



Application of Markov chains to Standardized Precipitation Index (SPI) in São Francisco River Basin

ARTICLES [doi:10.4136/ambi-agua.2311](https://doi.org/10.4136/ambi-agua.2311)

Received: 08 Aug. 2018; Accepted: 12 Apr. 2019

Esdras Adriano Barbosa dos Santos^{1*} ; **Tatijana Stosic²** 
Ikaro Daniel de Carvalho Barreto² ; **Laélia Campos³** 
Antonio Samuel Alves da Silva² 

¹Universidade Federal de Sergipe (UFS), São Cristóvão, SE, Brasil
Departamento de Estatística e Ciências Atuariais (DECAT).

E-mail: esdras.adriano@gmail.com

²Universidade Federal Rural de Pernambuco (UFRPE), Recife, PE, Brasil
Departamento de Estatística e Informática (DEINFO).

E-mail: tastosic@gmail.com, daniel.carvalho.ib@gmail.com,
samuelmathematical@gmail.com

³Universidade Federal de Sergipe (UFS), São Cristóvão, SE, Brasil
Departamento de Física (DFI). E-mail: lpbcampos@gmail.com

*Corresponding author

ABSTRACT

This work evaluated dry and rainy conditions in the subregions of the São Francisco River Basin (BHSF) using the Standardized Precipitation Index (SPI) and Markov chains. Each subregion of the BHSF has specific physical and climatic characteristics. The data was obtained from the National Water Agency (ANA), collected by four pluviometric stations (representative of each subregion), covering 46 years of data, from 1970 to 2015. The SPI was calculated for the time scales of six and twelve months and transition probabilities were obtained using the Markov chain. Transition matrices showed that, at both scales, if the climate conditions were severe drought or rainy, switching to another class would be unlikely in the short term. Correlating this information with the probabilities of the stationary distribution, it was possible to find the regions that are most likely to be under rainy or dry weather in the future. The recurrence times calculated for the stations that belong to the semi-arid region were smaller when compared to the value of the return period of the representative station of Upper São Francisco that has higher levels of precipitation, confirming the predisposition of the semi-arid region to present greater chances of future periods of drought.

Keywords: drought, Markov chains, standardized precipitation index.

Aplicação de Cadeias de Markov no Índice de Precipitação Padronizado (SPI) em estações da Bacia do Rio São Francisco

RESUMO

Este trabalho objetivou avaliar períodos secos e chuvosos nas sub-regiões da Bacia Hidrográfica do Rio São Francisco (BHSF) utilizando o Índice de Precipitação Padronizado (SPI) e cadeias de Markov. Cada sub-região da BHSF possui características físicas e climáticas



This is an Open Access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

específicas, dessa forma foram utilizadas quatro estações pluviométricas contendo dados com 46 anos, de 1970 a 2015, obtidas na Agência Nacional de Águas (ANA), representativas de cada sub-região. O SPI foi calculado para as escalas de tempo de seis e doze meses e as matrizes de probabilidades de transição foram obtidas utilizando as cadeias de Markov. As matrizes de transição mostraram que, em ambas as escalas, caso a condição climática estivesse em classe de seca severa ou chuvosa, a mudança para outra classe seria pouco provável a curto prazo. Correlacionando esta informação com as probabilidades da distribuição estacionária, foi possível encontrar as regiões que têm maiores possibilidades de no momento futuro estar sob clima chuvoso ou de seca. Os tempos de recorrência calculados para as estações inseridas no semiárido foram menores quando comparado com o valor do tempo da estação representativa do Alto São Francisco que possui maiores níveis de precipitação, confirmando a predisposição do semiárido em apresentar maiores chances de períodos futuros de seca.

Palavras-chave: cadeias de Markov, índice de precipitação padronizado, seca.

1. INTRODUCTION

Water is considered a renewable natural resource because it is constantly recovering through the processes of the hydrological cycle. However, due to its great potential for use, with demands for different uses, water grows increasingly more scarce. The São Francisco River Basin has experienced this shortage.

The basin is very important to Brazil, not only for the volume of water transported in a semi-arid region, but also for the potential uses of the water and for its historical and economic contributions to the region. The basin covers an area of about 634,000 km². The São Francisco River is approximately 2,700 km long and originates in the Serra da Canastra in Minas Gerais, flowing south-north through Bahia and Pernambuco, where it changes course until it reaches the Atlantic Ocean at the border between Alagoas and Sergipe (Brasil, 2018).

The Northeast Region of Brazil demands special attention in relation to water supply, in particular Ceará, Rio Grande do Norte, Paraíba and Pernambuco, which has 87.8% of its territory in the semi-arid region. Due to little rainfall, the irregularity of its regime, high temperatures throughout the year, strong insolation and high evaporation rates, in addition to some hydrogeological characteristics, water has been very scarce (Brasil, 2018).

Drought is a phenomenon characterized by the absence, scarcity, reduced frequency, limited quantity, and inadequate distribution of rainfall during a year. Drought is considered a natural disaster, being a phenomenon that affects economic, social and political areas. Drought is random phenomenon, since its onset, end and severity are unpredictable (Mishra and Singh, 2010). According to the Brazilian Atlas of Natural Disasters from 1991 to 2012 (UFSC, 2013), drought is the disaster that most affects the Brazilian population, because it is the most frequent, representing 51% of the total records, followed by flood, with 21%.

In order to reduce the vulnerability of populations affected by drought, it is of fundamental importance to develop techniques to monitor this phenomenon, proposing indexes to standardize drought periods (Blain *et al.*, 2010; Mckee *et al.*, 1993; Tsakiris and Vangelis, 2004; Tonkaz, 2006; Sadeghi and Shamseldin, 2014), or evaluating conditions to predict droughts using Markov chains (Keller Filho *et al.*, 2006; Paulo and Pereira, 2007; Lennartsson *et al.*, 2008; Sanusi *et al.*, 2015).

In this context, this work evaluated the periodicity of the drought, and classified the intensity of the drought classes using the Standardized Precipitation Index (SPI) and Markov chains applied to 4 meteorological stations distributed in the subregions of the São Francisco Basin.

2. MATERIALS AND METHODS

2.1. Study area and data

The study area is located in the São Francisco River Basin and the data was provided by the National Water Agency-ANA through the National Information System on Water Resources (SNIRH, 2018).

In this work, 4 stations (one station in each basin sub-region) were used with daily rainfall data for the period from January 1970 to December 2015, which corresponds to 46 years of rainfall records. Table 1 below presents the information about the chosen stations.

Table 1. Description of pluviometric stations.

Sub-region	Latitude	Longitude	Code	Station	State
Upper	-20.178889	-45.700278	2045002	Iguatama	MG
Middle	-10.002222	-42.474167	1042015	Pilão Arcado	BA
Sub-middle	-8.520278	-39.640833	839034	Fazenda Tapera	PE
Lower	-10.215278	-36.823889	1036048	Propriá	SE

2.2. Standardized Precipitation Index (SPI)

The Standardized Precipitation Index was proposed and developed by McKee *et al.* (1993) as a versatile tool in the monitoring and analysis of periods of drought and rain. The SPI is used to estimate the drought condition based on precipitation. To obtain the index, the monthly precipitation values for each meteorological station should be calculated for the desired time scale, which can be 1, 3, 6, 12 and 24 months (Mahmoudzadeh *et al.*, 2016).

Regarding the time scale, the SPI of one month is associated with short-term conditions, the three-month SPI depicts soil moisture conditions in the short- and medium term, in addition to estimating seasonal precipitation. The six-month SPI is associated with irregularities of water supply and river flows, and twelve- and twenty-four-month SPIs express long-term precipitation patterns and are directly associated with water scarcity, with water levels in groundwater flows (Sousa *et al.*, 2016).

With the accumulated monthly precipitation series, the SPI calculation consists of adjusting the Gamma probability density function to the frequency distribution of the precipitation of each of the rainfall stations. The Gamma distribution model has a good fit for continuous variables that do not have an upper limit, being widely used for studies of historical series of precipitation (Wilks, 2011). The probability density function Gamma is given by the Equation 1:

$$f(x) = \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x}{\beta}}, \quad x > 0 \quad (1)$$

Where $\alpha > 0$ is the shape parameter, $\beta > 0$ is the scale parameter, x is the precipitation and $\Gamma(\alpha)$ is Gamma function, defined as Equation 2:

$$\Gamma(\alpha) = \int_0^\infty y^{\alpha-1} e^{-y} dy \quad (2)$$

The parameters α e β of probability density function are estimated for each station by the Maximum Likelihood (Wilks, 2011) (Equations 3, 4 and 5).

$$\hat{\alpha} = \frac{1}{4A} \left(1 + \sqrt{1 + \frac{4A}{3}} \right) \quad (3)$$

$$\hat{\beta} = \frac{\hat{x}}{\hat{\alpha}} \quad (4)$$

$$A = \ln(\bar{x}) - \frac{\sum_i^n x_i}{n} \quad (5)$$

Where: n , x , x_i and A are, respectively, the sample size of the precipitation records, the arithmetic mean, the precipitations observed, and a measure of distribution asymmetry (Husak *et al.*, 2007).

In this work, the SPI was used to assess agricultural drought as well as water scarcity in these regions of the São Francisco Basin. In this way, the SPI was used for six- (SPI-06) and twelve months (SPI-12), being an important index in arid and semi-arid regions (Blain, 2012). In order to calculate the SPI, the Gamma probability density function must first be adjusted to the series of monthly precipitation totals. In addition, the Kolmogorov-Smirnov goodness of fit test adapted by Lilliefors, as described in Blain (2014), was used to test the suitability of the gamma distribution to the data set. The software used in the calculations was the R (R Core Team, 2019). Then, the cumulative probability of occurrence of each monthly total is estimated and the inverse Normal function is applied to this probability resulting in the SPI values. According to McKee *et al.* (1993), the drought period begins when the SPI becomes negative and ends when it returns to positive values (rainy season). Table 2 presents the classification of SPI values used in this study as states in a Markov chain.

Table 2. Classification of SPI values.

Class	SPI values
Extreme drought	$-\infty \leq SPI < -2$
Severe drought	$-2 \leq SPI < -1.5$
Moderate drought	$-1.5 \leq SPI < -1$
Mild drought	$-1 \leq SPI < 0$
Rain	$0 \leq SPI < \infty$

Source: McKee *et al.* (1993).

2.3. Markov Chains

A Markov chain is a stochastic process X_t , such as the probability that X_{t+j} takes a value j at time $t+1$ depends on the past only through its most recent value X_t at time t (Equation 6):

$$P(X_{t+1} = j | X_0, X_1, X_2, \dots, X_t) = P(X_{t+1} = j | X_t = i) = p_{ij} \quad (6)$$

For any one of the $i, j \in S$ and $t \in T$ (Paulo and Perreira, 2007). The transition probability p_{ij} is the probability that the Markov chain is in state j at the next time point, given that it is in state i at the present time point.

The transition probabilities can be expressed in the form of a matrix, namely (Equation 7):

$$p_{ij} = \begin{pmatrix} p_{11} & p_{12} & \dots & p_{1n} \\ p_{21} & p_{22} & \dots & p_{2n} \\ p_{31} & p_{32} & \dots & p_{3n} \\ \vdots & \vdots & \dots & \vdots \\ p_{n1} & p_{n2} & \dots & p_{nn} \end{pmatrix} \quad (7)$$

Where: $0 \leq p_{ij} \leq 1$ e $\sum p_{ij} = 1$; $i = 1, 2, 3, 4 \dots n$.

The Maximum Likelihood method was used to calculate the transition probabilities. The Maximum Likelihood Estimator (MLE) of p_{ij} is given as Equation 8.

$$p_{ij} = \frac{n_{ij}}{\sum_j n_{ij}}, \forall i, j \in S \quad (8)$$

Where n_{ij} is the number of times the observed data went from state i to state j (Mishra *et al.*, 2009).

In long-term processes, the transition probabilities of the Markov chain states are independent of the initial state. Thus, the equilibrium probabilities π_j could be obtained by successive multiplications of the transition matrix itself or by the resolution of the linear system of equations, namely (Equation 9 and 10):

$$\pi_j = \sum_{k \in S} \pi_k p_{kj}, j \in S \quad (9)$$

$$\sum_{j \in S} \pi_j = 1 \quad (10)$$

Where π_j is the long-term probability of drought class j .

The probability of stay, during m months, in class i , is calculated as (Equation 11):

$$P(X_1 = i | X_0 = i) \dots P(X_{m-2} = i | X_{m-1} = i) P(X_m \neq i | X_{m-1} = i) = P_{ii}^{m-1} (1 - P_{ii}) \quad (11)$$

Knowing the probabilities of stay in each drought class i during m months, the expected residence time in any class i , is (Equation 12):

$$E(T_i | X_0) = \sum_k k P(m = k | X_0 = i); m = 1, 2, \dots, k \quad (12)$$

An important measure that helps in the process of assessing drought periods is the recurrence time to a given state. This average time represents the time required for a specific drought class to occur again. This time is then calculated (Equation 13):

$$t_{ii} = \frac{1}{\pi_i} \quad (13)$$

3. RESULTS AND DISCUSSION

The selected stations are distributed in the four subregions of the basin, and therefore have different climates and rainfall levels. For each season, mean precipitations were calculated for all months of the year, as can be seen in the bar chart in Figure 1. Precipitation totals were: 1317 mm for Iguatama station, 715 mm for Pilão Arcado, 481 mm for Fazenda Tapera and 832 mm for the Propriá station, confirming the expected difference between the sub-regions of the basin (Brasil, 2018).

Observing Figure 1, the Iguatama station has higher average precipitations in the months of November to January, especially with average rainfall over 200 mm in the months of December and January. The Pilão Arcado and Fazenda Tapera stations present the same behavior with average monthly rainfall in these months less than 150 mm. The exception is the Propriá station (Lower São Francisco), which assumes its maximum precipitation values in the months of May to June, intercepting the months of least precipitation (from June to August) of the other stations. In particular, the Pilão Arcado, Fazenda Tapera and Propriá stations, although located in different subregions, (Middle, Sub-middle and Lower São Francisco, respectively), are part of the Northeastern Semi-arid, according to the new delimitation of the Semi-arid (Brasil, 2005).

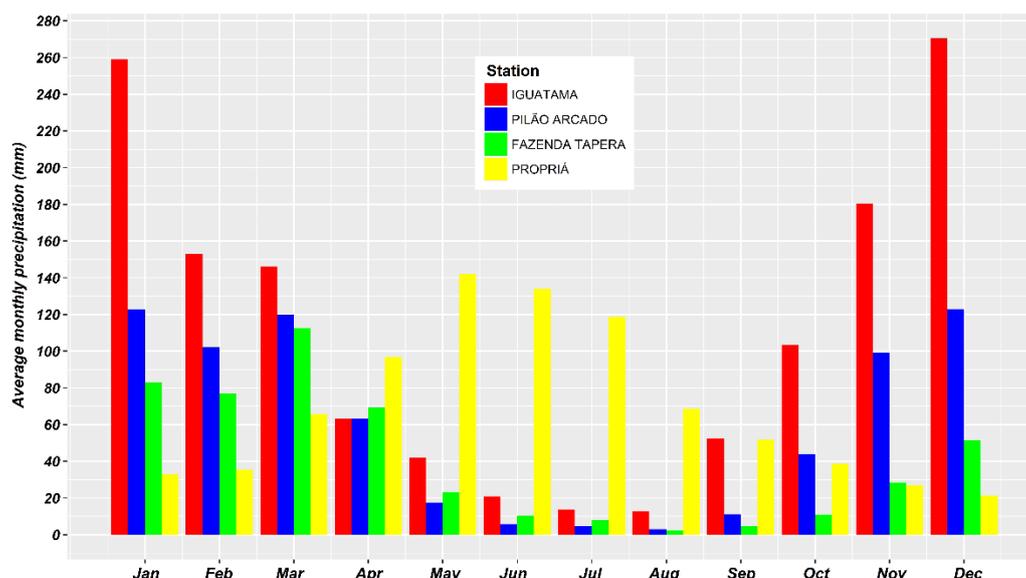


Figure 1. Average monthly precipitation of basin stations.

The SPI is widely used to characterize meteorological droughts (Blain, 2012). For the analysis of drought in the subregions of the basin, the time scales of six- and twelve months were chosen. The six-month SPI compares the precipitation of this period with the same six-month period over historical record, being associated with water storage irregularities and river outflows. The twelve-month SPI expresses long-term precipitation patterns and is directly related to water scarcity (WMO, 2012).

The adequacy of the historical series to the gamma distribution evaluated by the Kolmogorov-Smirnov test adapted by Lilliefors (Blain, 2014) indicated a good fit, as shown in Table 3.

Table 3. P-values of the Kolmogorov-Smirnov test adapted by Lilliefors to time scales of six- (SPI-06) and twelve-month (SPI-12) for basin stations.

Month	Iguatama		Pilão Arcado		Fazenda Tapera		Propriá	
	SPI-06	SPI-12	SPI-06	SPI-12	SPI-06	SPI-12	SPI-06	SPI-12
Jan	0.727	0.728	0.728	0.729	0.727	0.728	0.724	0.727
Feb	0.535	0.540	0.939	0.939	0.273	0.273	0.315	0.314
Mar	0.738	0.737	0.764	0.772	0.339	0.337	0.904	0.905
Apr	0.689	0.691	0.117	0.114	0.143	0.144	0.295	0.294
May	0.665	0.666	0.451	0.452	0.617	0.621	0.583	0.581
Jun	0.043	0.043	0.110	0.110	0.293	0.291	0.476	0.474
Jul	0.402	0.405	0.031	0.032	0.979	0.977	0.310	0.309
Aug	0.195	0.195	0.002	0.002	0.882	0.879	0.916	0.912
Sep	0.062	0.059	0.752	0.753	0.294	0.293	0.546	0.541
Oct	0.131	0.131	0.553	0.552	0.352	0.351	0.700	0.702
Nov	0.039	0.038	0.072	0.067	0.144	0.144	0.011	0.011
Dec	0.400	0.400	0.381	0.381	0.039	0.040	0.506	0.505

Next, SPI values will be presented using the six- (SPI-06) and twelve-month (SPI-12) time scales for the 4 stations representative of the subregions of the basin, for the period from 1970 to 2015, which totals a historical series of 46 years (Figures 2, 3, 4 and 5). For drought classes, the values of SPI are negative, whereas for the rainy condition the values are positive, according to the classification in Table 2.

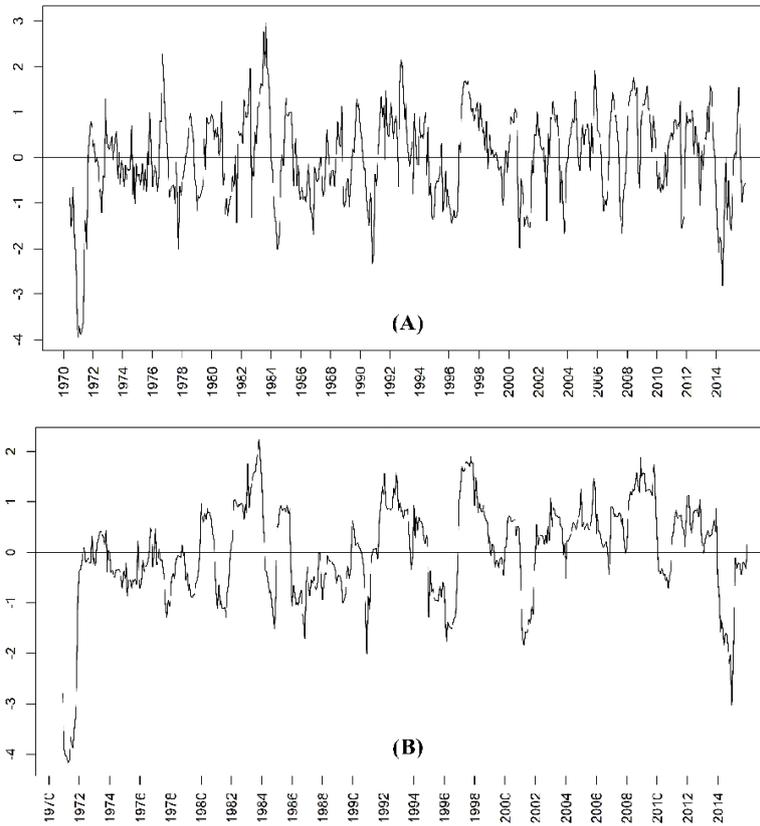


Figure 2. SPI-06 (A) and SPI-12 (B) for Iguatama station.

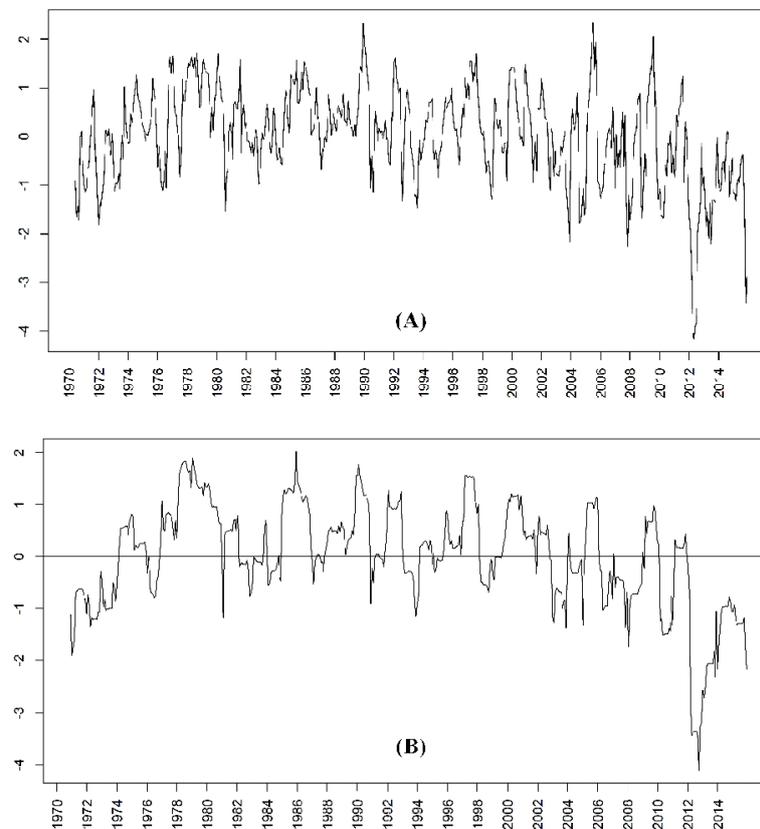


Figure 3. SPI-06 (A) and SPI-12 (B) for Pilão Arcado station.

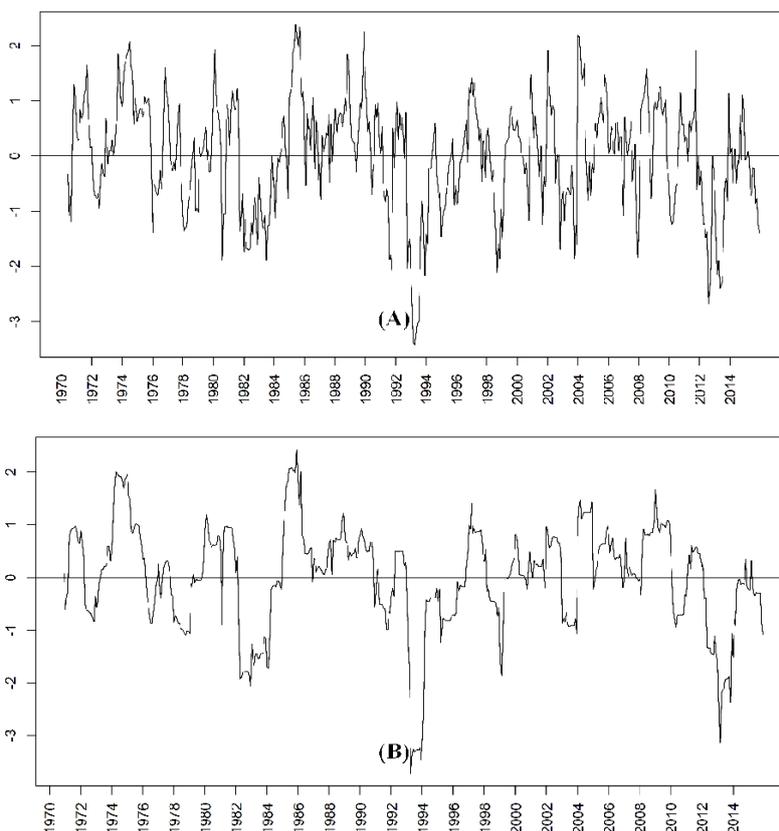


Figure 4. SPI-06 (A) and SPI-12 (B) for Fazenda Tapera station.

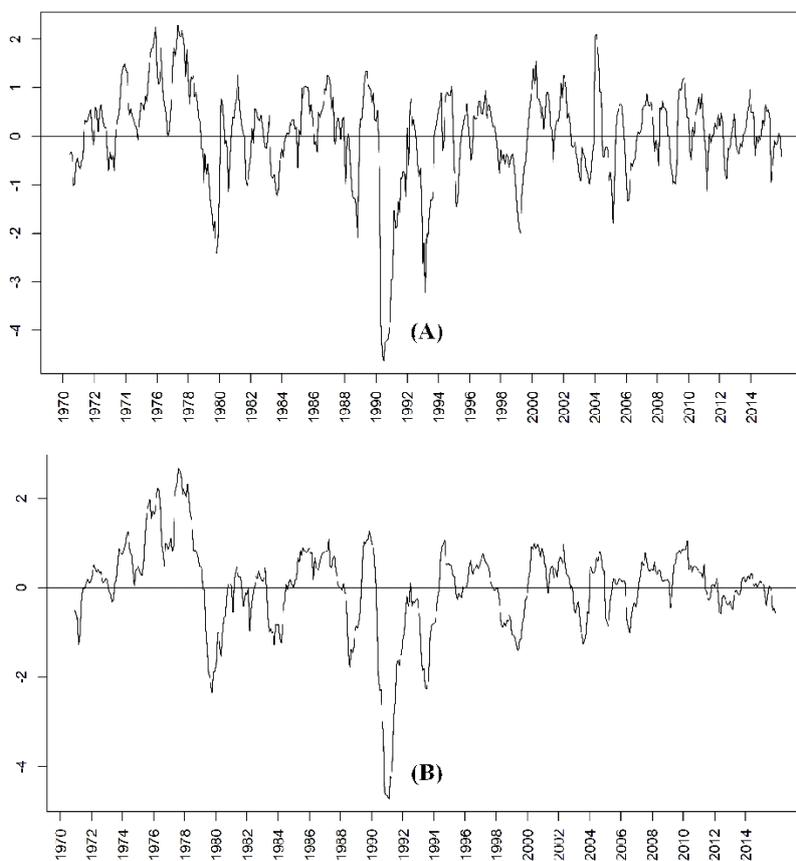


Figure 5. SPI-06 (A) and SPI-12 (B) for Propriá station.

In Figure 2 (A and B), it is possible to identify the drought years, due to the low values of SPI: the extreme drought in 1971 (with large negative peaks of SPI, -3.95 and -4.17 for the time scales of six and twelve months, respectively), and severe and moderate drought (SPI values between -1.0 and -2.0) in some years. It is important to emphasize that this station belongs to the Upper São Francisco region, where the occurrence of extreme and severe droughts is not expected.

In Figure 3 (A and B), for the Pilão Arcado station, four peaks of extreme droughts are identified for the years 2003, 2007, 2012 and 2013 with SPI values between -2.10 and -4.16 for the scale of period of six months (SPI-06). In the twelve-month scale (SPI-12), there are three extreme drought peaks for the years 2012, 2013 and 2014. The presence of severe dry and moderate drought can be observed in some years, with SPI values between -1.0 and -2.0. This station belongs to the region of the Middle San Francisco, where the occurrence of extreme and severe droughts is not expected, and the precipitation values are not so high.

In Figure 4 (A and B), the Fazenda Tapera station, there are several peaks of extreme drought. For the six-month scale (SPI-06) there are eight records of extreme dry spikes with values between -2.04 and -3.42 between 1991 and 2013; and also eight peaks for the twelve-month scale (SPI-12) between the years 1983 and 2013. In addition to extreme drought, several years presented severe drought and moderate drought with SPI values between -1.0 and -2.0, and mild drought with SPI ranging from 0.0 to -1.0. Since this station belongs to the sub-middle region, with a predominance of semi-arid and arid climate, extreme droughts, severe droughts and even mild droughts are expected due to the low precipitation values in this subregion of the basin.

In Figure 5 (A and B), three extreme drought peaks were identified, in 1980, 1991 and between 1993 and 1994 for the two time scales (SPI-06 and SPI-12). There are also some years with severe and moderate droughts in the categories of SPI-06 and SPI-12 with SPI values between -1.0 and -2.0, in addition to several years presenting mild drought with SPI between 0.0 and -1.0. It is worth noting that this station belongs to the Lower São Francisco subregion, whose expected occurrence of extreme and severe droughts is lower than in the Sub-middle, although Propriá is also part of the Northeastern Semi-arid (Brasil, 2005).

The transition probabilities for each of drought classes were calculated, representing the probabilities of state change from month to month, and presented in Table 4 (SPI-06) and Table 5 (SPI-12).

It is seen in Table 4 that for six-month, scale (SPI-06), the transition probabilities of the rainy state in the current month for the same rainy state in the subsequent month are greater than 0.800 (80%). It is also noted that the transition probabilities of drought conditions to rainy conditions are small, varying in the range of 0.00 (0.0%) to 0.262 (26.2%).

Another important fact that the matrix exposes is that if a locality is under the condition of extreme drought, the probability that in the following month the condition changes to rainy is zero. This is expected, because weather conditions do not change suddenly. In addition, the probability for an extreme drought to continue in the following month increases as it moves towards the river mouth, with values ranging from 0.500 (50%) to 0.789 (78.9%).

For the twelve-month scale (SPI-12), shown in Table 5, the probability that a locality is currently under a rainy state and in the next month remains in this same condition, is greater than 0.900 (90%). Likewise, the probability that a locality is under extreme drought conditions and the next month remains in this same condition, are higher than 0.778 (78.8%).

Table 4. Transition probability matrices for SPI-06 for basin stations.

Iguatama station Code: 2045002					
	Extreme drought	Severe drought	Moderate drought	Mild drought	Rainy
Extreme drought	0.500	0.286	0.000	0.143	0.071
Severe drought	0.176	0.118	0.412	0.294	0.000
Moderate drought	0.038	0.135	0.327	0.404	0.096
Mild drought	0.011	0.017	0.108	0.602	0.261
Rainy	0.000	0.003	0.031	0.146	0.819
Pilão Arcado station Code: 1042015					
	Extreme drought	Severe drought	Moderate drought	Mild drought	Rainy
Extreme drought	0.545	0.091	0.273	0.000	0.091
Severe drought	0.167	0.278	0.444	0.111	0.000
Moderate drought	0.057	0.151	0.340	0.377	0.075
Mild drought	0.000	0.018	0.114	0.611	0.257
Rainy	0.000	0.003	0.017	0.141	0.838
Fazenda Tapera station Code: 839034					
	Extreme drought	Severe drought	Moderate drought	Mild drought	Rainy
Extreme drought	0.556	0.222	0.167	0.000	0.056
Severe drought	0.095	0.238	0.381	0.238	0.048
Moderate drought	0.104	0.125	0.354	0.292	0.125
Mild drought	0.000	0.024	0.109	0.618	0.248
Rainy	0.003	0.007	0.010	0.146	0.833
Propriá station Code: 1036048					
	Extreme drought	Severe drought	Moderate drought	Mild drought	Rainy
Extreme drought	0.789	0.105	0.053	0.053	0.000
Severe drought	0.154	0.462	0.308	0.077	0.000
Moderate drought	0.037	0.148	0.296	0.370	0.148
Mild drought	0.006	0.006	0.081	0.705	0.202
Rainy	0.000	0.000	0.000	0.124	0.876

Table 5. Transition probability matrices for SPI-12 for basin stations.

Iguatama station Code: 2045002					
	Extreme drought	Severe drought	Moderate drought	Mild drought	Rainy
Extreme drought	0.824	0.000	0.059	0.118	0.000
Severe drought	0.063	0.563	0.250	0.125	0.000
Moderate drought	0.031	0.156	0.375	0.438	0.000
Mild drought	0.000	0.010	0.072	0.778	0.139
Rainy	0.000	0.000	0.004	0.089	0.907
Pilão Arcado station Code: 1042015					
	Extreme drought	Severe drought	Moderate drought	Mild drought	Rainy
Extreme drought	0.857	0.095	0.048	0.000	0.000
Severe drought	0.222	0.333	0.222	0.222	0.000
Moderate drought	0.022	0.067	0.556	0.333	0.022
Mild drought	0.005	0.005	0.077	0.798	0.115
Rainy	0.000	0.000	0.007	0.071	0.922
Fazenda Tapera station Code: 839034					
	Extreme drought	Severe drought	Moderate drought	Mild drought	Rainy
Extreme drought	0.778	0.056	0.111	0.056	0.000
Severe drought	0.167	0.583	0.167	0.083	0.000
Moderate drought	0.000	0.269	0.577	0.115	0.038
Mild drought	0.000	0.011	0.034	0.835	0.119
Rainy	0.000	0.000	0.000	0.078	0.922
Propriá station Code: 1036048					
	Extreme drought	Severe drought	Moderate drought	Mild drought	Rainy
Extreme drought	0.842	0.105	0.053	0.000	0.000
Severe drought	0.214	0.429	0.357	0.000	0.000
Moderate drought	0.000	0.152	0.485	0.364	0.000
Mild drought	0.000	0.006	0.071	0.753	0.169
Rainy	0.000	0.000	0.000	0.081	0.919

Tables 6 and 7 show stationary distribution for the six- and twelve-month scales. From Table 6, it can be seen that for all stations the probability for the rainy condition is higher than 0.526 (52.6%), and that the probabilities of drought conditions decrease with the degree of severity. Particularly for the extreme drought category, the probabilities are greater than 0.025 (2.5%) and increase as it moves towards the river mouth. The same occurs for the twelve-month scale whose probabilities for rainy condition are greater than 0.513 (51.3%), while for the extreme drought category the odds are greater than 0.021 (2.1%).

Table 6. Stationary distributions for SPI-06 for basin stations.

Station	Extreme drought	Severe drought	Moderate drought	Mild drought	Rainy
Iguatama	0.026	0.095	0.031	0.322	0.526
Pilão Arcado	0.025	0.099	0.034	0.302	0.541
Fazenda Tapera	0.034	0.091	0.039	0.299	0.537
Propriá	0.035	0.049	0.024	0.317	0.575

Table 7. Stationary distributions for SPI-12 for basin stations.

Station	Extreme drought	Severe drought	Moderate drought	Mild drought	Rainy
Iguatama	0.021	0.058	0.029	0.356	0.536
Pilão Arcado	0.054	0.081	0.019	0.333	0.513
Fazenda Tapera	0.038	0.056	0.05	0.327	0.53
Propriá	0.035	0.061	0.026	0.285	0.593

The recurrence time for each drought class is shown in Tables 8 and 9 for the six- and twelve-month scale, respectively.

Table 8. Recurrence time for SPI-06 for basin stations.

Station	Extreme drought	Severe drought	Moderate drought	Mild drought	Rainy
Iguatama	38.5	10.5	32.3	3.1	1.9
Pilão Arcado	40.0	10.1	29.4	3.3	1.8
Fazenda Tapera	29.4	11.0	25.6	3.3	1.9
Propriá	28.6	20.4	41.7	3.2	1.7

Table 9. Recurrence time for SPI-12 for basin stations.

Station	Extreme drought	Severe drought	Moderate drought	Mild drought	Rainy
Iguatama	47.6	17.2	34.5	2.8	1.9
Pilão Arcado	18.5	12.3	52.6	3.0	1.9
Fazenda Tapera	26.3	17.9	20.0	3.1	1.9
Propriá	28.6	16.4	38.5	3.5	1.7

At the six-month scale, the lowest recurrence times found for the extreme drought class were for the Fazenda Tapera and Propriá stations, with values close to 29.4 and 28.6, respectively. These stations belong to Sub-middle and Lower Sao Francisco, respectively where severe droughts occur frequently as a result of low rainfall and high evapotranspiration (Bezerra *et al.*, 2018).

For the twelve-month scale, the highest return time for the extreme drought class was found to be 47.6 months for the Iguatama station, meaning that among the other stations this is the one with the least possibility of being under this type of climate. It is important to point out that this station belongs to the Upper São Francisco subregion, with a predominance of tropical humid and temperate climate, reflecting a greater amount of precipitation, that is, the chances of drought must be lower than in other subregions.

Still in the twelve-month time scale (SPI-12), the lowest recurrence times for the extreme drought class were for the Pilão Arcado, Fazenda Tapera and Propriá stations. This fact indicates that these stations are more susceptible to drought, as expected, since they are located in the semi-arid region, according to the last update of the municipalities of that region made in 2005 by the Ministry of National Integration (Brasil, 2005).

The expected residence time (which is the average time the process stays in a particular drought class before migrating to another class and represents the duration of that drought class) is shown in Tables 10 and 11, for the six- and twelve-month scales, respectively.

Table 10. Expected residence times for SPI-06 for basin stations.

Station	Extreme drought	Severe drought	Moderate drought	Mild drought	Rainy
Iguatama	2.0	1.5	1.1	2.5	5.5
Pilão Arcado	2.2	1.5	1.4	2.6	6.2
Fazenda Tapera	2.3	1.6	1.3	2.6	6.0
Propriá	4.8	1.4	1.9	3.4	8.1

Table 11. Expected residence time for SPI-12 for basin stations.

Station	Extreme drought	Severe drought	Moderate drought	Mild drought	Rainy
Iguatama	5.7	1.6	2.3	4.5	10.8
Pilão Arcado	7.0	2.3	1.5	5.0	12.8
Fazenda Tapera	4.5	2.4	2.4	6.1	12.9
Propriá	6.3	1.9	1.8	4.1	12.3

It can be observed in Tables 10 and 11 that the longest residence time for both six- and twelve-month scales, is obtained for the rainy condition, greater than and equal to 5.5 months. This information can be used, for example, to plan the sowing of a given crop, since the beginning of the rainy season is known. Also, in water management projects, such as the planning of water drainage systems or irrigation projects, it is extremely important to know the probability of rainy conditions above or below a certain value (Pereira *et al.*, 2007).

4. CONCLUSIONS

The modeling of the SPI for six- and twelve-month scales through Markov chains proved useful in the search for a better understanding of the stochastic characteristics of the climatic behavior of the subregions of the São Francisco Basin, providing (through the probability analysis for each drought class) recurrence time and expected value of residence time in each class. It was possible to observe that the results for the two scales are similar, but each one has its specific purpose. The six-month scale is recommended for the observation of agricultural drought, while the twelve-month scale is better for hydrological drought assessment (WMO, 2012).

For both scales, the transition matrix showed that switching from severe drought or rainy conditions to another class is unlikely in the short term. Correlating this information with the probabilities of the stationary distribution, it is possible to find the regions that are most likely to be in the future under rainy weather or under drought conditions. In all cases studied, and for both scales, the probability of extreme drought was low even for the semi-arid region, represented by Fazenda Tapera station. The lowest recurrence times for the extreme drought class was found for Fazenda Tapera and Propriá stations that belong to the semi-arid region.

The evaluation of the expected value of the time of residence rectifies the information of the period of duration of each climatic condition. The shortest residence time for extreme drought and the longest residence time for rainy condition was again obtained for the Fazenda Tapera station, located in semi-arid region. This information can be used for planting a particular crop, knowing the beginning of the rainy season or for the planning of the use of water resources for the generation of energy or for reservoir- and dam construction.

5. ACKNOWLEDGEMENTS

The authors thank the Coordination for the Improvement of Higher Education Personnel (CAPES) for her PhD scholarship and the Interinstitutional Doctorate (DINTER) between Federal University of Sergipe (UFS) and Federal Rural University of Pernambuco (UFRPE).

6. REFERENCES

- BEZERRA, B. G. *et al.* Changes of precipitation extreme indices in São Francisco River Basin, Brazil from 1947 to 2012. **Theoretical and Applied Climatology**, p. 1-12, 2018. <https://doi.org/10.1007/s00704-018-2396-6>
- BLAIN, G. C. Revisiting the probabilistic definition of drought: strengths, limitations and an agrometeorological adaptation. **Bragantia**, v. 71, p. 132-141, 2012. <http://dx.doi.org/10.1590/S0006-87052012000100019>
- BLAIN, G. C. Revisiting the critical values of the Lilliefors test: towards the correct agrometeorological use of the Kolmogorov-Smirnov framework. **Bragantia**, v. 73, p. 192-202, 2014. <http://dx.doi.org/10.1590/brag.2014.015>
- BLAIN, G. C. *et al.* Índice padronizado de precipitação aplicado às condições de seca no Estado do Espírito Santo. **Revista Brasileira de Engenharia Agrícola e Ambiental**, v. 14, n. 10, p. 1067-1073, 2010. <https://dx.doi.org/10.1002/joc.1441>
- BRASIL. Ministério da Integração Nacional. **Rio e Seus Números**. Available at: <http://www.integracao.gov.br/web/projeto-sao-francisco/o-rio-e-seus-numeros> Access in: July 2018.
- BRASIL. Ministério da Integração Nacional. Secretaria de Políticas de Desenvolvimento Regional. **Nova Delimitação do Semi-Árido Brasileiro**. Brasília, 2005.
- HUSAK, G. J.; MICHAELSEN, J.; FUNK, C. Use of the gamma distribution to represent monthly rainfall in Africa for drought monitoring applications. **International Journal of Climatology**, v. 27, p. 935-944, 2007.
- KELLER FILHO, T. *et al.* Análise da Transição entre Dias Secos e Chuvosos Usando Cadeias de Markov de Terceira Ordem. **Pesquisa Agropecuária Brasileira**, v. 41, n. 9, p. 1341-1349, 2006.
- LENNARTSSON, L. *et al.* Modelling precipitation in Sweden using multiple step Markov chains and a composite model. **Journal of Hydrology**, v. 363, p. 42-59, 2008. <https://doi.org/10.1016/j.jhydrol.2008.10.003>
- MAHMOUDZADEH, H. *et al.* Applying First-Order Markov Chains and SPI Drought Index to Monitor and Forecast Drought in West Azerbaijan Province of Iran. **International Journal of Geosciences and Environment Planning**, v. 1, n. 2, p. 44-53, 2016.
- MCKEE, T. B. *et al.* The Relationship of Drought Frequency and Duration Times Scales. In: CONFERENCE ON APPLIED CLIMATOLOGY, 8., 17-22 Janvier, Anaheim. **Proceedings[...]** Massachusetts: AMS, 1993. p. 179-184.
- MISHRA, A. *et al.* Drought characterization: a probabilistic approach. **Stochastic Environmental Research and Risk Assessment**, v. 23, p. 41-55, 2009. <https://doi.org/10.1007/s00477-007-0194-2>
- MISHRA, A.K.; SINGH, V.P. A review of drought concepts. **Journal of Hydrology**, v. 391, n. 1-2, p. 202-216, 2010. <https://doi.org/10.1016/j.jhydrol.2010.07.012>
- PEREIRA, A. R. *et al.* **Meteorologia agrícola**. São Paulo: USP, 2007.
- PAULO, A. A.; PEREIRA, L. S. Prediction of SPI Drought Class Transitions Using Markov Chains. **Water Resources Management**, v. 21, n. 10, p. 1813-1827, 2007. <https://doi.org/10.1007/s11269-006-9129-9>

- SADEGHI, L.; SHAMSELDIN, A. Y. Application of the Standardized Precipitation Index (SPI) in Hawke's Bay, New Zealand. *In*: ANDREU, J. (Ed.). **Drought: Research and Science-Policy Interfacing**, 2014. p. 139.
- R CORE TEAM. **R**: A language and environment for statistical computing. Vienna, 2019. Available at: <https://www.R-project.org/>
- SANUSI, W. *et al.* The drought characteristics using the first-order homogeneous Markov chain of monthly rainfall data in peninsular Malaysia. **Water Resources Management**, v. 29, n. 5, p. 1523-1539, 2015. <https://doi.org/10.1007/s11269-014-0892-8>
- SNIRH. **Informações sobre recursos hídricos.** Available at: <http://www.snirh.gov.br/portal/snirh> Access in: July 2018.
- SOUSA, F. A. S. *et al.* O Índice de Precipitação Padronizada (IPP) na identificação de extremos de chuvas e secas na bacia do rio Paraguaçu (BA). **Ambiência Guarapuava (PR)**, v. 12, n. 2, p. 707-719, 2016.
- TONKAZ, T. Spatio-temporal assessment of historical droughts using SPI with GIS in GAP Region, Turkey. **Journal of Applied Sciences**, v. 12, n. 6, p. 2565-2571, 2006.
- TSAKIRIS, G.; VANGELIS, H. Towards a Drought Watch System based on Spatial SPI. **Water Resources Management**, v. 18, p. 1-12, 2004. <https://doi.org/10.1023/B:WARM.0000015410.47014.a4>
- UFSC. Centro Universitário de Estudos e Pesquisas sobre Desastres. **Atlas brasileiro de desastres naturais: 1991 a 2012**. 2nd ed. Volume Brasil. Florianópolis: CEPED UFSC, 2013.
- WILKS, D. S. **Statistical Methods in the Atmospheric Sciences**. 3rd Ed. Oxford: Elsevier, 2011.
- WMO. **Standardized precipitation index user guide**. Geneva, 2012. Available at: http://www.wamis.org/agm/pubs/SPI/WMO_1090_EN.pdf Access in: July 2018.