurrence of completely dry (zero rain) three-month periods. The higher the rate of occurrence was, the higher that the probability was of its rejecting the normality test. High sample skewness was the second influencing factor. Finally, a series where the Grubbs' test of identifying outliers was rejected was more likely to show the *SPI-3* series deviating from normality.

Keywords: monthly rainfall; drought; *SPI-3*; Shapiro–Wilk test; south Mediterranean region; gamma distribution

check for updates

Citation: Kebaili Bargaoui, Z.; Jemai, S. *SPI-3* Analysis of Medjerda River Basin and Gamma Model Limits in Semi-Arid and Arid Contexts. *Atmosphere* **2022**, *13*, 2021. https://doi.org/10.3390/atmos13122021

Academic Editors: Hao Guo, Guoxiong Zheng and Liangliang Jiang

Received: 2 September 2022
Accepted: 26 September 2022
Published: 1 December 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Rainfall observations are the basis for many scientific studies in various areas, such as climatology, hydrology, geography, and ecology. Rainfall is also a driver for economy in relation to rainfed crops, especially cereals. The investigation of drought occurrence is based on drought indices computed using a monthly rainfall series as well as an evapotranspiration series ([1–3]). A drought warning is crucial for cereal crops and food security [4]. The Standardized Precipitation Index (SPI) [5] was recommended by the World Meteorological Organization (WMO), among other indices [6]. Another source of data to elaborate drought indices is potential evapotranspiration, giving rise to the Standardized Precipitation Evapotranspiration Index ($SPEI$) [7]. High correlations were found between SPI and $SPEI$ in temperate and continental regions of Iran [2], while low correlations were found in arid climate regions. Considering this question, we focused on SPI with a three-month accumulation for quantifying the immediate drought impacts such as agricultural and soil moisture drought. In Tunisia, cereals are mostly rainfed, and the country has a high deficit in cereal trade balance. Thus, agricultural drought is of prime interest in the study region. Recently, a national insurance fund was implemented in 2019 to help farmers facing drought impacts on cereal crops, using SPI information as support for its decisions. The Standardized Precipitation Index *SPI-3* represents the number of standard deviations by which the normalized, cumulative precipitation deviates from the climatological median

for a three-month accumulation period [5]. In the parametric framework, *SPI* is obtained by first adopting an underlying parametric distribution to fit rainfall totals and then by transforming the estimated non-exceeding probability to make them normally distributed. Such a normalization is important for achieving spatial intercomparisons as well as for interpreting the drought index [1] in terms of drought severity. Several two-parameter and three-parameter candidate distributions were evaluated to assess whether they correctly represented the skewed rainfall totals. After considering various candidate distributions and ranking them based on several durations of accumulation periods (using $0.5^\circ \times 0.5^\circ$ gridded historical data based on ERA-40 reanalysis), researchers found it more accurate to use a simple distribution to estimate *SPI* series in order to avoid overfitting and interpretability issues [1]. The two-parameter gamma distribution was identified as the most likely to give acceptable results across all accumulation periods and regions of Europe [1]. Its success was explained by the flexibility of its shape parameter. A three-month resolution [8] found that, for global land grid points, two-parameter gamma distribution performed at a level of 94% accuracy in the category "best performance".

An issue pointed out by [1] was the method adopted for testing the adequacy of the candidate distribution. They argued that the lack of rejection of the Kolmogorov–Smirnov test does not ensure the normality of the transformed rainfall. They recommended testing *SPI* normality using the Shapiro–Wilk test [9]. Since normality is a necessary condition for interpretating *SPI* in terms of drought categories, as adopted by the WMO's *SPI* user guide [6], a default in normality will result in over- or underestimating drought severity, especially in the lower tail of the distribution. Another recommendation was to implement a limitation to interpret *SPI* estimates to intervals (−3, +3) since the likelihood of exceeding these limits is very low, especially considering that the commonly available lengths of samples are not very high [1].

Based on a comparison with crop yields, [10] found that *SPI* was overestimated for stations with a low rainfall average. Such an overestimation in the left tail of the rainfall distribution resulted in underestimating drought severity. They mentioned the failure in fulfilling the normality assumption as a reason for this overestimation. The Shapiro–Wilk rejection was 3–13% in the case of the gridded, historical ERA-40 reanalysis data over Europe [1]. In addition, the normality test of *SPI* derived from the gamma model failed for approximately 10% of 264 studied series in Brazil [11]. Studying rainfall data from the United States, Ref. [12] suggested that the normality test results were linked to local climate conditions and that rejection was more likely for stations characterized by a high probability of no-rain cases.

According to [1,12], one challenge in selecting a candidate distribution is to ensure that it captures the likelihood of zero precipitation. Thus, our aim was to assess the relevance of computing the Standardized Precipitation Index *SPI-3* using a two-parameter gamma distribution in the case of the Medjerda River Basin that, according to Koppen's climate classification [13], fits into the following climates: *Csa* (temperate, dry, and hot summer), *Bwh* (arid hot desert), and *Bsh* (arid hot steppe). The Medjerda River Basin is the most important perennial surface water provider in Tunisia. It flows to the South Mediterranean in Kallat Landlous, with a basin area of nearly 23,000 km^2. It is a transboundary river with its source in Algeria. In addition, it is one of the most important transboundary rivers in North Africa, feeding the West Mediterranean. To date, there has not been a study with *SPI-3* estimation using the potential rainfall network that is displayed in Tunisia. Conversely, the performance of the Standardized Precipitation Index using gamma distribution has not been estimated so far for this basin. The most recent research for the North Africa region used gamma distributions to estimate *SPI* time series in 16 stations of the Wadi Mina basin (4900 km^2) in northern Algeria [14], without elaborating on the conclusions about normality assumption. Additionally, Ref. [15] studied 194 rainfall series, including the entire domain of northern Algeria using a copula approach for estimating the *SPI-3* drought's severity, duration, extension area, and return levels (1970–2006). The advantage of the copula approach is to avoid estimating the univariate probability distribution functions

(pdfs). However, they did not mention any test for normality assumption. In [16], the authors studied the period 1960–2010 using 65 rainfall stations in the Cheliff-Zahrez basin (56,227 km^2) of northern Algeria with semi-arid to arid climate conditions. They adopted a non-parametric SPI estimation without mentioning the performance with respect to the normality of resulting $SPIs$. Thus, our study was based on a higher rainfall network density and covered the three types of climates of Tunisia in order to draw conclusions about the effectiveness of the normality test for SPI when using the gamma distribution as the underlying statistical model.

2. Materials and Methods

The data are from rainfall yearbooks published by the Tunisian National Hydrological Service, which is part of the Ministry of Agriculture, Hydraulics, and Fish Resources. There, one can find information about the size and composition of the rainfall network during a given hydrological year, which is assumed to extend from 1 September of one year through 31 August of the next year. The station names and identifiers are provided, as well as the elevation above sea level and the geographical localization for every station. Tunisia is subdivided into 7 main watersheds, with the Medjerda River Basin assigned to the fifth. In rainfall and hydrometric reports, the Medjerda River Basin is known as Basin 5 ($BV5$). Consequently, the stations of the Medjerda River Basin have identifiers beginning with '5 . . . '. It is worth noting that the labelling of the station names may present some small changes from one yearbook to another, due to different methods of translating the name from Arabic to French. This is why the most important reference for any station is its identifier.

The yearbooks report daily rainfall measurements and monthly and annual rainfall totals. They indicate daily and monthly missing data (gaps) using the symbol '-'. Gaps may be due to a temporary lack of observation, deterioration of the equipment, or loss of register. It may happen that a station is abandoned, mostly due to the lack of an observer (from death, a change of residence, or other personal reasons) or due to a decrease in database budgets, resulting in reduction of the network size. A closure is officially reported in a specific (not public) database containing the station's name, location, and elevation above sea level. When this is not reported, one cannot be certain of a station's closure. In the Medjerda River Basin, the first station was implemented in 1888. In this study, we focused on the analysis of the recent period 1950–2018 when the network was comprised of 144 stations. One single station (code number 53766) out of 144 had no metadata. The spatial locations of the remaining 143 stations are shown in Figure 1. Two stations are located out of the basin, showing that their coordinates are erroneous in the yearbook (Figure 1). However, this did not affect our results. The stations are represented within their climate classification, using Koppen's climate classification system (Figure 1). A total of 49% are *Csa* (temperate, dry, and hot summer), while 31% belong to *Bwh* (arid hot desert) and 20% to *Bsh* (arid hot steppe). In the Medjerda River Basin, the driest month is July, while the rainy season is from November to January. The transition season is August through October, and spring expends from March to May.

We first evaluated the network's spatiotemporal extension every month. The gaps in the monthly rainfall data were not reconstructed. Thus, the number of stations with complete monthly data changes from year to year. However, it cannot exceed 144, which is the size of the studied network. We then identified the 3-month total for every station. Any 3-month period containing a missing month (a gap) was assumed missing and was removed from the time series used to fit the underlying statistical distribution, since the gap meant that the resulting *SPI-3* predictions were missing in those cases. We further selected those stations with sample sizes of 3-month totals greater than 30. We then evaluated the frequency f_0 of occurrence of no rain during every 3-month period. The frequency f_0 was estimated as the ratio of the number of zeros in the series by the series size, as recommended in [6]. Additionally, data control was performed for the 3-month total using outlier detection. To do so, we used Grubbs' test [17] to assess whether the maximum was

an outlier. We use a two-sided test and the critical value G_{max} of the t-distribution at a significant level of 5% to evaluate the test statistic G.

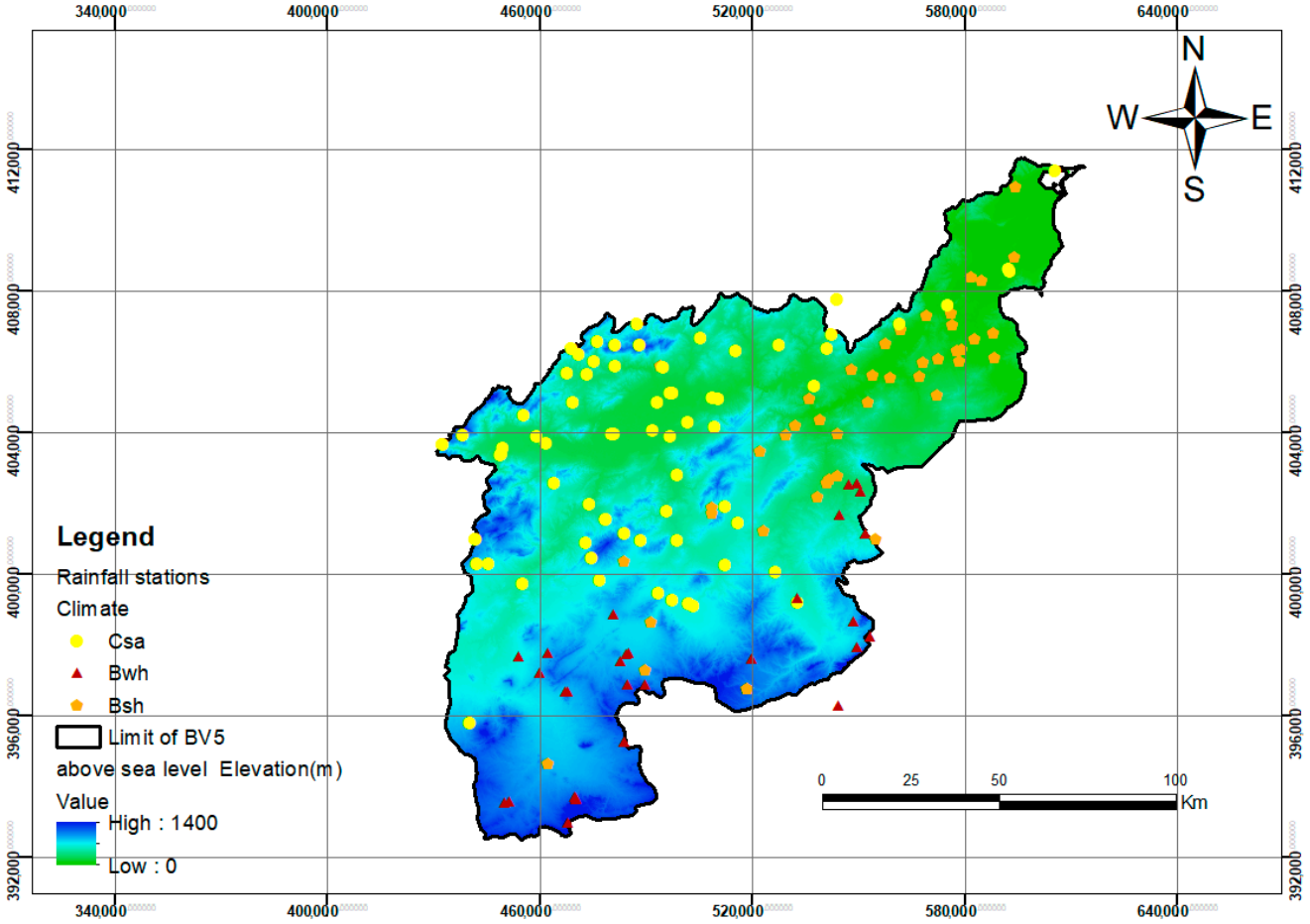


Figure 1. Spatial distribution of rainfall stations in the Medjerda River Basin from 1950 to 2018 and the climate classification for stations at elevations above sea level.

The parametric estimation of *SPI-3* required fitting a statistical distribution to the 3-month totals. As stated before, we adopted two-parameter gamma distribution, with a α scale and λ-shaped parameters [18,19]. The method of maximum likelihood, which uses the logarithmic transformation of data and is known to be more efficient, was not appropriate because of the existence of zeros in the data. The method of moments was used to estimate the model's parameters. Furthermore, based on α and λ estimators, the standardized gamma variable K_y associated with any observed 3-month total y was appraised using the following equation:

$$K_y = (y - (\lambda/\alpha))/(\lambda\gamma/\alpha^2) \quad (1)$$

Additionally, the Wilson–Hilferty transformation (see [18], Eq. 4.27, p. 33) was adopted to normalize quantiles. The standardized gamma quantile was related to the qth quantile of the standard normal distribution Φ_q by Equation (2), which states that:

$$K_y \approx (2/C_s)\left((C_s/6 * (\Phi_q - (C_s/6)) + 1)^3 - 1\right) \text{ for } C_s > 0 \quad (2)$$

where

C_s was the population asymmetry coefficient.

$$C_s = 2(\lambda)^{-0.5} \quad (3)$$

Estimating α and λ and inverting (2) helped to calculate *SPI-3*, taking into account the presence of zeros in the data and the fact that the gamma was not defined for $y = 0$ in [20].

$$\begin{array}{ll} SPI\text{-}3(y) = \Phi_q & \text{for } y \neq 0 \\ SPI\text{-}3(y) = \Phi^{-1}(f_0) & \text{for } y = 0 \end{array} \quad (4)$$

where y represents the 3-month rainfall total; $\Phi^{-1}()$ is the inverse of the standard normal cumulative distribution function ($\Phi q = \Phi^{-1}(q)$) [21].

Furthermore, the adequacy of the adjusted pdf model was tested using a Kolmogorov–Smirnov test at the 5% significance level. In addition, for those models that were not rejected for adequacy, a Shapiro–Wilk test [9] was implemented to check whether the *SPI-3* estimates were normally distributed. According to [22], the Shapiro–Wilk test achieves the best performance when compared with other normality tests.

Finally, the normality results were examined in the light of the following factors: the sample size, station location, Grubbs' test results, frequency f_0 of the occurrence of no rain, and population asymmetry coefficient C_s. Figure 2 shows the inputs of the study (monthly rainfall series of 144 stations), its outputs (*SPI-3* series and the percentage of the series with an accepted normality test), and the steps to bridging them.

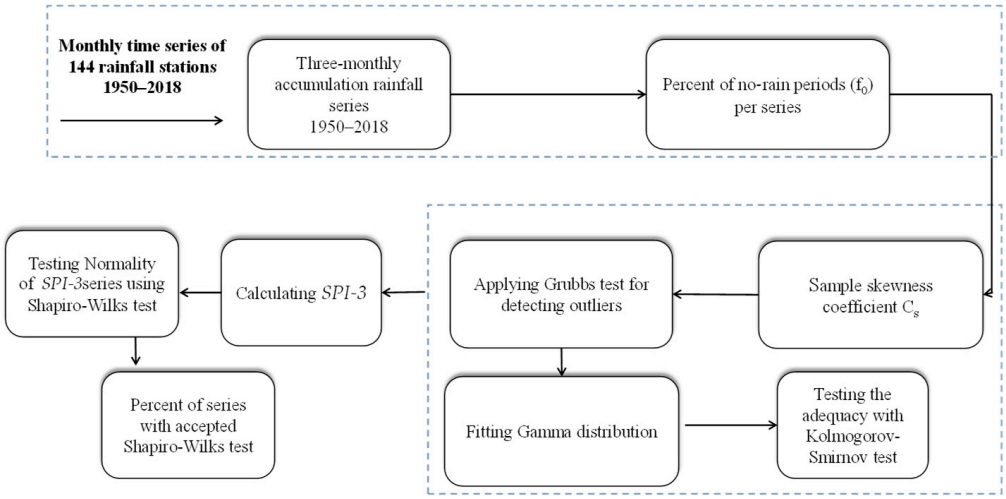


Figure 2. Flowchart showing inputs, methods, and outputs.

3. Results

3.1. Network Time Evolution, Percent of Gaps, and Frequency of Occurrence of No-Rain Periods

Figure 3 shows the spatial distribution of the rainfall network for the hydrological year 1950–1951, which was the starting year of our study period. It is comprised of 74 stations. An augmentation of the network size was observed between 1970 and 1990, with respectively 90 and 101 stations in operation during the hydrological years 1969–1970 and 1990–1991. However, the number of stations was again reduced to 74 in the hydrological year 2018–2019, which was the most recent year considered in this study (Figure 3). A study of network optimization was achieved in 2007 by the National Hydrological Service that resulted in an important network size decrease. The stations' locations are shown with respect to their elevations above sea level in Figures 1 and 3. Both reflect the acceptable representativity of the orography by the rainfall network.

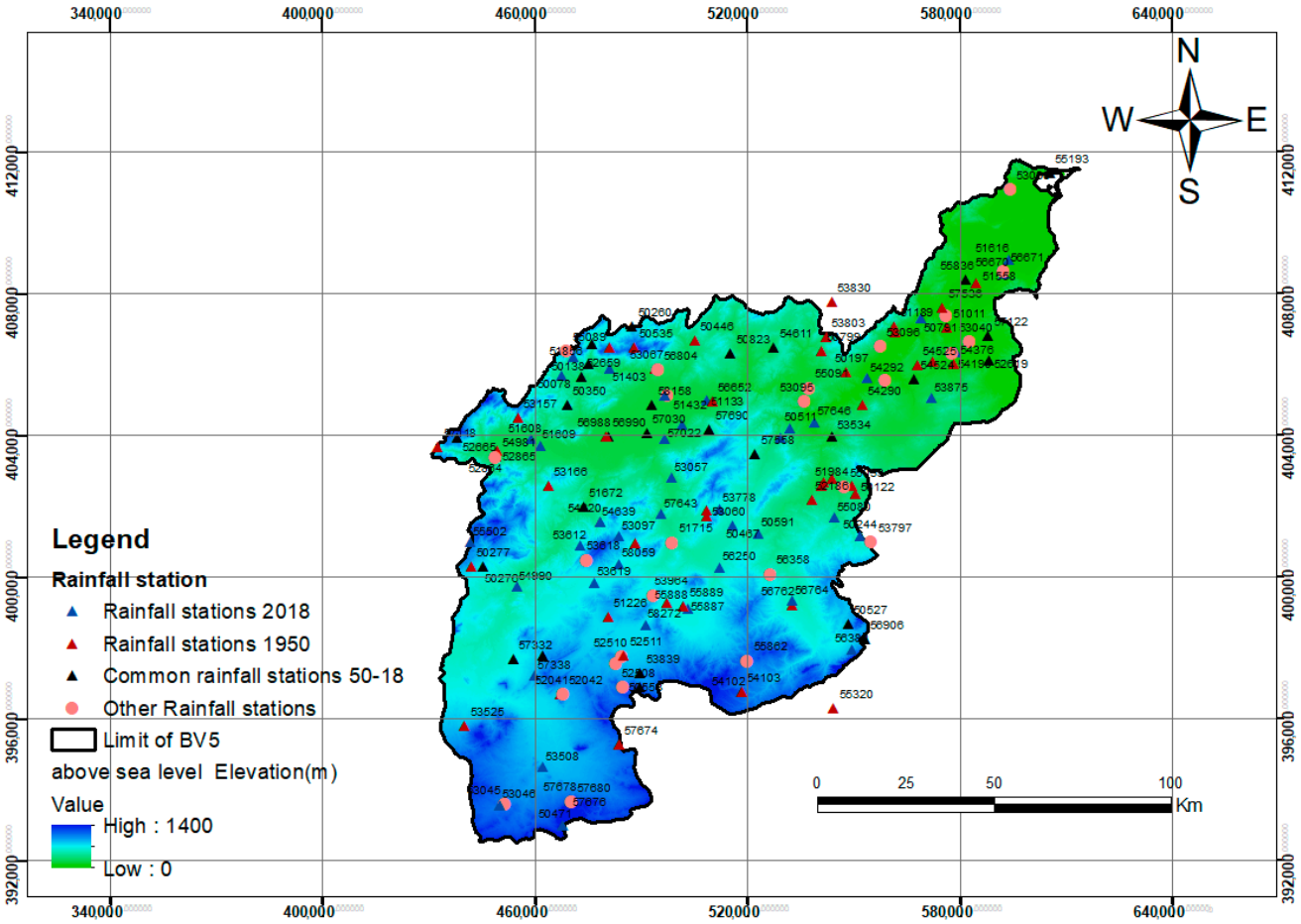


Figure 3. Rainfall stations' spatial cover in 1950–1951 and 2018–2019 and the stations' elevations above sea level.

Figure 4 shows the box plots of the monthly trimmed mean, considering the threshold of 3%. The seasonal variability is highlighted, with July as the driest month and January as the wettest month. In total, for the period of 1950–2018, 98 stations totaled more than 30 years of observations for every calendar month, representing 59,032 station-months at monthly level. The maximum monthly sample size was 69 years corresponding to a single station (51,672) without gaps. In contrast, 45 stations out of 144 had fewer than 31 years of observations per month during the period 1950–2018. Nineteen of them are stations that opened before 1931. For stations with more than 30 years of observations for every month, the ratio of the number of gaps to the number of station-months was 14%. At station scale, the median rate of gaps was 9%. Per month, the mean rate of gaps was 14–15%, except for July and August, where it was higher (17–18%).

The percentage of zero rain (f_0) during 3-month periods varies, as shown in Figure 5 as a box plot representation. A singular value exceeding 0.35 was found for one single station from July through September. For a few stations, f_0 exceeded 0.25 for June–August, which corresponded to the summer season. The lowest f_0 frequencies displaying the lowest variability corresponded to the 3-month period of February–April (Figure 5), which encompasses the rainy season.

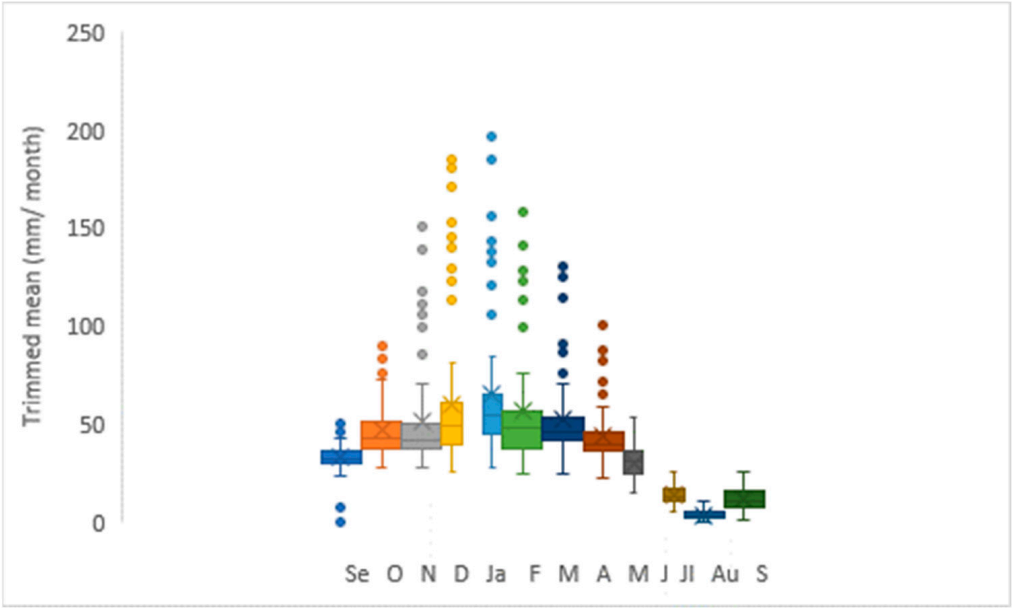


Figure 4. Box plots showing the seasonal and interstation variability of the monthly trimmed mean.

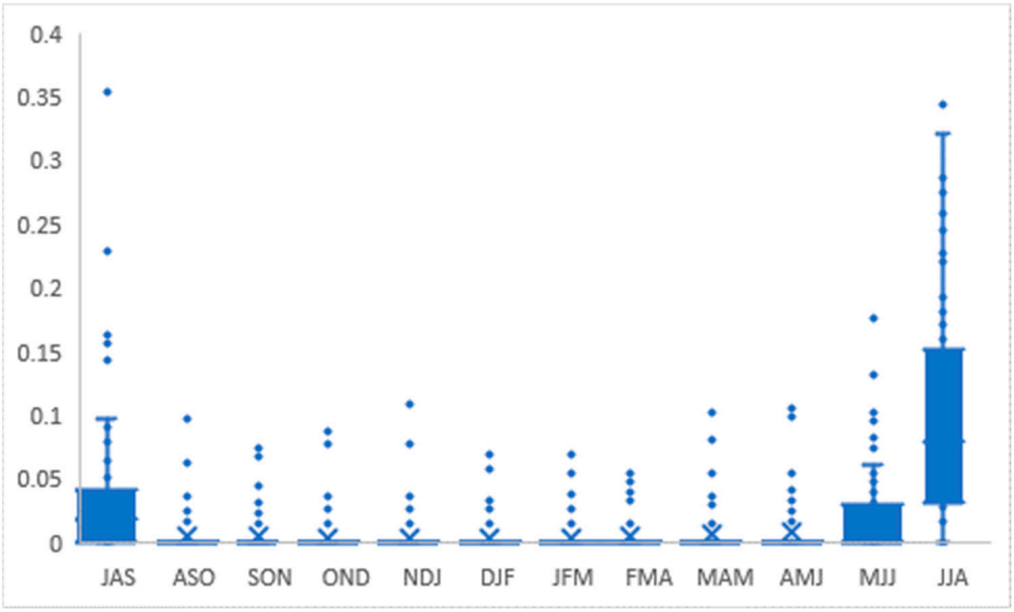


Figure 5. Box plots of the frequency of zero rainfall f_0 for 3-month periods.

3.2. Fitted Gamma Distributions

The fitting of the two-parameter gamma distribution for every 3-month period and the use of the Kolmogorov–Smirnov test at a 5% confidence level resulted in its acceptation for 1205 station-months. For 206 out of 1205 3-month periods (17%), the normality of *SPI-3* was rejected using the Shapiro–Wilk test. Figure 6 shows the sample sizes for series not rejected by the Shapiro–Wilk test. Figure 6 shows that the number of samples having more than 50 years of observations for 3-month periods was not trivial and was favorable for inferring a statistical analysis.

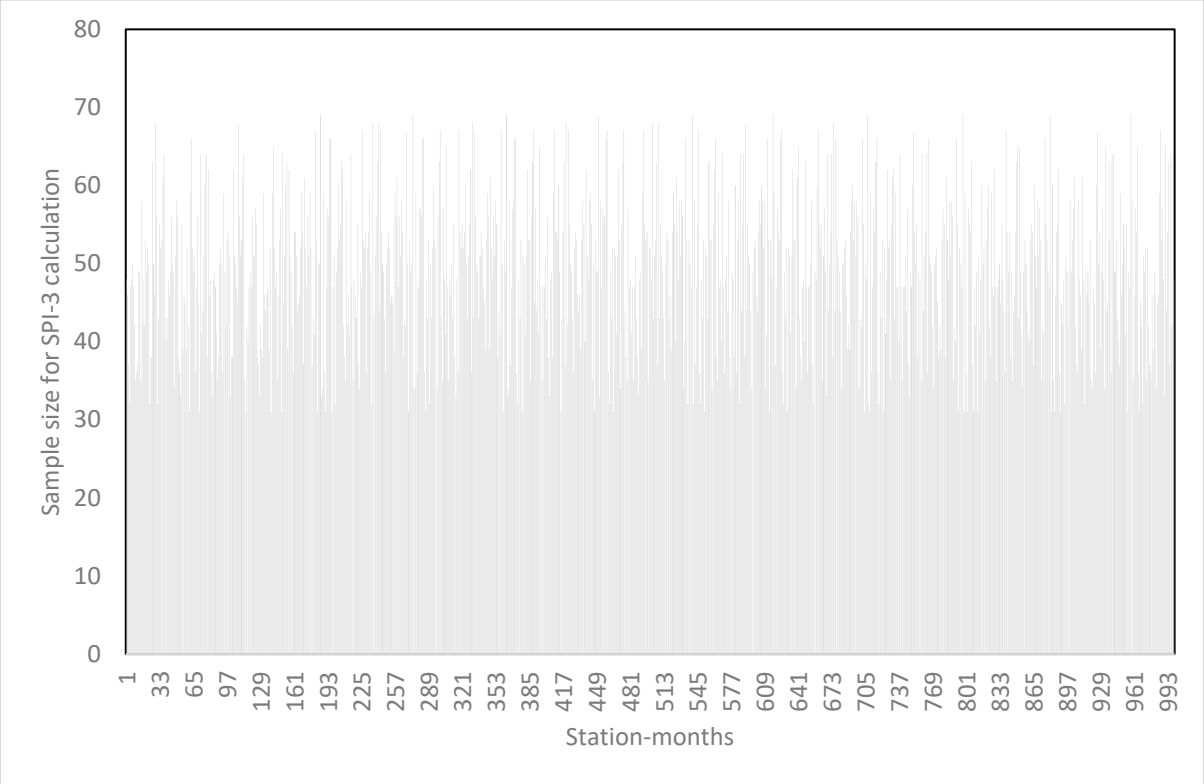
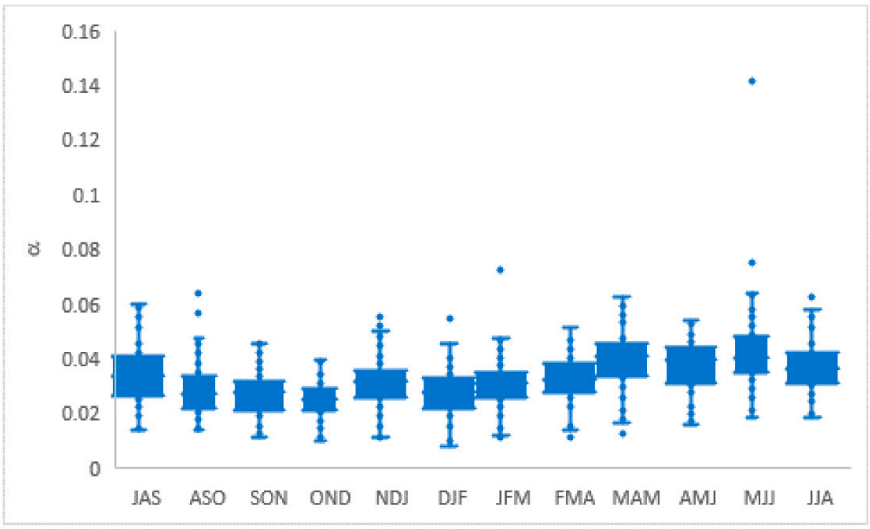


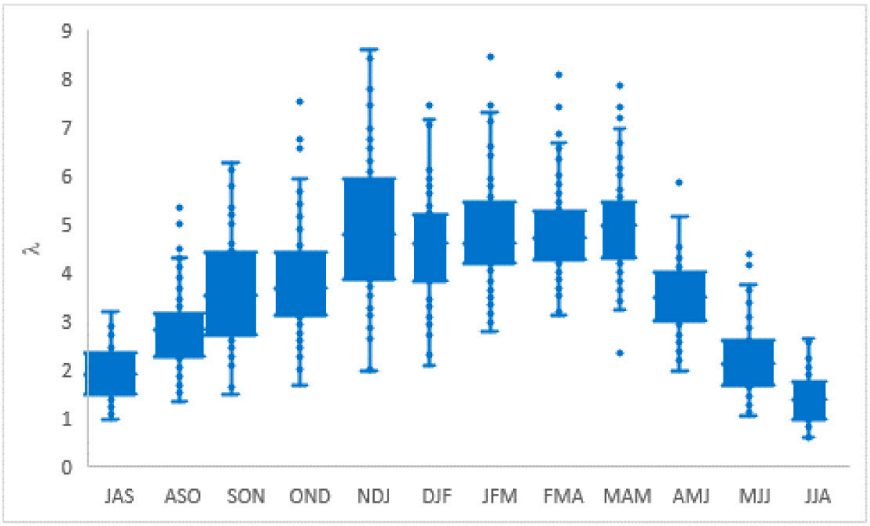
Figure 6. Sample sizes for 3-month totals for stations with an accepted Shapiro–Wilk test.

Figure 7a,b shows the temporal variability of the gamma distribution parameters for those cases where the normality of *SPI-3* was not rejected. There is a clear seasonal effect in both parameters. Figure 7a,b shows that the seasonal curve was more accentuated for the shape parameter than for the scale parameter. One single station (code number 51688) seemed singular for the scale parameter (0.14) corresponding to a 3-month period (May–July). This station had a sample size of 31 years with $f_0 = 0$. Its skewness coefficient was not very high ($C_s = 0.74$), but its Grubbs' test was rejected, which might explain this singularity.



(**a**)

Figure 7. *Cont.*



(**b**)

Figure 7. (**a**) Box plots and time variation of the α scale parameter for those stations with an accepted Shapiro–Wilk test. (**b**) Box plots and time variation of the λ shape parameter for those stations with an accepted Shapiro–Wilk test.

3.3. Identification of Factors Responsible for Decrease in Fitting Performance

A total of 84 stations have at least one 3-month period with rejected normality. Their geographical locations are shown in Figure 8. The 84 stations are equally distributed over the study area, suggesting no link to their location conditions.

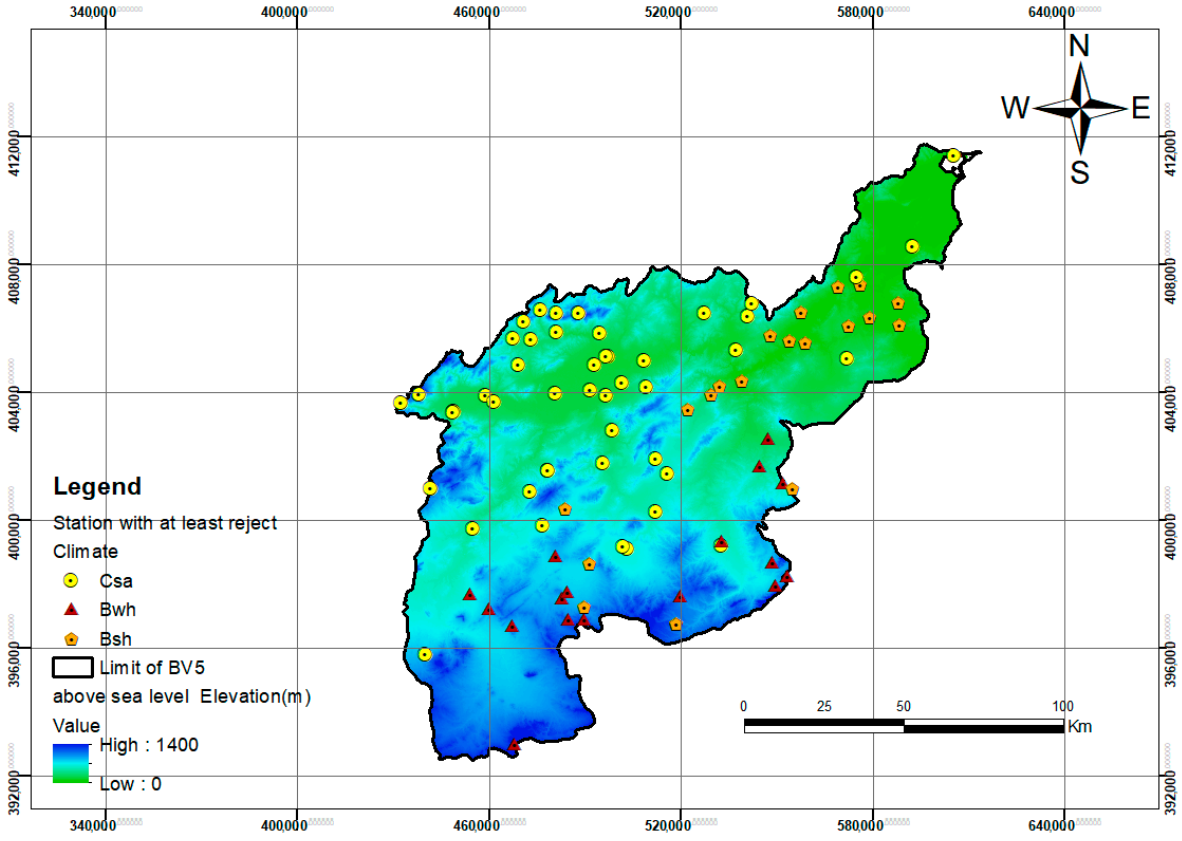


Figure 8. Spatial distribution of stations with a rejection of a Shapiro–Wilk test for at least one 3-month period and their climate categories.

Table 1 shows the number of station-months with an accepted Shapiro–Wilk test for every 3-month period. The corresponding quantity of data (48,835) was computed to be the sum of all sample sizes. Considering Grubbs' test results, the decision of rejection was 59.9% of the station-months with an accepted Shapiro–Wilk test for *SPI-3* (Table 1). Nevertheless, we did not censor the data since a lot of other tests exist for detecting outliers. In addition, outlier presence may be due to heavy-tailed distribution of the 3-month totals. Thus, more research is needed in this area.

Table 1. Characteristics of accepted periods with respect to the *SPI-3*'s Shapiro–Wilk test.

3-Month Period	Number of Station-Months with Accepted Shapiro–Wilk Test	Percent of the Total	Number of Data with Accepted Shapiro–Wilk Test	Number of Station-Months with Accepted Shapiro–Wilk Test and Rejected Grubbs' Test ($G > G_{max}$)
September-October-November	80	8.0	3782	54
October-November-December	76	7.6	3618	52
November-December-January	88	8.8	4362	49
December-January-February	90	9.0	4473	51
January-February-March	89	8.9	4429	60
February-March-April	87	8.7	4295	50
March-April-May	83	8.3	4065	56
April-May-June	85	8.5	4243	37
May-June-July	88	8.8	4325	32
June-July-August	86	8.6	4172	47
July-August-September	90	9.0	4338	70
August-September-October	57	5.7	2733	40
Total	*999*	100	48,835	*598*
%				*59.9*

After distinguishing between accepted (*A*) and rejected (*R*) 3-month periods, we built the box plots on the stations' elevations in both cases (not shown). No difference was found, suggesting that station elevation was not a factor in explaining normality rejection. In addition, we computed (Table 2) the respective number of station-months (*N*), contribution to the annual total *r* (%), mean f_0, mean sample size, and mean sample asymmetry coefficient (C_s).

Table 2 shows that the mean sample size was not sensitive. On the contrary, the series where a Shapiro–Wilk test was rejected are associated with a larger percentage of Grubbs' test rejection (64.1% vs. 59.9%). Additionally, their mean asymmetry coefficients were higher (1.41) in comparison with those corresponding to series with an accepted Shapiro–Wilk test (0.92). The rejected periods had generally a higher f_0, with a mean of 0.041, compared with that of accepted periods, with a mean f_0 = 0.011.

Most of the 999 station-months with an accepted Shapiro–Wilk test (Table 1) displayed a frequency of zero rainfall during 3-month periods f_0 = zero (761), representing nearly 76%. The corresponding rate was smaller in the case of station-months with a rejected Shapiro–Wilk test. A total of 138 out of 206 displayed f_0 = zero, representing nearly 67%. Thus, f_0 appeared to be the most sensitive factor, followed by C_s, and then by the existence of a maximum detected outlier.

Table 2. Characteristics of accepted (*A*) and rejected (*R*) 3-month periods; *N* number of station-months; contribution to the annual total. (%).

3-Month Period	*N* (*A*)	*r*% (*A*)	Mean f_0 (*A*)	Mean *N* (*A*)	Mean C_s (*A*)	*N* (*R*)	*r*	Mean f_0 (*R*)	Mean *N*(*R*)	Mean C_s (*R*)
September-October-November	80	8.0	0.027	47	1.18	16	7.8	0.054	42	1.56
October-November-December	76	7.6	0.005	48	1.06	20	9.7	0.000	42	2.56
November-December-January	88	8.8	0.005	50	0.88	14	6.8	0.000	46	1.95
December-January-February	90	9.0	0.004	50	0.86	12	5.8	0.000	46	0.77
January-February-March	89	8.9	0.002	50	0.93	12	5.8	0.015	48	1.13
February-March-April	87	8.7	0.004	49	0.82	15	7.3	0.002	49	1.05
March-April-May	83	8.3	0.003	49	0.85	18	8.7	0.005	52	1.18
April-May-June	85	8.5	0.004	50	0.66	16	7.8	0.007	46	1.08
May-June-July	88	8.8	0.006	49	0.60	13	6.3	0.004	49	0.70
June-July-August	86	8.6	0.008	49	0.83	15	7.3	0.002	49	1.08
July-August-September	90	9.0	0.019	48	1.24	11	5.3	0.042	48	2.09
August-September-October	57	5.7	0.054	48	1.33	44	21.4	0.150	48	1.39
Total	999	100				206	100			

The period August-September-October has the highest rate of rejection (21.4%) of *SPI-3* normality, representing 2.5 to 4 times the rejection rates of the other periods (Table 2). It is associated with the highest average f_0 (0.15). Figure 9 shows for the case of accepted series, and for every 3-month period of the year, the contribution to the annual total of station-months with accepted Shapiro-Wilk test (*r*%) versus the corresponding mean value of f_0 (as seen from Table 2). Expected acceptation rate is (1/12 = 8.3%). From Figure 9, the period August-September-October (r = 0.05%; f_0 = 5.7) is singular. It has much lower acceptation rates than expected.

The analysis of the percentage of rejection with respect to the climate category was undertaken. The network composition was 49%, 20%, and 31% of stations categorized as *Csa*, *Bsh*, and *Bwh*, respectively. In comparison, 56%, 24%, and 20% of the 84 rejected stations were in categories *Csa*, *Bsh*, and *Bwh*, respectively. Thus, the category *Csa* was over-represented in the sample of rejected cases, while *Bwh* was under-represented. Figure 10 compares the histograms when considering the entire network versus the histograms when considering stations with rejections.

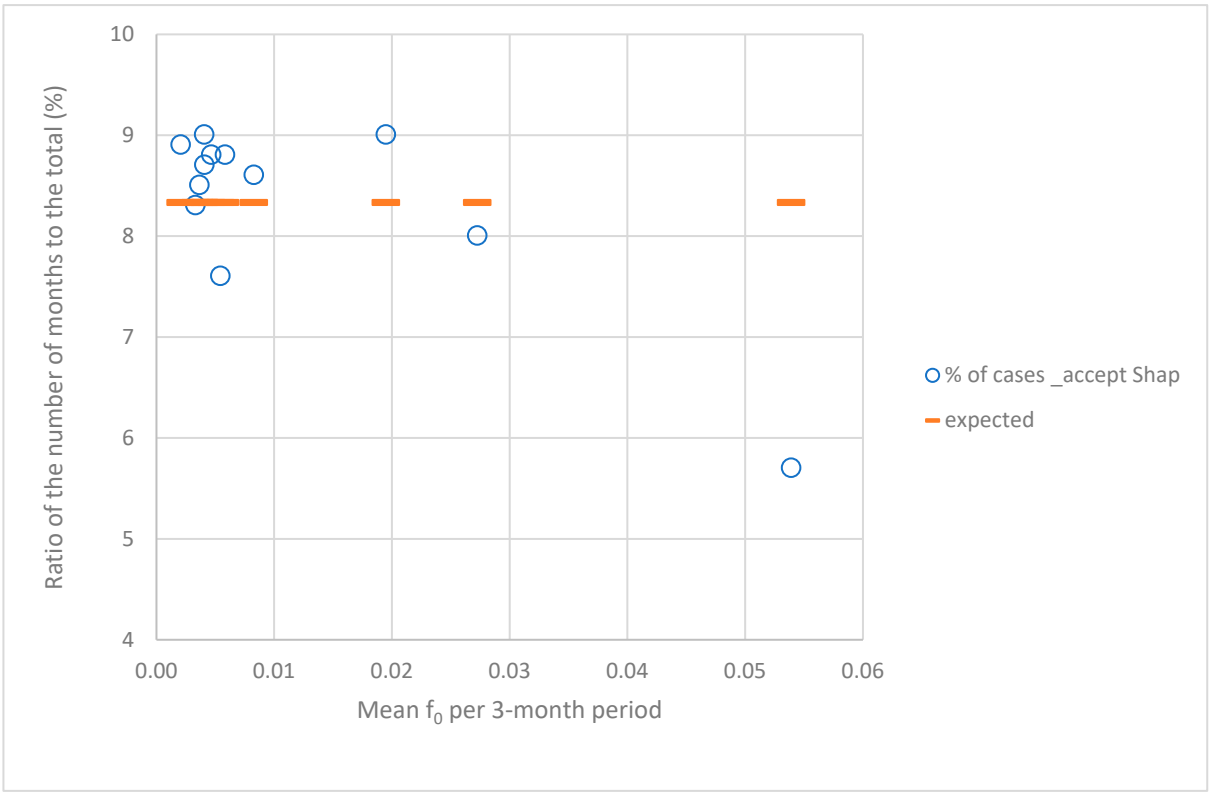


Figure 9. Contribution to the annual total r (%) of an accepted normality test for every 3-month period versus the corresponding mean value of f_0.

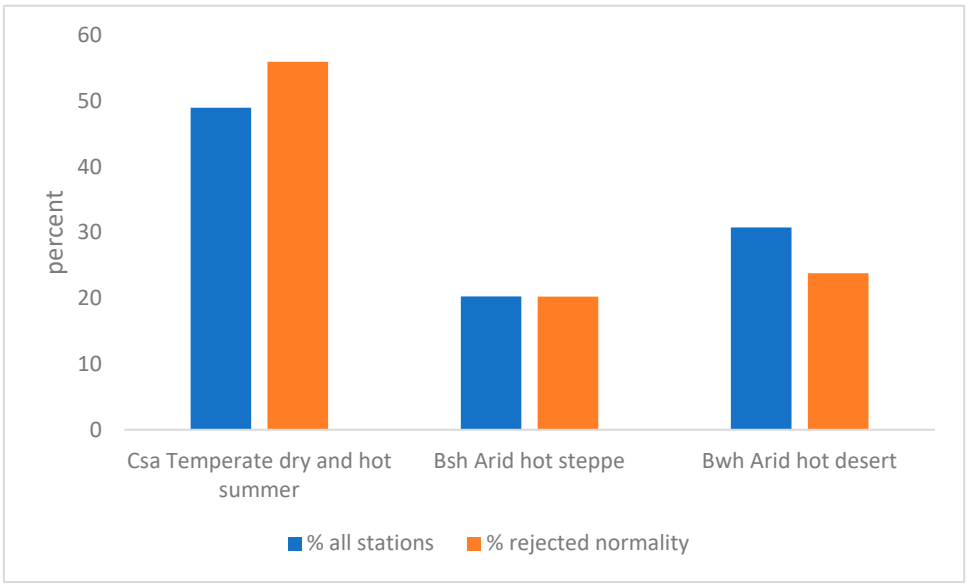


Figure 10. Percentage of stations for every climate category, considering either the entire network (all stations) or only the stations with a normality rejection.

Figure 11 shows the estimated *SPI-3* time series (all stations). As seen in Figure 10, the period of 1950–2018 experienced several droughts, with *SPI-3* less than −2 at station level, highlighting the fact that this region faced extreme agricultural drought periods.

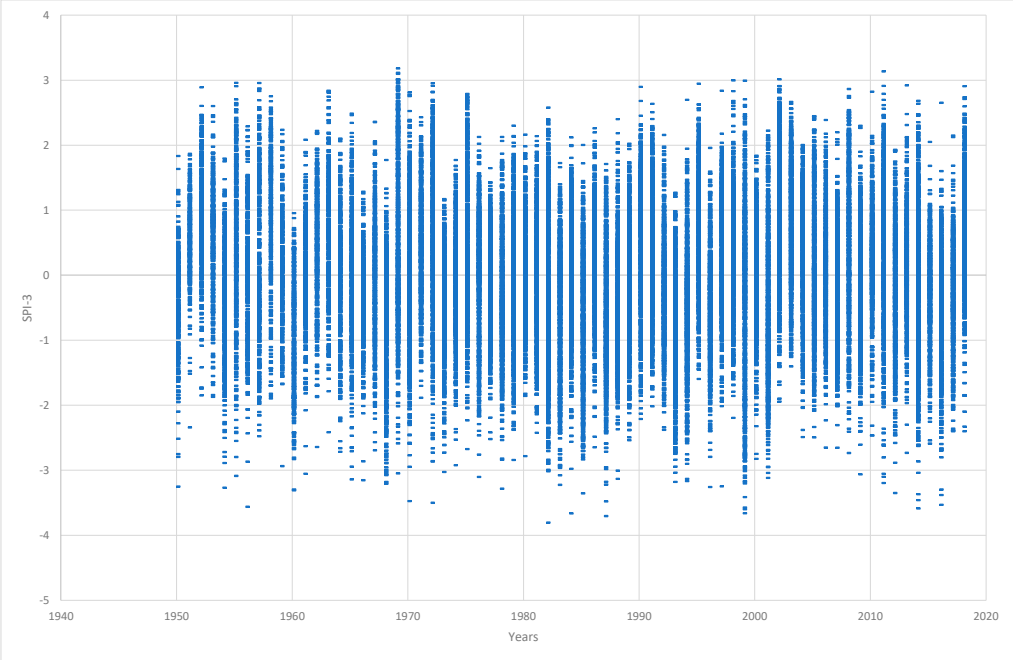


Figure 11. Contribution interstation and time variability of estimated *SPI-3* (1950–2018).

4. Discussion

When considering the non-sensitivity of normality results compared with an increase in the sample size, our results are in conformity with the literature [8]. The rejection of an *SPI-3* normality test was at a rate of 17%, while 3–13% were mentioned in [1], compared with 10% in [11]. One may argue that the gridded data used in [1] contained fewer zeros than that of historical data and that the geographical area studied in [1] included fewer stations under arid and semi-arid conditions. In fact, 0.01 mm was assumed to be the threshold for zero and had few cells with zero accumulation of precipitation in [1]. In addition, [1] fitted daily distributions and computed *SPIs* averaging on a monthly basis for their use of gridded data. When normalizing zero in the *SPI*, the Weibull plotting position is often used for estimating f_0. The Weibull plotting position estimations were averaged using the alternative form $f_0 = (m + 1)/(2N + 1)$, where m was the number of zeros and N was the sample size in [1]. In this method, a depletion operated in the high values of f_0. In our case, such a transformation would effectively be in favor of decreasing the number of cases of rejecting normality.

The percentage of *SPI-3* being less than −3 was 0.12% (63 out of 48,835). The total percentage of values out of the interval (−3, +3) was 0.14%, which is very low in accordance with literature recommendations [1]. The minimum estimated *SPI-3* was −3.803 for station 50535, corresponding to December–February 1982. A study of drought occurrence in the Mediterranean region using *SPI-12*, *SPEI-12*, and two other indicators showed that 1982 achieved the most severe drought in the North Africa region [23]. Conversely, only eight (8) *SPI-3* estimates were greater than +3, and the maximum was 3.181, which corresponded to period 3 (September–November) of 1969 at station 58272. The autumn of 1969 is well-known for its dramatic floods in Tunisia. Thus, our results are accurate in regards to *SPI-3* interpretation when considering the fact that we did not remove outliers.

5. Conclusions

The study of the rainfall network pattern and its time evolution in northern Tunisia within the case of the Medjerda River Basin resulted in retaining 98 out of 144 stations, totaling more than 30 years of observations for every 3-month duration during the period of 1950–2018. This represents a high spatial density of 0.4 station/100 km^2. Based on a massive amount of data (48,835 samples of three-months of rainfall accumulation with sample sizes

greater than 30, covering an area of 23,000 km^2), our study confirmed the potential of the two-parameter gamma distribution for estimating *SPI-3* in North Africa, as well as its drawbacks. We found that the category *Csa* was over-represented in the sample of rejected cases. The most important conclusions are that the presence of outliers as well as the grade of asymmetry coefficient C_s and the importance of f_0 are all responsible for the poor performance of gamma distribution with respect to the normality test of *SPI-3*. In fact, f_0 is the most sensitive factor, followed by C_s, and then by the presence of the maximum detected outlier. Thus, high f_0 and high skewness coefficients are more often attached to a rejection of a Shapiro–Wilk test for *SPI-3* under the gamma distribution hypothesis. Therefore, our results confirm the importance of rainfall intermittency in explaining the gamma model's deficiency in representing *SPI-3* in dry climate conditions. The size of the rainfall time series and the variability of climate conditions in the study region were large enough to extend these conclusions to other sites with similar climate conditions. We therefore propose the adoption of the exponentiated Weibull distribution as an alternative [8] to the gamma in order to estimate SPI. Prospecting heavy-tailed distributions may be another issue.

Author Contributions: Conceptualization, Z.K.B.; data curation, S.J.; formal analysis, S.J.; funding acquisition, Z.K.B.; methodology, Z.K.B.; project administration, Z.K.B.; software, Z.K.B.; validation, S.J.; visualization, S.J.; writing—original draft, Z.K.B. and S.J. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partially funded by the Ministry of Higher Education and Scientific Research (MHESR) of Tunisia through the SWATCH European project of the PRIMA MED program (Project funding reference: 06619.0047.019).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The authors acknowledge the administrative and technical support of the National Engineering School of Tunis.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Stagge, J.H.; Tallaksen, L.M.; Gudmundsson, L.; Van Loon, A.F.; Stahl, K. Candidate Distributions for Climatological Drought Indices (SPI and SPEI). *Int. J. Clim.* **2015**, *35*, 4027–4040. [CrossRef]
2. Lotfirad, M.; Esmaeili-Gisavandani, H.; Adib, A. Drought monitoring and prediction using SPI, SPEI, and random forest model in various climates of Iran. *J. Water Clim. Chang.* **2021**, *13*, 383–406. [CrossRef]
3. Reyes, L.J.C.; Rangel, Á.; Herazo, H.L.C.S. Adjustment of the Standardized Precipitation Index (SPI) for the Evaluation of Drought in the Arroyo Pechelín Basin, Colombia, under Zero Monthly Precipitation Conditions. *Atmosphere* **2022**, *13*, 236. [CrossRef]
4. Enenkel, M.; See, L.; Bonifacio, R.; Boken, V.; Chaney, N.; Vinck, P.; You, L.; Dutra, E.; Anderson, M. Drought and food security —Improving decision-support via new technologies and innovative collaboration. *Glob. Food Secur.* **2015**, *4*, 51–55. [CrossRef]
5. McKee, T.B.; Doesken, N.J.; Kleist, J. The relationship of drought frequency and duration to time scale. In Proceedings of the Eighth Conference on Applied Climatology 1993, Anaheim, CA, USA, 17–22 January 1993; American Meteorological Society: Boston, MA, USA; pp. 179–184.
6. Svoboda, M.; Hayes, M.; Wood, D. *Standardized Precipitation Index User Guide*; WMO-No. 1090; World Meteorological Organization: Geneva, Switzerland, 2012.
7. Vicente-Serrano, S.M.; Begueria, S.; Lopez-Moreno, J.I. A multi-scalar drought index sensitive to global warming: The standardized precipitation evapotranspiration index-SPEI. *J. Clim.* **2010**, *23*, 1696–1718. [CrossRef]
8. Pieper, P.; Düsterhus, A.; Baehr, J. A universal Standardized Precipitation Index candidate distribution function for observations and simulations. *Hydrol. Earth Syst. Sci.* **2020**, *24*, 4541–4565. [CrossRef]
9. Shapiro, S.S.; Wilk, M.B. An analysis of variance test for normality (complete samples). *Biometrika* **1965**, *52*, 591–611. [CrossRef]
10. Naresh Kumar, M.; Murthy, C.S.; Sesha Sai, M.V.R.; Roy, P.S. On the use of Standardized Precipitation Index (SPI) for drought intensity assessment. *Meteorol. Appl.* **2009**, *16*, 381–389. [CrossRef]
11. Blain, G.C.; Meschiatti, M.C. Inadequacy of the gamma distribution to calculate the Standardized Precipitation Index. *Rev. Bras. Eng. Agric. Ambient.* **2015**, *19*, 1129–1135. [CrossRef]

12. Wu, H.; Svoboda, M.D.; Hayes, M.J.; Wilhite, D.A.; Wen, F. Appropriate application of the standardized precipitation index in arid locations and dry seasons. *Int. J. Clim.* **2006**, *27*, 65–79. Available online: http://digitalcommons.unl.edu/droughtfacpub/42 (accessed on 1 September 2017). [CrossRef]
13. Peel, M.C.; Finlayson, B.L.; McMahon, T.A. Updated world map of the Koppen-Geiger climate classification. *Hydrol. Earth Syst. Sci.* **2007**, *11*, 1633–1644. [CrossRef]
14. Achite, M.; Wałęga, A.; Toubal, A.; Mansour, H.; Krakauer, N. Spatiotemporal Characteristics and Trends of Meteorological Droughts in the Wadi Mina Basin, Northwest Algeria. *Water* **2021**, *13*, 3103. [CrossRef]
15. Mellak, S.; Souag-Gamane, D. Spatio-temporal analysis of maximum drought severity using Copulas in Northern Algeria. *J. Water Clim. Change* **2020**, *11*, 68–84. [CrossRef]
16. Habibi, B.; Meddi, M.; Torfs, P.J.; Remaoun, M.; Van Lanen, H.A. Characterisation and prediction of meteorological drought using stochastic models in the semi-arid Chéliff–Zahrez basin (Algeria). *J. Hydrol. Reg. Stud.* **2018**, *16*, 15–31. [CrossRef]
17. Grubbs, F.E. Procedures for Detecting Outlying Observations in Samples. *Technometrics* **1969**, *11*, 1–21. [CrossRef]
18. Bobée, B.; Ashkar, F. The Gamma Family and Derived Distributions Applied in Hydrology. *Water Res. Publ.* **1991**. 203 pages.
19. Thom, H.C.S. A Note on the Gamma Distribution. *Mon. Weather Rev.* **1958**, *86*, 117–122. [CrossRef]
20. Edwards, D.C.; McKee, T.B. Characteristics of 20th Century Drought in the United States at Multiple Times Scales. *Atmos. Sci.* **1997**, *634*, 1–30.
21. Abramowitz, M.; Stegun, I.A. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*; Applied Mathematics Series; National Bureau of Standards: Gaithersburg, MI, USA, 1964; Volume 55.
22. Razali, N.M.; Wah, Y.P. Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests. *J. Stat. Model. Anal.* **2011**, *2*, 21–33.
23. Mathbout, S.; Lopez-Bustins, J.A.; Royé, D.; Martin-Vide, J. Mediterranean-Scale Drought: Regional Datasets for Exceptional Meteorological Drought Events during 1975–2019. *Atmosphere* **2021**, *12*, 941. [CrossRef]