



# Fuzzy rule-based forecast of meteorological drought in western Niger

Zakari Seybou Abdourahamane<sup>1</sup> · Reşat Acar<sup>1</sup>

Received: 6 July 2017 / Accepted: 29 December 2017 / Published online: 12 January 2018  
© Springer-Verlag GmbH Austria, part of Springer Nature 2018

## Abstract

Understanding the causes of rainfall anomalies in the West African Sahel to effectively predict drought events remains a challenge. The physical mechanisms that influence precipitation in this region are complex, uncertain, and imprecise in nature. Fuzzy logic techniques are renowned to be highly efficient in modeling such dynamics. This paper attempts to forecast meteorological drought in Western Niger using fuzzy rule-based modeling techniques. The 3-month scale standardized precipitation index (SPI-3) of four rainfall stations was used as predictand. Monthly data of southern oscillation index (SOI), South Atlantic sea surface temperature (SST), relative humidity (RH), and Atlantic sea level pressure (SLP), sourced from the National Oceanic and Atmosphere Administration (NOAA), were used as predictors. Fuzzy rules and membership functions were generated using fuzzy c-means clustering approach, expert decision, and literature review. For a minimum lead time of 1 month, the model has a coefficient of determination  $R^2$  between 0.80 and 0.88, mean square error (MSE) below 0.17, and Nash-Sutcliffe efficiency (NSE) ranging between 0.79 and 0.87. The empirical frequency distributions of the predicted and the observed drought classes are equal at the 99% of confidence level based on two-sample  $t$  test. Results also revealed the discrepancy in the influence of SOI and SLP on drought occurrence at the four stations while the effect of SST and RH are space independent, being both significantly correlated (at  $\alpha < 0.05$  level) to the SPI-3. Moreover, the implemented fuzzy model compared to decision tree-based forecast model shows better forecast skills.

## 1 Introduction

The Sahel region is an eco-climate zone located on the southern edge of the Sahara Desert. It is a semi-arid area where precipitations are characterized by strong seasonal and interannual variabilities with a short rainy season (4 months) and a long dry season. The average rainfall amount is about 200 to 800 mm from north to south (Djibo et al., 2015). In an inter-seasonal timescale, the precipitations over the West African Sahel are controlled by three main processes: a flow of moist air from the south associated with the West African Monsoon (WAM) onset, the seasonal movement of the ITCZ (inter-tropical convergence zone), and a dry (and aerosol rich) advection from the Sahara (Buontempo, 2010).

In the last three decades of the twentieth century, the Sahel region has experienced a significant drought which corresponds to a discontinuity in the rainfall series. Dai et al. (2004) classified the decay in the precipitation series and the consequent droughts among the most unquestionable and major contemporary changes in climate that have been observed by the scientific community. Drought investigation over the region has therefore become an important component of scientific research; consequently, temporal rainfall variability has been related to the teleconnection with the oceanic basins (Dieppo et al., 2015; Folland et al., 1986; Giannini et al., 2008).

Several drought forecasting models have been developed for the Sahel region using large climate variables as predictors. The methods commonly employed in the prediction models comprises dynamical models (Folland et al., 1991) and statistical methods (Badr and Zaitchik, 2014; Djibo et al., 2015; Lodoun et al., 2014). Hence, skilful predictability was achieved by considering rainfall amount as a single mean value over the whole region. However, the connections between the oceanic-atmospheric dynamics and the Sahelian rainfall vary strongly in space and time, and even the sign of the effect can switch from season to season and between sub-regions of the same country. In addition, the climate change feedback adds to the daunting task

---

✉ Zakari Seybou Abdourahamane  
abdourahamane.zakari.seybou14@ogr.atauni.edu.tr

Reşat Acar  
racar@atauni.edu.tr

<sup>1</sup> Department of Civil Engineering, Atatürk University,  
25030 Erzurum, BP, Turkey

of drought forecast, and the complexity and uncertainty that characterize the predictors make the forecast results hard to transform into traditional mathematical models and therefore are incomprehensible in linguistic form.

Recently, artificial intelligence such as ANN (artificial neural network), fuzzy logic, data mining, wavelet transform, and hybrid models have become popular in drought forecasting and hydrologic parameters modeling. Fuzzy logic (FL) theories were conceived and initiated by Zadeh (1965). Fuzzy theory appears to be extremely effective at handling dynamic, nonlinear, and noisy data, especially when the underlying physical relationships are not fully understood (Nayak and Sudheer, 2008). Similarly, FL models have the facility to include expert understanding of climate processes in linguistic form during the inference process. FL technique has the potential to produce models that are robust, less complex, and more comprehensible in human language. Fuzzy algorithm relies on a systematic use of linguistic expressions to characterize the values of variables and relations between them. It has been successfully used to predict weather events, to forecast drought and rainfall amount, and to model rainfall-runoff processes. Pongracz et al. (1999) applied fuzzy rule-based modeling technique to the prediction of regional drought. Bardossy et al. (2005) utilized a fuzzy rule-based methodology to downscale local hydrological variables from large-scale atmospheric circulation. Ozger et al. (2012) applied a hybrid wavelet-fuzzy logic to forecast long lead drought in Texas. Ozger et al. (2011) predicted the Palmer drought severity index using wavelet-FL model with large-scale climate indices as predictors. Nayak et al. (2014) implemented a fuzzy model identification based on cluster estimation for reservoir inflow forecasting in the Narmada basin, India.

In this paper, we attempt to forecast drought characterized by 3-month scale standardized precipitation index (SPI-3) of four rainfall stations of western Niger using FL modeling techniques. Monthly Southern Oscillation Index (SOI) data obtained from the National Oceanic and Atmospheric Administration Earth System Research Laboratory (NOAA ESRL), and monthly re-analysis data (Kalnay et al., 1996) of South Atlantic sea surface temperature (SST), relative humidity (RH), and sea level pressure (SLP) were used as drought driving forces. A detailed description of the study area as well as the different variables is achieved. FL and the fuzzy c-means clustering methods were described. Finally, the implemented FL forecast model is compared to decision tree-based forecast modeling.

Previous studies have used fuzzy logic to forecast drought by applying empirical methods or by automatically generating the fuzzy rules and membership functions. In this work, in addition to the fuzzy c-means clustering, expert decision and literature review were applied to determine the fuzzy rules and membership functions.

An efficient drought forecast model for Niger can help policy makers, water resource managers, and stakeholders to take the appropriate and sustainable actions for drought preparedness and

mitigation. This paper, presents the perspectives for the implementation of a drought early warning system for the country.

## 2 Materials and methods

### 2.1 Study area

The four rainfall stations, considered in this work, are located in two administrative regions of Western Niger, namely Tillabery and Tahoua (Fig. 1). Both regions are situated between latitudes 12°–18° north and longitudes 0°–16.5° east.

Tillabery region covers 7% of the country's total area, in which lives 17.7% of the country's population (INS-Niger, 2016). The main economic activities are agriculture, livestock raising, and fishery production. It is one of the two regions crossed by the River Niger, which is the only permanent watercourse of the country. The hydrographic network is characterized by an important network of watershed and several seasonal watercourses known as Gorouol, Sirba, Mékrou, Tapoa, and many others. Precipitations are typically Sahelian and are determined by the movement of the ITCZ. In this region, the passage of the ITCZ occurs in June (the beginning of the rainy season) and returns in late September, corresponding to the end of the wet season. The mean annual rainfall is about 400 mm whereas daily mean temperature varies between 17 and 42 °C according to the season (DMN, 2015).

Tahoua, in contrast, is not crossed by the River Niger, thus has no permanent watercourse. Pastoralism and agriculture constitute the main economic activities. The region covers 8% of the country's total area and has 18.5% of the country population (INS-Niger, 2016). The population in its majority is settled in sparse valleys, where small irrigated agriculture is practiced. The region is well known for its violet onion production. The climate is Sahelo-Saharan and seasonal rainfall is about 360 mm. Daily mean temperature fluctuates from 15 to 41 °C depending on the season (DMN, 2015).

In both administrative regions, desertification is a constant treat. For this reason, one of the government and local population's main priorities is to mitigate this phenomenon by implementing land recuperation and reforestation techniques.

### 2.2 Data description

#### 2.2.1 Rainfall Data

The rainfall data used in this paper are monthly precipitation series of four rainfall stations spanning the period 1950–2013 (Table 1). The rainfall stations are located in Western Niger (Fig. 1), and the rainfall series were provided by the National Meteorological Office of Niger. The missing values in the rainfall series represent less than 10% of the total data, and the ratio varies from 3 to 6% for the four stations.

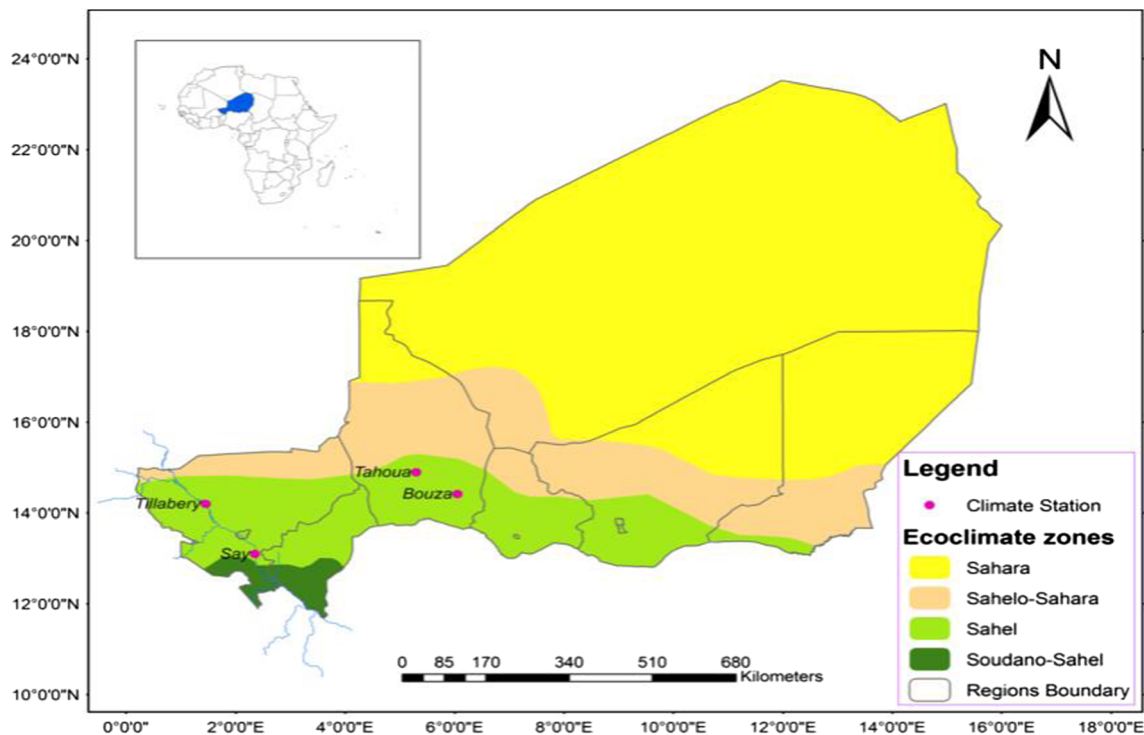


Fig. 1. Location of the rainfall stations

### 2.2.2 Large-scale climate indices

The sea surface temperature (SST) is the most important surface condition affecting the climate. The global warming has undoubtedly increased atmospheric moisture demand and has altered atmospheric-oceanic circulation patterns, thus enhances drought in the Sahel region (Dai, 2011). Researchers have found that El Niño-Southern Oscillation (ENSO) and Tropical Atlantic SSTs had played a significant role in the Sahelian prolonged drought of the 1970s and 1980s (Folland et al., 1991; Folland et al., 1986; Giannini et al., 2008; Giannini et al., 2003). The tropical Atlantic Ocean is considered as the main source of moisture for West Africa; however, after the twentieth century persistent drought, moisture transport over the region has also been related to the Mediterranean basin. Warmer conditions in the Mediterranean Sea enhance Sahelian rainfall (Gaetani et al., 2010), and an increase in the Mediterranean evaporation leads to higher local humidity across the Sahara, therefore providing a supplementary source of moisture for the Sahel (Rowell, 2003). Additionally, for a tropical region such as the Sahel, the assimilation of SLP

besides SST is effective in constraining precipitation values. SLP distribution over the Sahara is linked to temperature difference between the continents and the oceans; therefore, a rise in the surface air temperature across the Sahara in respect to the surrounding oceans reduces the SLP over the Sahara, thus increasing precipitation in the Sahel region (Haarsma et al., 2005).

Based on a literature survey, in this work, the Southern Oscillation Index (a common index representing the ENSO phenomenon), SST from South Atlantic Ocean, relative humidity from Mediterranean basin, and SLP from Atlantic Ocean were considered as drought predictors. The spatial and temporal coverage as well as the details of each predictor are summarized in Table 2.

## 2.3 Data analysis

### 2.3.1 The SPI

Drought as a complex climate phenomenon is difficult to define. Although drought is commonly known as a significant

Table 1 Rainfall stations

Station name	Station type	Station code	Longitude (°)	Latitude (°)	Altitude (m)
Say	Climate station	6,105,400	2.347	13.096	182
Tahoua	METAR/SYNOP	61,043, DRRT	5.250	14.900	386
Tillabery	METAR/SYNOP	61,036, DRRL	1.450	14.200	209
Bouza	Climate station	1,320,003,400	6050	14.417	300

**Table 2** Predictors description

Predictor	Unit	Level	Data source	Region	Spatial coverage	Times series length
SOI	–	–	NOAA ESRL	Equatorial Pacific Ocean	–	January 01, 1949–December 31, 2013
SST	°C	Surface	NOAA NCEP-DOE	South Atlantic Ocean	1°S–20°S; 30°W–10°E	
RH	%	1000 mb	Reanalysis	Mediterranean Basin	40°N–30°N; 20°E–35°E	
SLP	Pa	1000 mb		Atlantic Ocean	15°N–45°S; 60°W–10°E	

decrease in rainfall compared to the normal over an area, determining how to objectively characterize it for users’ activity planning and water resource management is a challenging issue. One of the most commonly used index for meteorological drought characterization is the standardized precipitation index (SPI) developed by McKee et al. (1993). It is a useful index that helps understand the onset and end of meteorological drought, determining so, the duration and severity of dry and wet spells. The SPI also has the flexibility to be computed at a desired time scale (3, 6, 12, 24, or 48 months). Its computation is simple because precipitation data is the only input variable. The available long-term rainfall data is fitted to a probability distribution. The cumulative probability distribution is then transformed to the standard normal distribution which is the SPI value. Dry or wet conditions are thus classified according to the SPI value presented in Table 3 as proposed by McKee et al. (1993). A large number of references are available on the SPI computation; for more details, the reader can refer to McKee et al. (1993), Guttman (1999), and Hayes et al. (1999).

In this paper, 3-month scale SPI was computed by fitting the accumulated rainfall amount of July–September (JAS) to a two-parameter Gamma distribution. The parameters of the distribution were computed using maximum likelihood method. The three-month scale standardized precipitation index (SPI-3) reflects short- and medium-term moisture conditions and provides a seasonal estimation of precipitation (WMO, 2012). Additionally, 80% of the annual rainfall over the Sahel region is recorded during July–September (Djibo et al., 2015). Moreover,

the two-parameter Gamma distribution is the commonly applied distribution for rainfall data (Karavitis et al., 2011). The cumulative probability of the two-parameter Gamma distribution is

$$G(x) = \frac{\int_0^x \alpha^{\alpha-1} e^{-\frac{x}{\beta}} dx}{\beta^{\alpha} \Gamma(\alpha)} \tag{1}$$

where  $\alpha$  is the shape parameter,  $\beta$  is the scale parameter, and  $\Gamma$  is the gamma function. The gamma function is undefined for  $x = 0$  and a precipitation distribution may contain zeros; therefore, the cumulative probability becomes

$$H(x) = q + (1-q)G(x), \tag{2}$$

where  $q$  is the probability of zero.

**2.3.2 Fuzzy logic forecast approach**

Mendel (1995) defined the fuzzy logic system (FLS) as a non-linear representation of an explanatory variable vector into a scalar response variable. The FLS generally consists of three principal components: input/output membership functions (with values ranging between 0 and 1), fuzzy rules (FR), and an inference engine (Mamdani, 1974).

The implementation of a fuzzy logic forecast model based on multiple inputs/single output generally follows these steps: Input and output data are first partitioned into *Learning* and *Test* set, then the variables of the learning set or universe of discourse (crisp sets or conventional set)  $U^P$  are partitioned into fuzzy subsets  $A_n^p$ ; each fuzzy subset is characterized by a membership function  $\mu_A, \mu_A = U^P \rightarrow (0, 1)$ , and the membership function defines the degree of belonging of any  $x \in U^P$  to a fuzzy subset  $A_n^p$ . A linguistic variable is assigned to each fuzzy subset  $A_n^p$ . Subsequently, the *IF and Then* rules are used to define the existing relationship between individual input and output variables. The resulting fuzzy subsets of the output variable are then transformed into a conventional (classical) set throughout a defuzzification technique. In the last step, the forecast model is estimated or validated by using the unseen values of the *test* set.

To implement the forecast model, a matrix database  $D$ , comprising the input and the output data, was created.

**Table 3** Drought classification based on SPI

SPI interval	Drought class
2.0 and more	Extremely wet
1.5 to 1.99	Very wet
1.0 to 1.49	Moderately wet
– 0.99 to .99	Near normal
– 1.0 to – 1.49	Moderately dry
– 1.5 to – 1.99	Severely dry
– 2 and less	Extremely dry

**Table 4** Linguistic variable of input and output data

Data	Linguistic variables					
SOI	Strong El Niño	Weak El Niño	Neutral	Weak La Niña	Strong La Niña	Extreme La Niña
SST	Very low	Low	Medium	High	Very high	Extremely high
SLP	Very low	Low	Medium	High	Very high	Extremely high
RH	Very low	Low	Medium	High	Very high	Extremely high
SPI-3	Severely dry	Dry	Near normal	Wet	Very wet	Extremely wet
Cluster	1	2	3	4	5	6

Seventy percent (70%) of the data was used as *training* set (1949–1995) and 30% as *test* set (1996–2013):

$$D = \begin{pmatrix} 1950 & X1_{1950} & X2_{1950} & X3_{1950} & X4_{1950} & Y_{1950} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 2013 & X1_{2013} & X2_{2013} & X3_{2013} & X4_{2013} & Y_{2013} \end{pmatrix}$$

The first column of the matrix D represents the year; *X1*, *X2*, *X3*, and *X4* are the climate indices used as predictors; and *Y* is the SPI-3 index (the predictand).

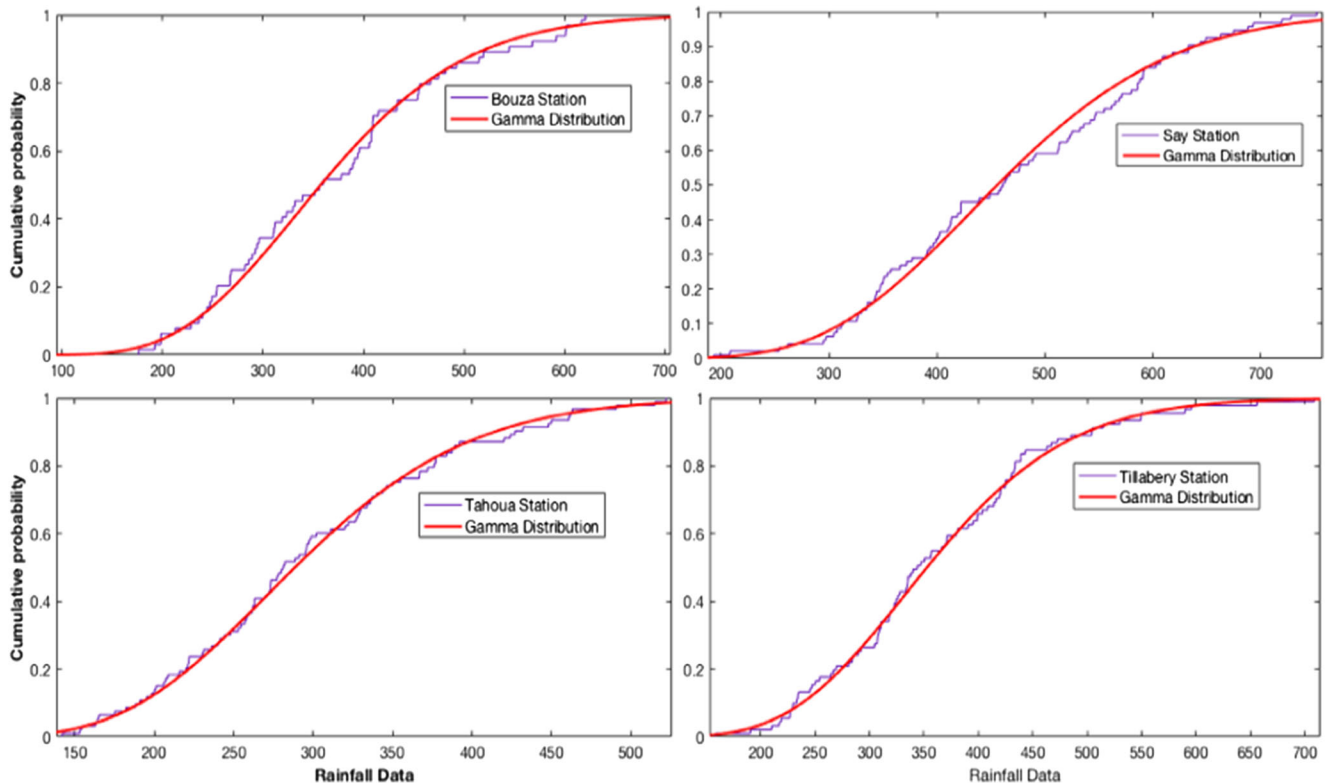
The variables in the training set were then fuzzified using fuzzy c-means clustering techniques, expert decision, and literature review. Fuzzy c-means (FCM) is a method of clustering which allows one piece of data to belong to two or more

clusters. It was first reported in the literature by Dunn (1973) and was improved by Bezdek (1973). The FCM clustering algorithm involves unsupervised learning based on the Euclidean distance between a data vector and a cluster center. It successfully clusters the samples and helps in identifying inherent laws in various samples.

The clustering algorithm of the FCM computes the optimal memberships by minimizing the objective function in Eq. 3 (Suganya and Shanthi 2012):

$$J_m = \sum_{i=1}^N \sum_{j=1}^C U_{ij}^m \|x_i - C_j\|^2, \quad 1 \leq m \leq \infty \quad (3)$$

where *m* is any real number greater than 1, *U<sub>ij</sub>* is the degree of membership of *x<sub>i</sub>* in cluster *j*, *x<sub>i</sub>* is the *i*th of *d*-



**Fig. 2** Rainfall series fitted to the Gamma two-parameter distribution

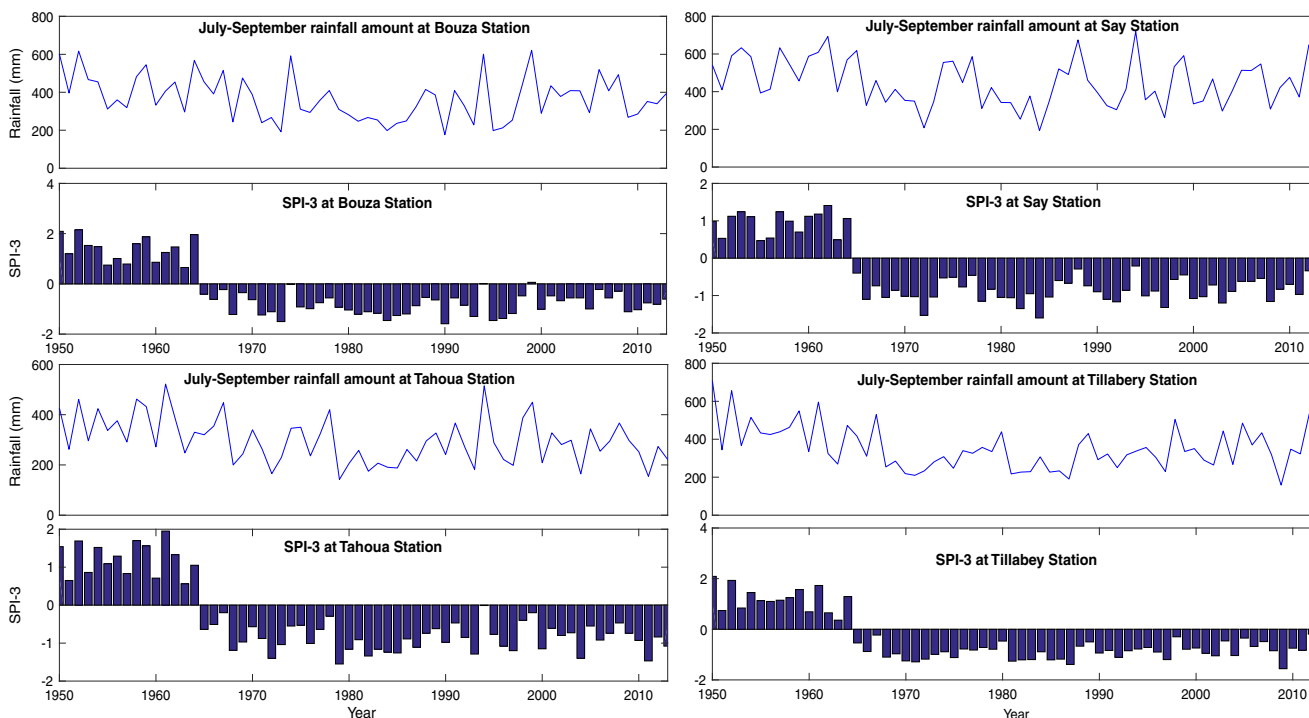


Fig. 3 Seasonal rainfall and SPI-3 series

dimensional measured data,  $C_j$  is the center of cluster  $j$ , and  $\|x_i - C_j\|$  are Euclidean distances between the data point and the cluster centers. Fuzzy partitioning is carried out through an iterative optimization of the objective function shown in Eq. 3, with the update of the degree of membership  $U_{ij}$  and the cluster centers  $C_j$  by

$$U_{ij} = \frac{1}{\sum_k^C \left( \frac{\|x_i - C_j\|}{\|x_i - C_k\|} \right)^{\frac{2}{m-1}}} \tag{4}$$

$$C_j = \frac{\sum_{i=1}^N U_{ij}^m * x_i}{\sum_{i=1}^N U_{ij}^m} \tag{5}$$

This iteration will stop when  $\max_{ij} \left\{ \left| U_{ij}^{(k+1)} - U_{ij}^{(k)} \right| \right\} < \epsilon$ , where  $\epsilon$  is a termination criterion between 0 and 1, whereas  $k$  is the iteration steps. This procedure converges to a local minimum or a saddle point of the cluster  $J$ . The number of clusters

determines the number of rules and membership functions in the generated FIS.

To transform the fuzzified results, obtained from the fuzzy rules into conventional set, centroid defuzzification method was used. The centroid defuzzification method defines the centroid coordinate of a fuzzy number  $\hat{A}$  in horizontal axis as its defuzzified value.

### 2.3.3 Expert decision and literature review

Membership functions (MFs) and number of clusters were obtained through the FCM clustering technique.

The interval domain of each cluster and MF were adjusted using expert decision considering the nature of the variable. For the SOI variable, the values that define El Niño and La Niña phases were used beside the FCM method as an additional criterion to determine the MF domain of definition. A similar classification to the SOI classes shown in Table 4 was previously used by Pongracz

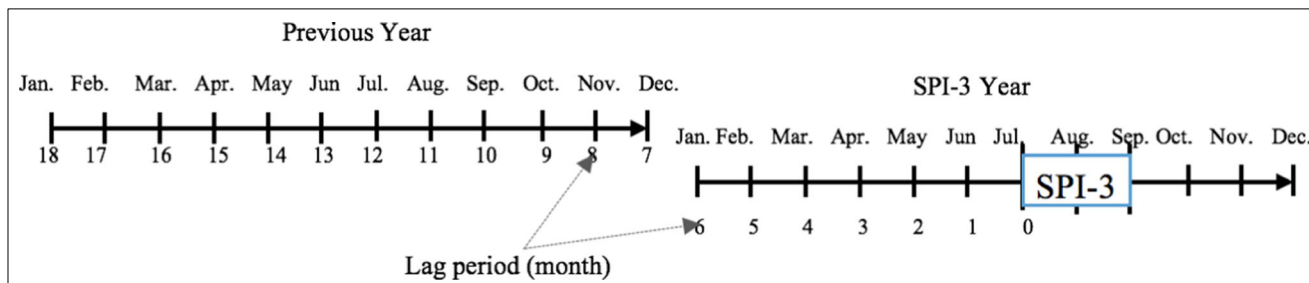


Fig. 4 Lag time between SPI-3 and predictors

et al. (2001) in the prediction of monthly precipitation in Hungary. The partition of the SST, SLP, and RH variables into fuzzy subsets was adjusted using the mean, the maximum, and the minimum values as a supplementary criterion to the FCM approach. As for the SPI-3 variable, the drought classification developed by McKee et al. (1993) was used to interpret the clusters of the output variable generated by the FCM. Hence, a linguistic variable was assigned to each cluster, and the combination of the linguistic variables of the input and output data was analyzed by means of the *IF and Then* rules.

The model was developed in MATLAB\_R2015a. The linguistic variables representing each cluster are shown in Table 4.

### 3 Results and discussions

#### 3.1 Three-month scale standardized precipitation index

The three-month scale standardized precipitation index (SPI-3) was used as drought quantifying parameter in a fuzzy logic-based forecast of meteorological drought in Western Niger. The SPI-3 was computed using two-parameter Gamma distribution. The goodness of fit of the rainfall series to the Gamma two-parameter distribution can be seen in Fig. 2.

The SPI-3 series are presented in Fig. 3 in the form of bar chart. In this figure, drought onset and intensities during the Sahelian prolonged drying are clearly displayed. The drought precisely started in 1966 at Say and in 1968 at the other three stations. As we can observe, no recovery in drought intensities was observed. Yet, lower values of SPI-3 were observed after the 1990s that corresponds to the recovery period mentioned in the introduction. These results agrees with L'Hote et al. (2002) who analyzed a Sahelian rainfall index from 1896 to 2000 and found that the drought persists after the year 2000. The findings of L'Hote et al. (2002) were further confirmed by Ali and Lebel (2009). The drought index used in this paper is

only based on precipitation data; therefore, to make a well-founded inference on the persistence of the Sahelian prolonged drought, other parameters such as evapotranspiration and soil moisture need to be considered.

#### 3.2 Relationship between the SPI-3 and the climate indices

In a predictor screening process carried out throughout literature review, a selection of skilful predictors among the climate indices commonly used as Sahelian driving forces was carried out. From this selection, the SOI, RH, SST, and SLP were retained. The selected predictors were individually lagged from 0 to 18 months in relation to the SPI-3. Lag 0 corresponds to July of the SPI-3 year and lag 18 to January of the previous year as explained schematically in the diagram of Fig. 4. The fuzzy rule-based model was then simulated for each lag period.

After running different simulations, the lead time at which the model displays the highest performance was retained as optimal lag time for its corresponding predictor. For a clear understanding of the link predictor-predictand, the correlation coefficients R and Kendal  $\tau$  were computed. The optimal lag time and the correlation coefficients are presented in Table 5. As can be seen in the afore-mentioned table, the SOI and the SLP optimal lag times are station-dependent. The influence of the SOI appears to be delayed to 1 month at Bouza, Tahoua and Tillabery stations whereas the lead time is 3 months at Say station. At the optimal lag time, the SOI is negatively correlated to the SPI-3 of Tahoua while it is positive for the other stations.

Although it is well documented that the connexion between Sahelian summer rainfall and SOI has increased during the last three decades of the twentieth century (Janicot et al., 1996), the spatial and temporal variation of this relationship has not been investigated in detail. The results obtained here reaffirm the spatiotemporal variability of the ENSO effects over the Sahel region at a rainfall station scale.

The SLP influence on the SPI-3 occurs at a lead time of 8 months at Bouza, 10 months at Say and Tahoua, and

**Table 5** Lag time and correlation predictors-predictand

Predictor	Stations											
	SPI of Bouza			SPI of Say			SPI of Tahoua			SPI of Tillabery		
	R	$\tau$	Lag	R	$\tau$	Lag	R	$\tau$	Lag	R	$\tau$	Lag
SOI	0.18	0.11	-1	0.32*	0.24	-3	-0.26*	-0.18	-1	0.25*	0.15	-1
SST	-0.36*	-0.18*	-3	-0.35*	-0.19	-3	-0.40**	-0.19*	-3	-0.37**	-0.25**	-3
RH	0.68**	0.33*	-1	0.64**	0.35**	-1	0.63**	0.31**	-1	0.68**	0.32**	-1
SLP	-0.36**	-0.22*	-8	-0.32*	-0.22*	-10	-0.36**	-0.25**	-10	-0.32*	0.18*	-15

Note: R is used for the Pearson correlation coefficient and  $\tau$  represents the Kendal  $\tau$

\*Significant at 5% level; \*\*significant at 1% level

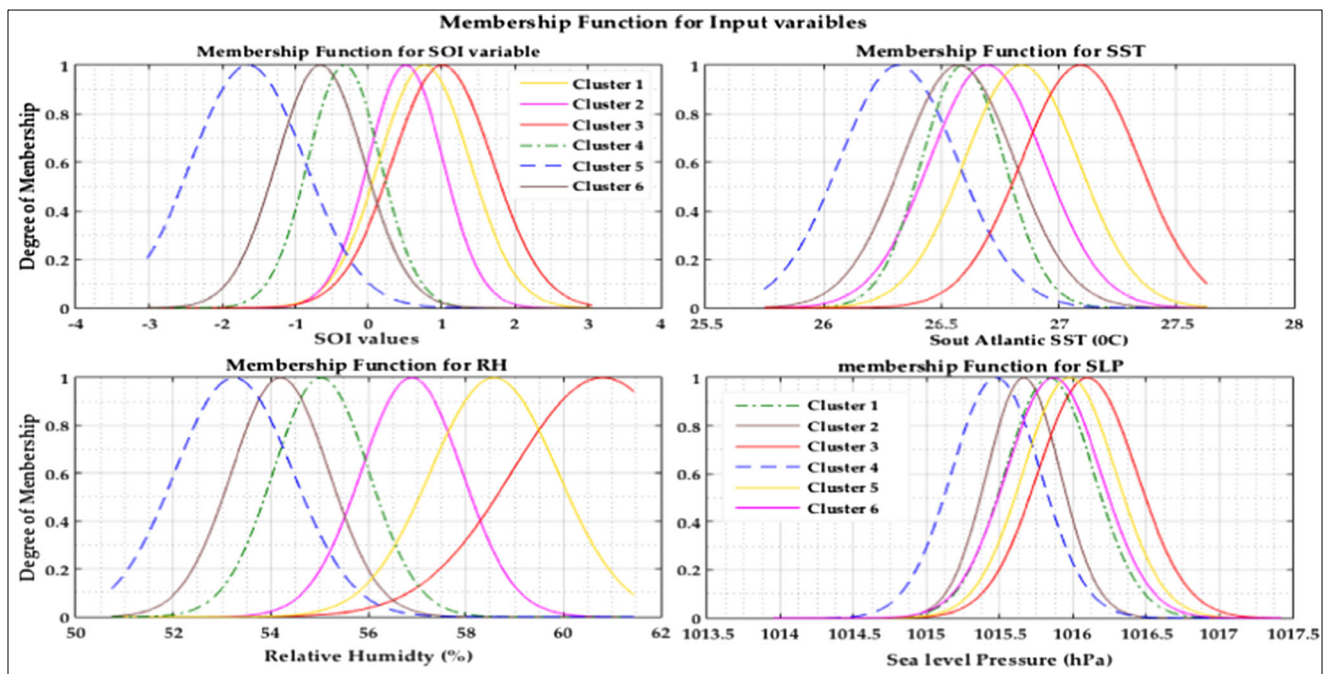


Fig. 5 Membership function for input variables

15 months at Tillabery station. This confirms the temporal variability of the SLP influence on Nigerien rainfall pattern. However, the significant and negative correlation coefficient (Table 5) attests to the spatial uniformity of the positive impact of the SLP on precipitation over the country. The optimal lag time for the SST is 3 months with a negative correlation to the SPI-3 at all the stations—that is to say, warmer phases of SST from the South Atlantic in April enhances drought over Niger during July–September. Similarly, in a previous study using the climate model simulation of Goddard Institute for Space Studies, Druyan (1991) found that warmer SST in the South Atlantic, starting from 1st of March, tends to reduce June–August rainfall over the Sahel. The negative influences of this SST on Sahelian rainfall have also been emphasized in several studies (Hastenrath, 1984; Janicot, 1992; Rodriguez-Fonseca et al., 2015).

The relationship between the RH and the SPI-3 was found to be the strongest with a lead time of 1 month. The link between RH and SPI-3 is positive and significant at the 1% level. Therefore, RH from the Mediterranean

basin enhances rainfall over Niger. The robust connexion between Sahelian rainfall and RH from the Mediterranean was proved by Djibo et al. (2015) who qualified this climate variable as one of the best predictors for seasonal rainfall forecasting over the Sahel.

From the above results, the effects of the relative humidity from the Mediterranean basin and the South Atlantic SST on Western Niger SPI-3 index are space independent and both variables can be used as skilful predictors. However, the use of El Niño and SLP for drought predictions over Niger should be handled with caution. Additionally, a minimum lead time of 1 month was obtained for the selected predictors. One month can be sufficient for drought risk managers and stakeholders to take appropriate actions for drought preparedness and water resource planning. In addition, the input variables used are forecasted each year by the NOAA and are opened to the public (available online at <http://www.esrl.noaa.gov>). Therefore, the implemented FL forecast model can be run at the beginning of each calendar year.

### 3.3 Model efficiency

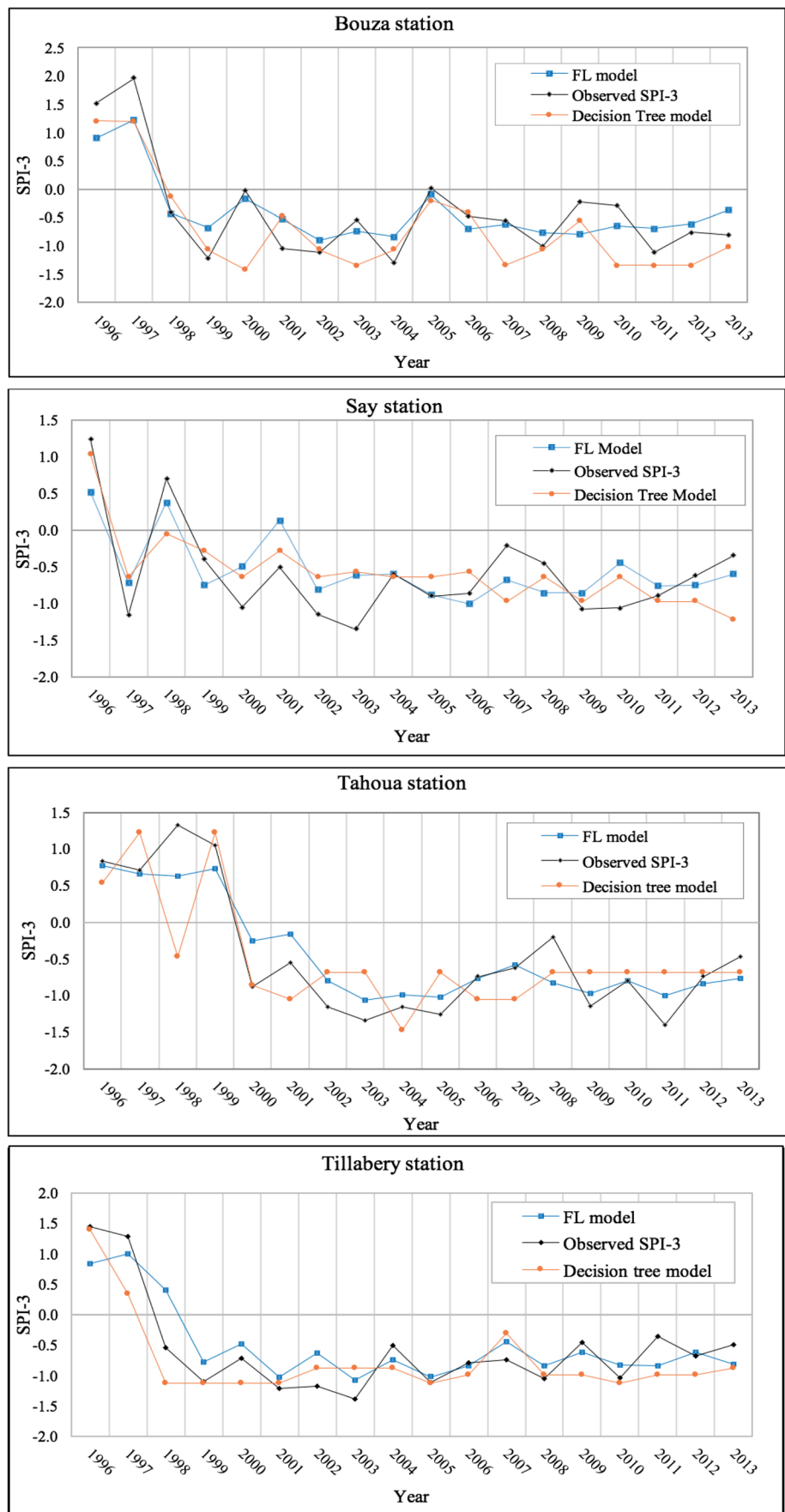
The number of membership function as well as the FIS type was obtained using fuzzy C-means clustering approach adjusted by expert decision and literature review. Mamdani FIS type with six membership functions (Fig. 5) for input and output variables were found to perform the best forecast results. The model efficiency was assessed by computing the mean absolute error (MAE), Nash-Sutcliffe efficiency (NSE), and coefficient of determination  $R^2$  (Table 6). The FL model was then

Table 6 Model efficiency

Climate Station	Fuzzy Logic Model			Decision tree model		
	MSE	NSE	$R^2$	MSE	NSE	$R^2$
Bouza	0.15	0.87	0.88	0.36	0.55	0.69
Say	0.17	0.81	0.81	0.22	0.53	0.55
Tahoua	0.14	0.81	0.84	0.33	0.55	0.56
Tillabery	0.14	0.79	0.80	0.18	0.73	0.79



**Fig. 6** Predicted and observed SPI-3



compared to decision tree (DT)-based forecast model implemented in MATLAB\_R2015a with the function “*classregtree*.” A comparison between the MAE, NSE, and  $R^2$ , presented in Table 6, shows clearly that the FL model yields the best forecast skills at the four rainfall stations. Moreover, the frequency distribution of the forecasted and observed drought classes are computed and presented in Fig. 6. It resulted that the frequency distribution of predicted and observed drought classes for the FL model are equal at 99% of level of confidence. Furthermore, the forecast model is found to be more skilful for *near normal* (NN) and *moderate dry* (MD) drought class prediction compared to the other drought classes (Fig. 7).

The implemented forecast model has an average performance of  $R^2 = 0.83$ , using the rainfall data of four stations. The model has the ability to overcome the complex relationship between drought occurrence in Western Niger and large-scale climate variables using an unsupervised algorithm (fuzzy C-means clustering technique) and can achieve acceptable results. Therefore, it can be applied in other drought-prone areas using different type of predictors. The model has the advantage of being simple and comprehensible in common language; thus, its understanding does not require mathematical background knowledge. Among the limitations of this model is the decrease in its efficiency when predicting extreme events. This is due to the presence of the few number of *extreme wet* and *severely dry* classes in the training set compared to the other classes during the study period. In the other hand, the predictor selection carried out through running different simulations of the model can be time consuming against the presence a large number of predictors.

### 4 Conclusions

Since the Sahelian prolonged drought, the implementation of a skilful drought forecast model for this region has been a challenging issue. The teleconnection between this part of Africa and the oceanic basins have been subjected to strong variabilities, certainly because of the global warming. As a result, the selection of skilful predictor which is the key factor to achieve a good forecast model becomes more and more complex. The model presented in this paper can help to overcome this issue. The application of fuzzy c-means clustering combined with expert decision and literature survey to generate membership functions and fuzzy rules, in a fuzzy logic model for meteorological drought forecasting, yields satisfactory results. The large-scale climate indices, namely, SOI, South Atlantic SST, RH from the Mediterranean basin, and Atlantic SLP, used to implement the model, were found to be suitable drought predictors for the study area. Additionally, it was possible to understand the spatial and temporal variability of the teleconnection between the oceanic basins and drought occurrence in Niger by using the SPI-3 of four different rainfall stations as predictand.

The present model, in contrast to the conventional statistical methods mentioned in the introduction, is simple and comprehensible in human language. It has the ability to incorporate any kind of predictor regardless to the nature (linear or non-linear) of its relationship with drought occurrence. Therefore, the model can be adapted to new predictors that may incorporate the impact of human activities (such as wood cutting, greenhouse effects, urbanization, over grazing, and agriculture).

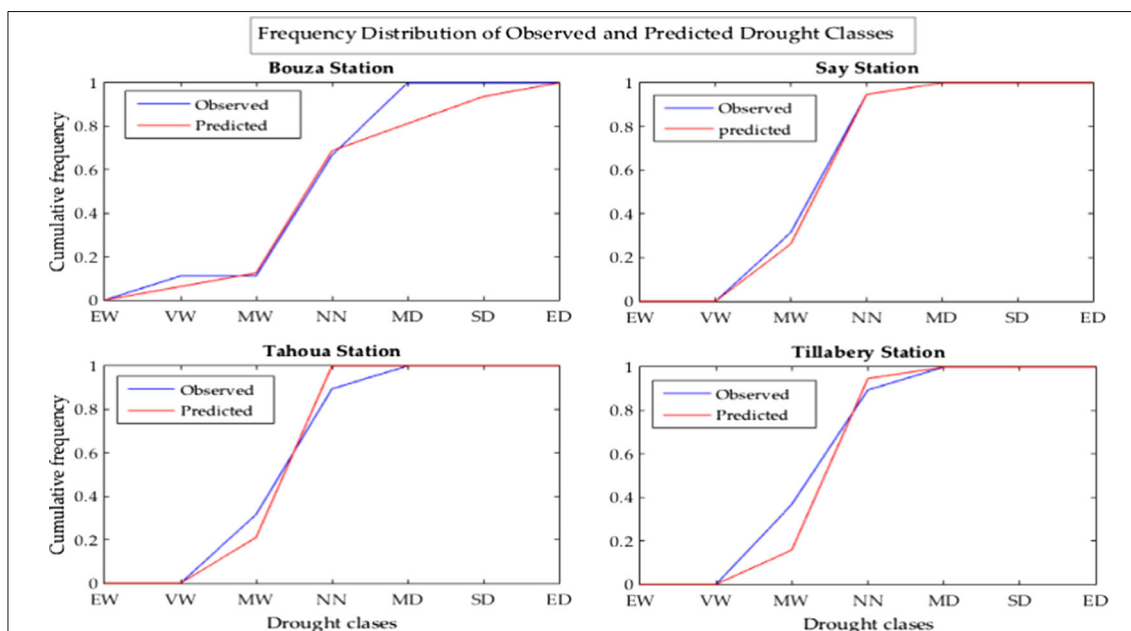


Fig. 7 Frequency distribution of predicted and observed drought classes. Note: In the above figure, EW extremely wet, VW very wet, MW moderately wet, NN near normal, MD moderately dry, SD severely dry, ED extremely dry

The application of this model to the rainfall data of all the available rainfall stations over Niger can be an important step for the elaboration of a drought early warning system for the country. Thus, drought occurrence in the country may be prevented from turning into disaster or state of emergency.

**Acknowledgments** The authors would like to thank the National Meteorological Institute of Niger for providing the monthly rainfall data. The NCEP Reanalysis 2 data were provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, from their Web site at <http://www.esrl.noaa.gov/psd/>

## References

- Ali A, Lebel T (2009) The Sahelian standardized rainfall index revisited. *Int J Climatol* 29(12):1705–1714. <https://doi.org/10.1002/joc.1832>
- Badr HS, Zaitchik BF (2014) Application of statistical models to the prediction of seasonal rainfall anomalies over the Sahel. *J Appl Meteorol Climatol* 53(3):614–636. <https://doi.org/10.1175/jamc-d-13-0181.1>
- Bardossy A, Bogardi I, Matyasovszky I (2005) Fuzzy rule-based downscaling of precipitation. *Theor Appl Climatol* 82(1–2):119–129. <https://doi.org/10.1007/s00704-004-0121-0>
- Bezdek J (1973) *Fuzzy Mathematics in Pattern Classification*. (Ph. D.). Cornell University, Ithaca
- Buontempo, C. (2010). Sahelian climate: past, current, projections. Retrieved from United Kingdom:
- Dai AG (2011) Drought under global warming: a review. *Wiley Interdiscip Rev Clim Chang* 2(1):45–65. <https://doi.org/10.1002/wcc.81>
- Dai A, Lamb PJ, Trenberth KE, Hulme M, Jones PD, Xie P (2004) The recent Sahel drought is real. *Int J Climatol* 24(11):1323–1331. <https://doi.org/10.1002/joc.1083>
- Dieppois B, Durand A, Fournier M, Diedhiou A, Fontaine B, Massei N, Nouaceur Z, Sebag D (2015) Low-frequency variability and zonal contrast in Sahel rainfall and Atlantic sea surface temperature teleconnections during the last century. *Theor Appl Climatol* 121(1–2):139–155. <https://doi.org/10.1007/s00704-014-1229-5>
- Djibo A, Karambiri H, Seidou O, Sittichok K, Philippon N, Paturel J, Saley H (2015) Linear and non-linear approaches for statistical seasonal rainfall forecast in the Sirba watershed region (SAHEL). *Climate* 3(3):727–752. <https://doi.org/10.3390/cli3030727>
- DMN. (2015). Direction de la Météorologie Nationale. Retrieved from <http://www.meteo-niger.org/>
- Druryan LM (1991) The sensitivity of sub-Saharan precipitation to ATLANTIC SST. *Clim Chang* 18(1):17–36. <https://doi.org/10.1007/bf00142503>
- Dunn, J. C. (1973). A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters [- 1973/01/01]. *Journal of Cybernetics*, - 3(- 3), - 57
- Folland CK, Palmer TN, Parker DE (1986) Sahel rainfall and worldwide sea temperatures, 1901–85. *Nature* 320(6063):602–607. <https://doi.org/10.1038/320602a0>
- Folland CK, Owen J, Ward MN, Colman A (1991) Prediction of seasonal rainfall in the Sahel region using empirical and dynamical methods. *J Forecast* 10(1–2):21–56. <https://doi.org/10.1002/for.3980100104>
- Gaetani M, Fontaine B, Roucou P, Baldi M (2010) Influence of the Mediterranean Sea on the West African monsoon: intraseasonal variability in numerical simulations. *J Geophys Res-Atmos* 115(D24):17. <https://doi.org/10.1029/2010jd014436>
- Giannini A, Saravanan R, Chang P (2003) Oceanic forcing of Sahel rainfall on interannual to interdecadal time scales. *Science* 302(5647):1027–1030. <https://doi.org/10.1126/science.1089357>
- Giannini A, Michel M, Verstraete M (2008) A climate model-based review of drought in the Sahel: desertification, the re-greening and climate change. *Glob Planet Chang* 64:119–128. <https://doi.org/10.1016/j.gloplacha.2008.05.004>
- Guttman NB (1999) Accepting the standardized precipitation index: a calculation algorithm. *J Am Water Resour Assoc* 35(2):311–322. <https://doi.org/10.1111/j.1752-1688.1999.tb03592.x>
- Haarsma RJ, Selten FM, Weber SL, Kliphuis M (2005) Sahel rainfall variability and response to greenhouse warming. *Geophys Res Lett* 32(17):4. <https://doi.org/10.1029/2005gl023232>
- Hastenrath S (1984) Interannual variability and annual cycle—mechanisms of circulation and climate in the tropical Atlantic sector. *Mon Weather Rev* 112(6):1097–1107. [https://doi.org/10.1175/1520-0493\(1984\)112<1097:ivaacm>2.0.co;2](https://doi.org/10.1175/1520-0493(1984)112<1097:ivaacm>2.0.co;2)
- Hayes MJ, Svoboda MD, Wilhite DA, Vanyarkho OV (1999) Monitoring the 1996 drought using the standardized precipitation index. *Bull Am Meteorol Soc* 80(3):429–438. [https://doi.org/10.1175/1520-0477\(1999\)080<0429:MTDUTS>2.0.CO;2](https://doi.org/10.1175/1520-0477(1999)080<0429:MTDUTS>2.0.CO;2)
- INS-Niger. (2016). Tablo de Bord Social. Retrieved from Niamey, Niger:
- Janicot, S. (1992). SPATIOTEMPORAL VARIABILITY OF WEST AFRICAN RAINFALL .1. REGIONALIZATIONS AND TYPINGS. *J Clim*, 5, 489–497. doi:[https://doi.org/10.1175/1520-0442\(1992\)005<0489:svowar>2.0.co;2](https://doi.org/10.1175/1520-0442(1992)005<0489:svowar>2.0.co;2), Spatiotemporal variability of west African rainfall. Part I: regionalizations and typings, 5
- Janicot S, Moron V, Fontaine B (1996) Sahel droughts and ENSO dynamics. *Geophys Res Lett* 23(5):515–518. <https://doi.org/10.1029/96gl00246>
- Kalnay E, Kanamitsu M, Kistler R, Collins W, Deaven D, Gandin L, Iredell M, Saha S, White G, Woollen J, Zhu Y, Leetmaa A, Reynolds R, Chelliah M, Ebisuzaki W, Higgins W, Janowiak J, Mo KC, Ropelewski C, Wang J, Jenne R, Joseph D (1996) The NCEP/NCAR 40-year reanalysis project. *Bull Am Meteorol Soc* 77(3):437–471. [https://doi.org/10.1175/1520-0477\(1996\)077<0437:tntyp>2.0.co;2](https://doi.org/10.1175/1520-0477(1996)077<0437:tntyp>2.0.co;2)
- Karavitis CA, Alexandris S, Tsesmelis DE, Athanasopoulos G (2011) Application of the standardized precipitation index (SPI) in Greece. *Water* 3(3):787–805. <https://doi.org/10.3390/w3030787>
- L'Hote Y, Mahe G, Some B, Triboulet JP (2002) Analysis of a Sahelian annual rainfall index from 1896 to 2000; the drought continues. *Hydrol Sci J-J Des Sci Hydrol* 47(4):563–572. <https://doi.org/10.1080/02626660209492960>
- Lodoun T, Sanon M, Giannini A, Traore PS, Some L, Rasolodimby JM (2014) Seasonal forecasts in the Sahel region: the use of rainfall-based predictive variables. *Theor Appl Climatol* 117(3–4):485–494. <https://doi.org/10.1007/s00704-013-1002-1>
- Mamdani EH (1974) Application of fuzzy algorithms for control of simple dynamic plant. *Proc Inst Electr Eng-London* 121(12):1585–1588. <https://doi.org/10.1049/piee.1974.0328>
- McKee TB, Doesken NJ, Kleist J (1993) Relationship of drought frequency. *Am Meteorol Soc* 179–186
- Mendel JM (1995) Fuzzy-logic systems for engineering—a tutorial (Vol 83, PG 345, 1995). *Proc IEEE* 83(9):1293–1293
- Nayak PC, Sudheer KP (2008) Fuzzy model identification based on cluster estimation for reservoir inflow forecasting. *Hydrol Process* 22(6):827–841. <https://doi.org/10.1002/hyp.6644>
- Nayak PC, Sudheer KP, Jain SK (2014) River flow forecasting through nonlinear local approximation in a fuzzy model. *Neural Comput & Applic* 25(7–8):1951–1965. <https://doi.org/10.1007/s00521-014-1684-z>
- Ozger M, Mishra AK, Singh VP (2011) Estimating palmer drought severity index using a wavelet fuzzy logic model based on meteorological variables. *Int J Climatol* 31(13):2021–2032. <https://doi.org/10.1002/joc.2215>

- Ozger M, Mishra AK, Singh VP (2012) Long lead time drought forecasting using a wavelet and fuzzy logic combination model: a case study in Texas. *J Hydrometeorol* 13(1):284–297. <https://doi.org/10.1175/jhm-d-10-05007.1>
- Pongracz R, Bogardi I, Duckstein L (1999) Application of fuzzy rule-based modeling technique to regional drought. *J Hydrol* 224(3–4):100–114. [https://doi.org/10.1016/S0022-1694\(99\)00131-6](https://doi.org/10.1016/S0022-1694(99)00131-6)
- Pongracz R, Bartholy J, Bogardi I (2001) Fuzzy rule-based prediction of monthly precipitation. *Phys Chem Earth Part B-Hydrol Oceans Atmos* 26(9):663–667. [https://doi.org/10.1016/S1464-1909\(01\)00066-1](https://doi.org/10.1016/S1464-1909(01)00066-1)
- Rodriguez-Fonseca B, Mohino E, Mechoso CR, Caminade C, Biasutti M, Gaetani M et al (2015) Variability and predictability of west African droughts: a review on the role of sea surface temperature anomalies. *J Clim* 28(10):4034–4060. <https://doi.org/10.1175/jcli-d-14-00130.1>
- Rowell DP (2003) The impact of Mediterranean SSTs on the Sahelian rainfall season. *J Clim* 16(5):849–862. [https://doi.org/10.1175/1520-0442\(2003\)016<0849:tiomso>2.0.co;2](https://doi.org/10.1175/1520-0442(2003)016<0849:tiomso>2.0.co;2)
- Suganya R, Shanthi R (2012) Fuzzy C- means algorithm- a review. *International Journal of Scientific and Research Publications* 2
- WMO. (2012). Standardized Precipitation Index User Guide. *World Meteorological Organization, WMO-No. 1090*
- Zadeh LA (1965) Fuzzy sets. *Inf Control* 8(3):338–353. [https://doi.org/10.1016/S0019-9958\(65\)90241-X](https://doi.org/10.1016/S0019-9958(65)90241-X)