Revista Brasileira de Meteorologia, **v. 39,** e39240050, 2024
DOI: http://dx.doi.org/10.1590/0102-77863910050
rbmet.org.br

Article

Mapping Rainfall Variability in the São Francisco River Basin: Insights for Water Resource Management

Elias Silva de Medeiros[1], Willian dos Santos Oliveira[2], Paulo Vinicius da Silva[3], Amaury de Souza[4], Carolina Cristina Bicalho[5], Munir Mauad[3], Sidnei Azevedo de Souza[1]

[1]Faculdade de Ciências Exatas e Tecnologia, Universidade Federal da Grande Dourados, Dourados, MS, Brazil.
[2]Faculdade de Engenharia, Universidade Federal da Grande Dourados, Dourados, MS, Brazil.
[3]Faculdade de Ciências Agrárias, Universidade Federal da Grande Dourados, Dourados, MS, Brazil.
[4]Departamento de Física, Universidade Federal de Mato Grosso do Sul, Campo Grande, MS, Brazil.
[5]Departamento de Matemática, Universidade Estadual de Mato Grosso do Sul, Campo Grande, MS, Brazil.

Received: 30 October 2023 - Accepted: 21 February 2024

Abstract

The São Francisco River Basin (SFRB) is extremely important for both the local rural and urban population, running through several states and municipalities in Brazil. Thus, the study of the spatial distribution of rainfall in the SFRB, supported by the application of geostatistics, offers valuable insights for water management and planning. In this study, data from 341 rainfall stations were used, corresponding to the period between 1989 and 2018, and monthly accumulated rainfall analyzed. About the trends, positive estimates for latitude, except for April, highlight that the accumulated precipitation decreases from south to north each month. In December, 85.10% of the spatial variability of rainfall was explained by the trend of rainfall concerning geographical coordinates. Concerning spatial dependence, the analyzed data predominantly exhibited the exponential variogram model. The conducted analyses offered a holistic view of rainfall in the region, facilitating the understanding of spatial patterns and shedding light on challenges in the face of climate change. Our research limited on the spatial modeling of rainfall in the SFRB, and future work should include modeling socio-economic and environmental impacts in the region concerning the spatial irregularity of rainfall.

Keywords: climate change, geostatistics, rainfall, universal kriging.

Mapeando a Variabilidade das Chuvas na Bacia do Rio São Francisco: Insights para a Gestão de Recursos Hídricos

Resumo

A Bacia Hidrográfica do Rio São Francisco (SFRB) é de extrema importância tanto para a população rural quanto urbana local, percorrendo diversos estados e municípios do Brasil. O estudo da distribuição espacial das chuvas na SFRB, apoiado na aplicação da geoestatística, oferece insights valiosos para a gestão e planejamento hídrico. Nesse estudo foram utilizados dados de 341 estações pluviométricas, correspondentes ao período entre 1989 e 2018, e analisadas as chuvas acumuladas mensais. Sobre a tendência, estimativas positivas para a latitude, com exceção de abril, ressaltam que a precipitação acumulada em cada mês diminui de sul para norte. No mês de dezembro, 85,10% da variabilidade espacial das chuvas foi explicada pela tendência das chuvas em relação às coordenadas geográficas. Na dependência espacial, os dados analisados mostram predomínio do modelo de variograma exponencial. As análises realizadas

Autor de correspondência: Elias Silva de Medeiros, eliasmedeiros@ufgd.edu.br.

proporcionaram uma visão holística das chuvas na região, permitindo a compreensão dos padrões espaciais e lançando luz sobre os desafios apresentados em um cenário de mudanças climáticas. Essa pesquisa limitou-se à modelagem espacial das chuvas na SFRB, e uma modelagem dos impactos socioeconômicos e ambientais na região em relação à irregularidade espacial das chuvas é essencial para trabalhos futuros.

Palavras-chave: mudanças climáticas, geoestatística, chuvas, krigagem universal.

1. Introduction

The São Francisco River Basin (SFRB) holds tremendous importance for both the local rural and urban populations, spanning several states and municipalities in Brazil. The SFRB is utilized for human consumption, power generation, irrigation, and livestock (Santos *et al.*, 2020). The basin contributes to the development of the states it traverses, and understanding its hydrological variability is of utmost importance, as it allows us to provide information on the availability of water resources, thereby helping to identify areas that may face scarcity of these resources (Pereira *et al.*, 2007).

Analyzing the spatial distribution of rainfall reveals substantial climate variability, resulting in scenarios of insufficient or excessive rainfall. These complex climatic dynamics not only directly influence the region's hydrological regime, but also trigger wide-ranging consequences, such as prolonged droughts or devastating floods. The occurrence of extreme weather events plays an incredibly relevant role in the daily lives of communities, noteworthy for both the severity of the weather conditions and the increasing recurrence of these events. This interaction between extreme weather phenomena and the socio-environmental vulnerability of the region underscores the pressing need for a deeper understanding of these weather patterns and adaptation strategies that can mitigate impacts on populations and the surrounding environment (Magerski and Virgens Filho, 2021).

The semi-arid regions of Northeast Brazil (NEB) are characterized by climate variability, mainly in relation to the spatiotemporal distribution of rainfall (Lyra *et al.*, 2014; Lyra *et al.*, 2017; Ruffato-Ferreira *et al.*, 2017; Santos *et al.*, 2020; Lucas *et al.*, 2020; Rocha *et al.*, 2022). The occurrence of such variability and extreme drought events characterizes the level of vulnerability faced by the population, generating uncertainties related to potential climate changes (Sobral *et al.*, 2018). Obtaining highly detailed maps of rainfall in a region plays a crucial role in hydrological planning. By enabling more efficient management of water resources, these maps contribute significantly to the mitigation of risks associated with natural disasters, such as floods, erosion, and droughts (Parker *et al.*, 2019).

The rainfall variability in the SFRB is influenced by various factors and processes. The South Atlantic Convergence Zone (SACZ) stands out as the primary meteorological system responsible for rainfall in the SFRB (Silva and Brito, 2008), with its intense activity impacting the reservoir levels of hydroelectric power plants in the basin, leading to fluctuations in energy prices (Vilar *et al.*, 2020). Additionally, through a Wavelet analysis, it was observed that phenomena such as the El Niño Southern Oscillation (ENSO), extended ENSO, and the Pacific Decadal Oscillation (PDO) directly influence the occurrence of extreme climatic events in the basin, causing periods of both drought and rainfall (Rocha *et al.*, 2022).

The spatial distribution of rainfall in the SFRB is a topic of great relevance for hydrological planning and the management of water resources. It is known that rainfall levels are available only in places where rainfall stations are installed. To estimate values in unsampled locations and obtain interpolated high-resolution maps for the entire study region, geostatistics can be used (Medeiros *et al.*, 2019; Barros *et al.*, 2020; Ferreira *et al.*, 2021; Oliveira *et al.*, 2022).

Given this context, the significant importance of the present study on the spatial distribution of rainfall in the SFRB is highly emphasized due to the vast territorial expanse of Brazil, where many regions of interest lack information from pluviometers. Particularly notable in developing countries, the limited infrastructure, density, and frequency of measurements can result in a monitoring deficit. In order to overcome this gap, hydrological regionalization, which encompasses interpolation maps, emerges as a vital solution for obtaining information on watercourse flows in areas lacking measured data (Charles *et al.*, 2022). Thus, the study of the spatial distribution of rainfall in the SFRB, supported by the application of geostatistics, offers valuable insights for water management and planning, enabling the precise identification of vulnerable areas and significantly contributing to the prevention of adverse impacts.

The present study has the following main objectives: i) to investigate the spatial distribution of rainfall in the SFRB through the application of geostatistics, aiming to obtain high-resolution maps of rainfall indices in non-sampled locations. ii) analyze the hydrological variability of the region and thus understand the temporal and spatial variations of the available water resources, which are fundamental for the planning and sustainable management of water resources. With this, it is expected that this study can make significant contributions to the understanding of rainfall distribution in the SFRB, as well as aid in decision-making and in the development of strategies to deal with the challenges posed by climate variability and the management of water resources within one of the most important regions of Brazil.

In this study, we address a critical research gap pertaining to the spatial distribution of rainfall in the SFRB. Despite the region's significant socio-economic and environmental importance, there exists a notable scarcity of detailed information on rainfall patterns, especially in areas lacking pluviometer coverage. The limited availability of such crucial data poses challenges for effective hydrological planning and water resource management.

The primary objectives of this study are: i) to investigate the spatial distribution of rainfall in the SFRB through geostatistics, aiming to generate high-resolution rainfall index maps for non-sampled locations. ii) analyze the hydrological variability of the region, understanding temporal and spatial variations of available water resources crucial for planning and sustainable management. It is anticipated that this study will make substantial contributions to understanding rainfall distribution in the SFRB, informing decision-making and strategies for addressing challenges posed by climate variability and water resource management in one of Brazil's most crucial regions.

2. Materials and Methods

The SFRB is located between geographic coordinates 7°17’ S and 20°50’ S latitude, and between 36°15’ W and 47°39’ W longitude. It encompasses a territorial extension of 639,219 km^2, which corresponds to 8% of the national territory. This area covers 503 municipalities, part of the Federal District and also six other federal units in Brazil: Minas Gerais, Goiás, Bahia, Pernambuco, Sergipe and Alagoas. The Hydrographic Region under analysis is divided into four distinct physiographic regions, known as São Francisco Upper, São Francisco Middle, São Francisco Lower-middle and São Francisco Lower (Fig. 1).

The basin is divided into four distinct regions. In the first region, called Alto São Francisco, which extends from the source in Serra da Canastra, in Minas Gerais, to the confluence with the Jequitaí River, also in Minas Gerais, we find a predominantly humid climate, interspersed with stretches of sub-humid and dry sub-humid climates, playing a crucial role in the generation of flows. The second region, Middle São Francisco, covers 1,152 km in length, extending from the confluence of the Jequitaí River to the dam of the Sobradinho hydroelectric plant, on the border between Pernambuco and Bahia. Here, areas characterized by a dry sub-humid and semi-arid climate impact the flow, influenced by the regulation of the Três Marias reservoir. The third region, Submédio São Francisco, with 568 km of length, covers from the Sobradinho dam to Belo Monte, Alagoas, and is characterized by prevailing an arid to semi-arid climate. Flow control is carried out by the reservoirs of the Sobradinho and Itaparica plants, complemented by a cascade of hydroelectric plants due to the slope of the original river bed. Finally, the fourth region, Baixo São Francisco, spans from the downstream stretch of Belo Monte to its mouth in the Atlantic Ocean, on the border between Sergipe and Alagoas, and is

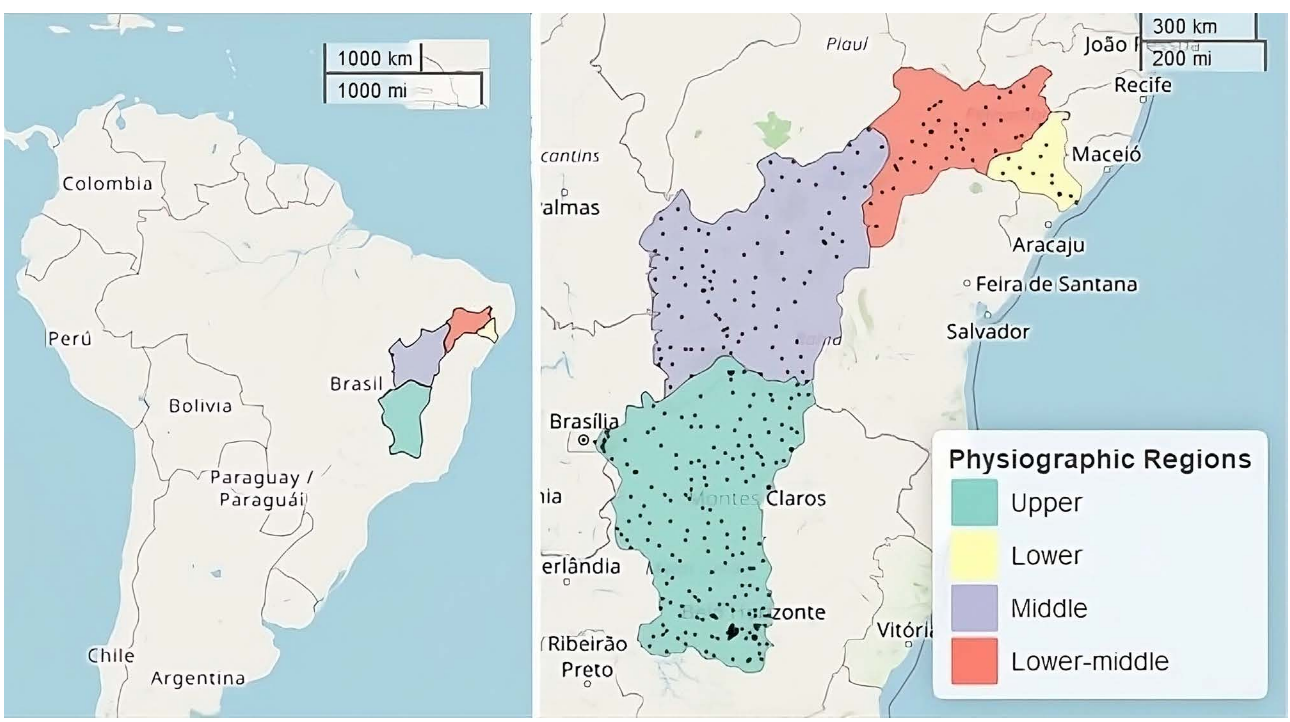

Figure 1 - Map of South America highlighting the São Francisco River Basin (left side) and the map of this basin highlighting the physiographic regions and the location of the 341 pluviometric stations (right side).

characterized by the river conformation and flow regulation through the reservoirs of the Sobradinho and Itaparica (CBHSF, 2016).

For statistical analysis, the dataset used consisted of 341 pluviometric stations located in the SFRB with monthly records of accumulated rainfall from 1989 to 2018. The dataset was acquired from the HydroWeb Portal, which is part of the National Water Resources Information System (SNIRH) and coordinated by the National Water Agency (ANA). Initially, the dataset was available with information on daily rainfall totals (24-hour rainfall) for each of the 341 stations. Subsequently, for stations with missing values, proportional interpolation was conducted based on the total number of days with information in the respective month under analysis. The data were aggregated by month, and the average total rainfall was calculated based on the 30-year record of the historical series. Thus, the spatial statistics analyses were conducted individually for each of the twelve months.

For the geostatistical analysis of average monthly rainfall, it was considered that the function $Z(s)$ was decomposed by the trend component $Y(s)$ and by the stochastic residue $\varepsilon(s)$.

$$Z(s) = Y(s) + \varepsilon(s). \tag{1}$$

In Eq. (1), the mean monthly rainfall is represented by $Z(s)$. For modeling the variogram in the residuals, it is crucial to remove the trend from the attribute under investigation, allowing a more faithful representation of the remaining spatial variability structure. One way to achieve this removal is through the use of multiple regression, considering the geographical coordinates of the sampled points as explanatory variables (Oliveira *et al.*, 2022). In this study, the trend component was adjusted using the multiple regression model described by:

$$Y_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \varepsilon_i. \tag{2}$$

For the trend, geographic coordinates were inserted as explanatory variables, considering the linear effect of latitude (X_1) and the linear (X_2) and quadratic (X_3) effects of longitude (Eq. 2). In this equation, the terms $\beta_0, \beta_1, \beta_2$and β_3 represent the parameters that will be estimated by the method of least squares. After properly eliminating trends by Eq. (2), a variographic analysis of the residues was carried out using the following empirical variogram:

$$\hat{\gamma}(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} \left[\hat{\varepsilon}_i(s_i + h) - \hat{\varepsilon}_i(s_i)\right]^2. \tag{3}$$

The terms $\hat{\gamma}(h)$ represent the value of the empirical variogram and $N(h)$ is the number of pairs of separated points for a distance h. The expressions $\hat{\varepsilon}_i(s_i + h) \text{and} \hat{\varepsilon}_i(s_i)$ refer to the residual values at locations $s_i + h \text{and} s_i$, respectively (Eq. (3)). In the context of the empirical variogram pseudo-data, the theoretical models of Spherical, Exponential and Gaussian variogram were proposed (Table 1). These models have been widely used in studies for modeling the spatial distribution of rainfall (Medeiros *et al.*, 2019; Comisso and Medeiros, 2021; Oliveira *et al.*, 2022), as they demonstrate a high capability to represent structures of spatial variability, thus enabling a comprehensive analysis of patterns present in the data. In Table 1, the parameter r represents the spatial range and the terms C_0 and C_1 represent the nugget and partial sill parameters, respectively.

To obtain the parameters of each model (Table 1), the weighted least squares method was used. In this process, weighting was determined based on the number of points $N(h)$ located at a distance h, following the ratio $N(h)/h^2$ (Comisso and Medeiros, 2021).

The spatial dependence index (SDI) presented in this research consists of the ratio between the sill and nugget parameter estimates, and can be classified between SDI ≤ 25% - strong spatial dependence; 25% < SDI < 75% - moderate spatial dependence and; SDI ≥ 75% - weak spatial dependence (Cambardella *et al.*, 1994).

To perform the interpolation of rainfall, the Universal Kriging (KU) method was adopted. In this approach, the trend component can be treated as a multiple regression, in which the geographic coordinates themselves act as regression variables (Bhattacharjee *et al.*, 2019). The choice of KU was based on its suitability as the most appropriate geostatistical interpolation method in modeling non-stationary phenomena, which exhibit spatial trends (Santos *et al.*, 2011). The interpolated precipitation value ($\hat{Z}$) for a specific point (s_0) is derived by the following expression:

$$\hat{Z}(s_0) = \sum_{i=1}^{N} \lambda_i f_k(s_i). \tag{4}$$

Where N represents the number of sampled rainfall stations ($i = 1, 2, \ldots, 341$), $f_k(s_i)$ refers to predefined functions that incorporate geographic coordinates (Eq. (4)). The term k denotes the number of parameters present in the

Table 1 - Variogram models fitted to pseudo-data from the empirical variogram followed by the spatial dependence index (SDI). The terms C_0 represents the nugget effect, $C_0 + C_1$ the sill and r the spatial range.

Model	Equation	SDI (%)
Spherical	$\gamma(h) = C_0 + C_1\left[\frac{3h}{2r} - \frac{1}{2}\left(\frac{h}{r}\right)^3\right], 0 < h \le r$ $\gamma(h) = C_0 + C_1, h > r$	$\frac{C_0}{C_0 + C_1} \times 100\%$
Exponential	$\gamma(h) = C_0 + C_1\left[1 - exp\left(-\frac{h}{r}\right)\right], h > 0$	
Gaussian	$\gamma(h) = C_0 + C_1\left[1 - exp\left(-\frac{h^2}{r^2}\right)\right], 0 < h \le r$ $\gamma(h) = C_0 + C, h > r$	

trend component. Obtaining the weights, denoted as λ_i, requires the solution of a system of linear equations represented by $A\lambda = b$ (Webster and Oliver, 2007). The matrix A and the vectors λ and b are constructed as follows:

$$A = \begin{bmatrix} \gamma(s_1,s_1) & \gamma(s_1,s_2) & \cdots & \gamma(s_1,x_N) & 1 & f_1(s_1) & f_2(s_1) & \cdots & f_K(s_1) \\ \gamma(s_2,s_1) & \gamma(s_2,s_2) & \cdots & \gamma(s_2,x_N) & 1 & f_1(s_2) & f_2(s_2) & \cdots & f_K(s_2) \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ \gamma(s_N,s_1) & \gamma(s_N,s_2) & \cdots & \gamma(s_N,x_N) & 1 & f_1(s_N) & f_2(s_N) & \cdots & f_K(s_N) \\ 1 & 1 & \cdots & 1 & 0 & 0 & 0 & \cdots & 0 \\ f_1(s_1) & f_1(s_2) & \cdots & f_1(s_N) & 0 & 0 & 0 & \cdots & 0 \\ f_2(s_1) & f_2(s_2) & \cdots & f_2(s_N) & 0 & 0 & 0 & \cdots & 0 \\ \vdots & 1 & \cdots & 1 & \vdots & \vdots & \vdots & \cdots & \vdots \\ f_K(s_1) & f_K(s_2) & \cdots & f_K(s_N) & 0 & 0 & 0 & \cdots & 0 \end{bmatrix}$$

$$\lambda = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_N \\ \psi_0 \\ \psi_1 \\ \psi_2 \\ \vdots \\ \psi_k \end{bmatrix} \text{ and } b = \begin{bmatrix} \gamma(s_1,s_0) \\ \gamma(s_2,s_0) \\ \vdots \\ \gamma(s_N,s_0) \\ 1 \\ f_1(s_0) \\ f_2(s_0) \\ \vdots \\ f_K(s_0) \end{bmatrix}.$$

In the system of linear equations, $\gamma(s_j, s_0)$ are the values of the variances obtained between the sampled point (s_j) and the interpolated one (s_0). The component $\gamma(s_i, s_j)$ indicates the variance of the residuals between points s_i and s_j. The function $f_k(s), k = 1, 2, \ldots, K$, represents the adjustment to the trend, where k is the number of parameters estimated in the adjustment. Furthermore, the Lagrange multiplier in Universal Kriging ψ_k is used to find a balance between the observed data and the overall pattern of the data, ensuring that the interpolation is consistent and effective in representing the spatial characteristics of the data. Additionally, the kriging variance can be determined by the $\sigma^2_{KU} = b^T\lambda$ system.

Through the *leave-one-out* cross-validation during the KU, the theoretical model of the variogram was selected, this validation takes place through the withdrawal of values observed in the geographic coordinates s_i $(i = 1, 2, \ldots, 341)$ and then these values are interpolated by kriging, based on the chosen variogram models (Gois *et al.*, 2015). Therefore, to choose the best variogram model for interpolation, the root-mean squared error (RMSE), mean absolute error (MAE) and coefficient of determination (R^2) statistics were used.

All statistical analyzes were performed in the R software (R Core Team, 2023) with support from the gstat (Gräler *et al.*, 2016) and ggplot2 (Wickham, 2016) libraries.

3. Results

Table 2 displays the descriptive statistics obtained from the 30-year time series (1989 to 2018), based on information from 341 rainfall stations within the SFRB. The month of December stands out, with the highest average (201.47 mm), and the lowest average occurred in August (7.83 mm) during the study period. Regarding the coefficient of variation (CV, %), there was greater variability than 30% in all periods, indicating high spatial variability, and the CV% in the months of June, July, and August was greater than 100%, once again indicating high spatial dispersion of rainfall during this period.

The Lower-middle São Francisco region is characterized by extremely irregular rainfall, with the main rainy

Table 2 - Descriptive analysis of the 341 rainfall stations located in the São Francisco River Basin with records from 1989 to 2018.

Month	Minimum	Mean	Median	Maximum	Standard deviation	CV (%)
January	19.49	174.46	169.16	451.72	81.89	46.94%
February	16.3	125.04	125.18	308.17	42.79	34.27%
March	25.91	139.7	142.23	383.26	46.69	33.42%
April	26.29	59.8	55.71	147.58	20.46	34.22%
May	1.44	23.79	18.82	206.58	23.49	9871,0%
June	0.00	11.05	5.06	199.63	21.67	195.90%
July	0.00	8.14	3.4	193.2	19.83	243.62%
August	0.00	7.83	4.59	112.97	12.08	154.25%
September	0.04	22.66	15.63	67.65	18.45	81.44%
October	3.82	65.08	69.11	157.7	31.69	48.69%
November	5.007	163.67	188.81	591.39	74.32	45.41%
December	13.27	201.47	216.88	610.09	93.48	46.40%

season occurring from January to April (Sobral *et al.*, 2018). In the study conducted by these authors, it was verified that the average annual rainfall varied between 300 and 1200 mm, with the highest values in the Alto Sertão of Pernambuco (> 600 mm), and the lowest in São Francisco de Pernambuco and Sertão in Bahia, with average totals ranging between 300 and 600 mm.

The monthly climate transitions in the SFRB are related to the seasons. In summer, which runs from December to March, average rainfall is higher. In autumn and winter, there is less rainfall, with values falling below 10 mm. In spring, a further increase in these average values can be noted.

In Fig. 2, a depiction of the rainfall relationship with respect to geographic coordinates is provided for each month of the year, based on data derived from the 341 pluviometric stations situated within the SFRB.

Through scatter plots (Fig. 2), the evident trend relationship of rainfall with geographical coordinates (latitude and longitude) becomes apparent. As a result, the adjustment of multiple regression with a linear effect for latitude and both linear and quadratic effects for longitude was proposed (Table 3).

The estimates from the multiple regression model were all significant ($p < 0.05$). The positive estimates for latitude, with the exception of April, underscore that accumulated rainfall in each month decreases from south to north. April exhibited the lowest R^2 value, indicating that 34.94% of the spatial variability of rainfall is accounted for by the trend adjustment. Conversely, for the month of December, 85.10% of the spatial variability of rainfall can be explained by the trend of rainfall in relation to geographical coordinates.

Table 4 presents the results of the “leave-one-out” cross-validation, based on the adjustments of variograms (Spherical, Exponential, and Gaussian), for the accumulated monthly rainfall data after trend removal. The Exponential variogram model yielded the most favorable outcomes in nine out of the twelve months of the year, followed by the Spherical model for May and October, and the Gaussian model solely for December.

Comisso and Medeiros (2021), when modeling rainfall data in the state of Alagoas, employed the adjustments of the Spherical, Exponential, and Gaussian models. The authors concluded that the Exponential model was the best fit for both dry and rainy periods in the region. Barros *et al.* (2020), meanwhile, assessed monthly pluviometric data from 150 pluviometric stations located in the state of Pernambuco. They found that the Exponential model yielded the lowest values for the AIC and BIC statistics when fitting the annual average pluviometric data in the study region. However, in the aforementioned study, the authors did not conduct cross-validation, a fundamental procedure for assessing the quality of interpolation in kriging.

Table 5 show the estimates of the adjusted theoretical variogram model for each month. In January, the highest sill estimate (1417.081) and a range of 73.086 km were observed, possibly attributed to January having one of the highest rainfall means (Table 2). In contrast, August, which had the lowest sill estimate (18.36) and a range of 42.57 km, exhibited the lowest average rainfall.

In Fig. 3, during the months of January to April (summer-autumn), there is a substantial distribution of rainfall across the regions from Upper to Lower São Francisco. In the months from May to August (autumn-winter), rainfall in the basin is scarce, particularly in the Upper, Middle, and southern parts of the Lower-Middle region, characterizing a dry period. From September to December (winter-spring), a reversal in rainfall is observed where a decrease begins in the Lower and northern part of the Lower-Middle region, and greater rainfall occurs from the

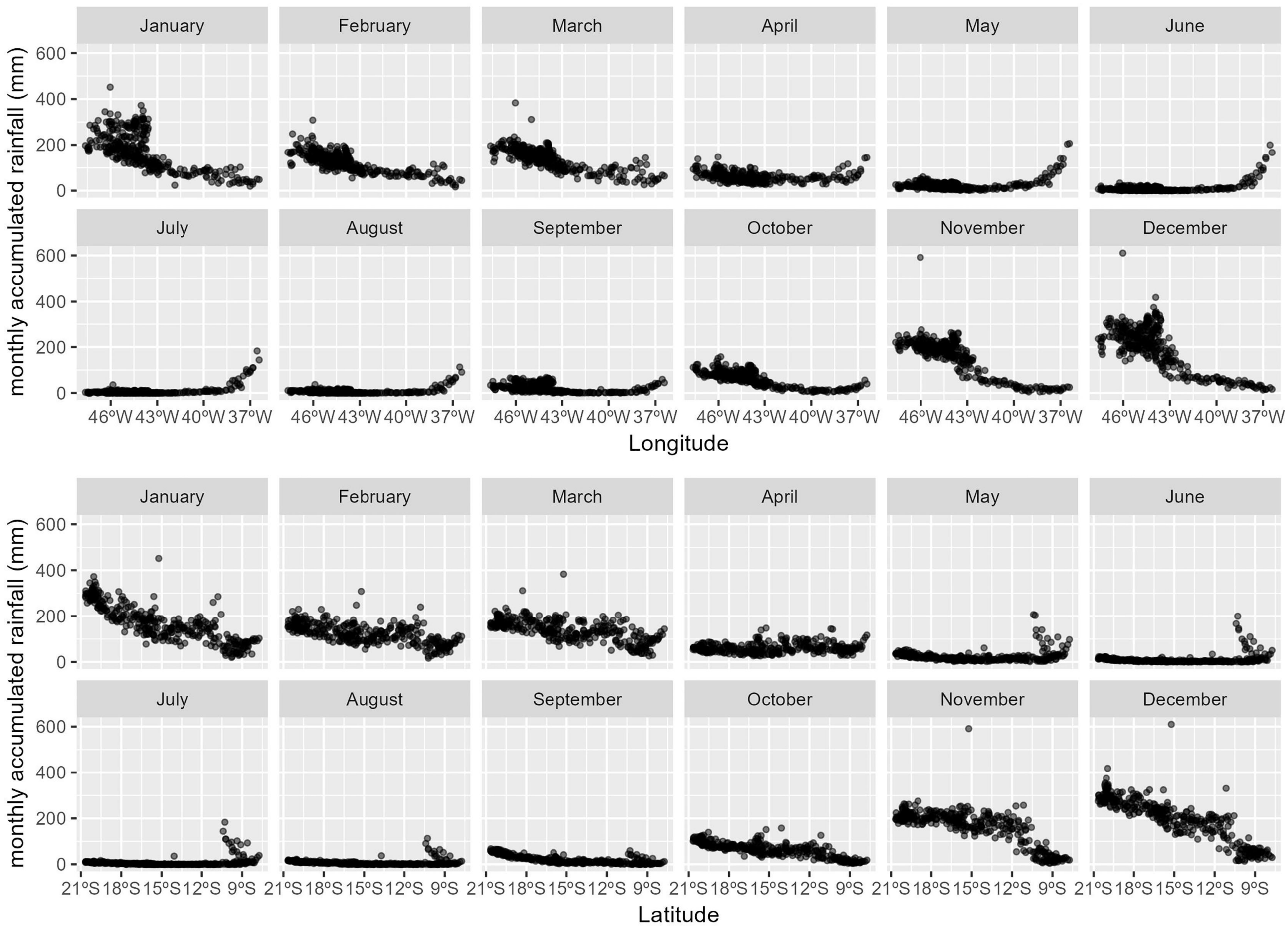

Figure 2 - Examination of the correlation between accumulated rainfall and geographical coordinates within the São Francisco River Basin for the period from 1989 to 2018.

Table 3 - Results of the multiple regression model for the modeling of accumulated rainfall in the São Francisco River Basin for the period from 1989 to 2018.

Month	Variable				R^2
	Intercept	Latitude	Longitude	Longitude^2	
January	1651.06**	-15.39**	88.73**	1.13**	75.34%
February	1381.15**	-3.35**	72.91**	0.98**	65.02%
March	1512.26**	-3.98**	79.46**	1.06**	65.42%
April	2465.96**	1.50**	116.11**	1.40**	34.94%
May	4053.11**	-3.054**	185.60**	2.10**	69.71%
June	3495.67**	-2.44**	159.40**	1.80**	70.72%
July	3124.92**	-1.98**	142.05**	1.60**	73.41%
August	2001.21**	-1.96**	91.89**	1.04**	67.99%
September	1498.28**	-5.10**	70.50**	0.79**	77.57%
October	1363.19**	-4.91**	70.52**	0.89**	80.25%
November	-654.00**	-7.41**	-13.55**	0.057**	81.27%
December	-579.70**	-16.33**	-11.55**	0.015**	85.1%

**, the estimate is significant at the 5% significance level according to the t-test; R^2, coefficient of determination.

Table 4 - Leave-one-out cross-validation results for accumulated rainfall data in the São Francisco River Basin for the period from 1989 to 2018.

Month	Spherical			Exponential			Gaussian		
	MAE	RMSE	R^2	MAE	RMSE	R^2	MAE	RMSE	R^2
January	18.66	28.98	87.58%	17.74	27.69	88.61%	194.50	683.70	4.07%
February	13.43	20.86	76.68%	13.38	20.74	76.89%	9480	93600	~ 0.00
March	15.83	23.48	75.01%	15.06	22.27	77.31%	1941.56	8452.37	0.09%
April	8.73	12.24	64.52%	8.43	11.89	66.35%	84.99	282.08	0.02%
May	4.22	6.49	92.34%	4.23	6.49	92.34%	1103.39	5293.21	0.05%
June	2.61	5.30	94.07%	2.59	5.23	94.20%	1147.89	5538.84	0.06%
July	2.26	5.85	91.38%	2.22	5.77	91.57%	184.87	1002.87	0.02%
August	2.27	4.60	85.55%	1.89	3.80	90.08%	10.05	33.37	26.65%
September	3.62	5.29	91.80%	3.45	5.07	92.47%	212.06	1148.69	0.87%
October	7.57	1.36	87.20%	7.46	11.47	86.95%	73.55	344.49	1.39%
November	15.72	29.70	84.22%	15.22	28.87	85.01%	264.00	12200.00	0.08%
December	16.67	23.65	93.23%	16.24	23.05	93.56%	16.05	22.87	93.66%

Table 5 - Estimates of the parameters of the selected variogram model in cross-validation during Universal Kriging for accumulated rainfall data in the São Francisco River Basin for the period from 1989 to 2018.

Month	Variogram	Nugget	Sill	Range (km)
January	Exponential	0.000	1.417.081	73.086
February	Exponential	0.000	835.396	176.430
March	Exponential	0.000	822.054	74.586
April	Exponential	0.000	249.213	65.629
May	Spherical	0.000	107.524	274.445
June	Exponential	0.000	84.73	194.16
July	Exponential	0.000	81.428	267.30
August	Exponential	0.000	18.36	42.57
September	Exponential	0.000	59.07	74.89
October	Spherical	0.000	153.93	63.83
November	Exponential	0.000	1260.58	88.92
December	Gaussian	0.000	225.63	83.52

southern part of the Lower-Middle to the Upper São Francisco.

4. Discussions

A comparative analysis reveals the association between high altitudes and substantial levels of rainfall in the regions. This relationship is influenced by multiple factors, including the climatic characteristics of the São Francisco River Basin (SFRB) and the influence of the maritime environment near the mouth of the São Francisco River (Oliveira *et al.*, 2022). The Upper São Francisco, situated in the state of Minas Gerais, boasts elevated altitudes and a humid climate classification. The Middle São Francisco, representing the largest physiographic region of the basin, showcases diverse climatic types. It shares a climate similar to that of the Upper São Francisco, but precipitation levels decrease upon entering the semi-arid region, with climates then classified as dry and semi-arid. The middle-lower portion of the basin is recognized as the driest, with annual rainfall ranging from 350 to 800 mm and an average annual temperature of 27 °C, thus defined as a semi-arid and arid climate. In the Lower São Francisco, owing to its proximity to the ocean, the climate is somewhat milder, with an annual temperature of 25 °C and an average annual rainfall varying between 800 and 1,300 mm. The climate here varies from semi-arid to sub-humid.

In this study, we found that the trend component, adjusted for the effects of latitude and longitude, explained from 35% to 85% of the spatial variability of rainfall in the months of April and December, respectively. Comisso and

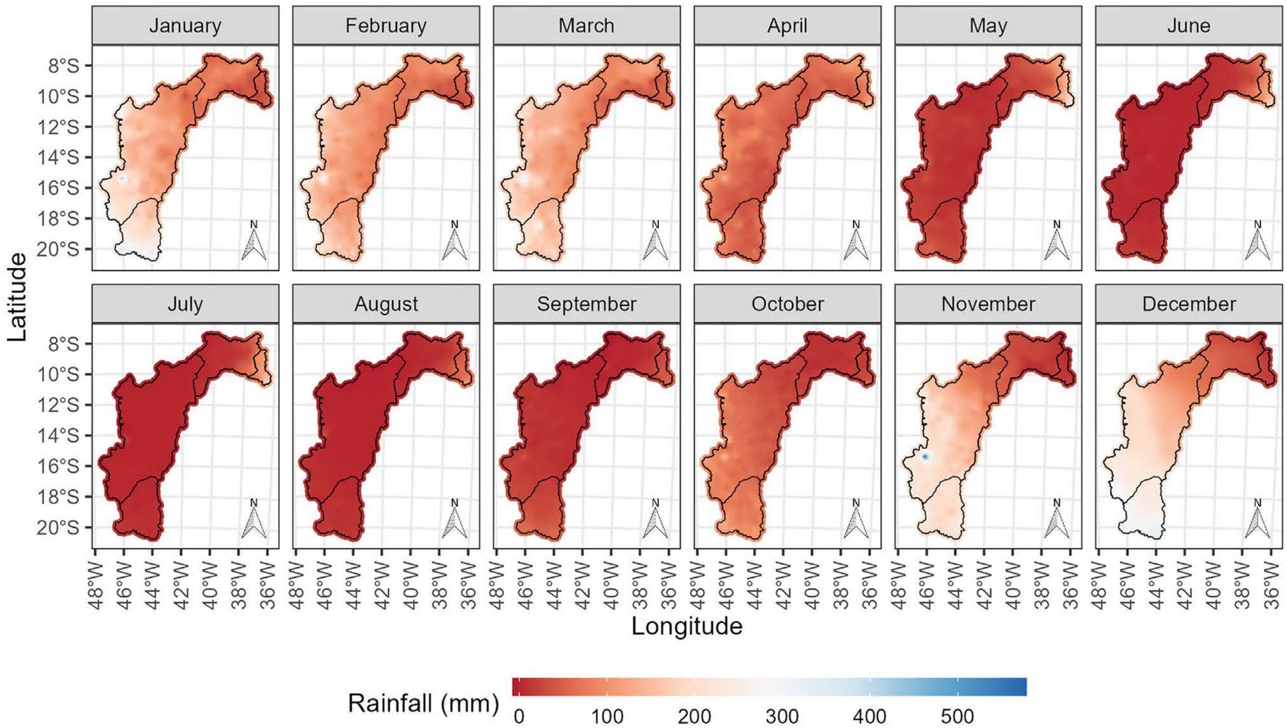

Figure 3 - Maps illustrating the spatial distribution of accumulated rainfall in the São Francisco River Basin for the period from 1989 to 2018.

Medeiros (2021), in their analysis of the spatial distribution of rainfall in the state of Alagoas, Brazil, highlighted that 68.5% of the spatial variability of rainfall during the rainy season in the region can be explained by the model considering longitude as an explanatory variable. Additionally, in a stochastic process, the trend component may account for the large-scale variation of the investigated phenomenon, necessitating a detailed analysis of this component (Medeiros *et al.*, 2022).

The rainfall interpolations and climate projections highlight areas that may face significant reductions in water availability. These identified vulnerable areas can be targeted for specific interventions to enhance water resilience, such as implementing water conservation practices or investing in water infrastructure. Therefore, given the projection of drier conditions, the study's conclusions can inform the development of more efficient and sustainable irrigation systems. This may involve adopting smart irrigation practices, such as drip irrigation, and implementing technologies that maximize water efficiency.

In alignment with the themes addressed in our study, we delve into the discussion of rainfall interpolation maps in the São Francisco River Basin. In this context, we outline that prospective forecasts delineate a trajectory in which the climate in the basin leans towards an even drier state, largely owing to its location within the semi-arid region of the country. Research has unveiled projections indicating an annual decrease of approximately 9% in water availability in the area, further aggravated by an additional increment of about 12% per year during dry periods (Marengo *et al.*, 2017). Thus, emerges a scenario where imminent climate shifts hold the intrinsic potential to exacerbate the pre-existing levels of exposure and socio-economic vulnerability in the region. Such impacts are notably evident in the more frequent occurrences of droughts and subsequent water scarcity, thereby delineating a pivotal intersection between rainfall interpolation analyses and ongoing climate transformation projections.

In convergence with the perspectives that underlie our investigations, we initiate an intrinsic discussion concerning rainfall interpolation maps within the basin. In this context, it becomes evident that the ramifications of adverse climatic reality will substantially reverberate across a multitude of economic sectors. Among them, hydroelectric power generation, agriculture, livestock, and industry emerge, each subject to the vicissitudes of this climatic dynamic. We emphasize that in the São Francisco basin, agriculture is primarily rooted in subsistence practices, characterized by rainfed cultivation (Vieira *et al.*, 2021). Thus, the inherent social relevance of this agricultural framework is undeniable, constituting a substantial pillar for the sustainability of rural communities in the region (Marengo *et al.*, 2020). By establishing these connections, we solidify a crucial dialogue between pluviometric interpolation analyses and the multifaceted repercussions that resonate in the delineated sectors, bestowing a deeper layer of understanding onto the complex intersection of climate, economy, and socio-environmental dynamics.

Amidst these climatic transformations, a clear concern arises regarding the escalating socio-economic vulnerability of the region. Climate change projections can further exacerbate the challenges faced by local communities, given that their subsistence is intricately tied to agriculture and limited water resources. It is crucial to recognize that climate change impacts not only environmental aspects but also holds profound implications across social, economic, and cultural domains. In this context, the implementation of adaptive measures and policies for natural resource management becomes imperative to mitigate adverse effects and safeguard the sustainability of populations and economic activities in the São Francisco River Basin.

In the 6th Report of the Intergovernmental Panel on Climate Change (IPCC), it was highlighted that in the context of the Northeast region, a reduction in the already scarce rainfall could reach up to 22%. This arid scenario, combined with a temperature increase of 3 °C to 4 °C, has the potential to transform the region into a semi-desert, provided that emission levels remain at elevated levels. Presently, certain areas of the Northeast have already experienced a 20% decrease in rainfall and an average temperature increase of 2.4 °C. The exacerbation of climate change amplifies the possibility of this scenario becoming even more critical in terms of sustainability for the populations in the northeastern region of Brazil. Consequently, the country will face the necessity to conceive strategies to tackle the challenges imposed upon the millions of Brazilians residing in the arid Northeast, whose subsistence capacity could be substantially impaired in that area (Prizibisczki, 2022).

5. Final Remarks

The presented study comprehensively addressed the analysis of rainfall in the São Francisco River Basin over a 30-year period. Through the analysis of descriptive statistics from rainfall stations, it was possible to identify spatial variations in precipitation throughout the year. The coefficient of variation highlighted the high spatial variability of rainfall, especially in the months of June, July, and August, with values exceeding 100%.

In this study, we identified that geographic coordinates, latitude, and longitude, largely explain the spatial variability of rainfall in the basin. Additionally, high-resolution kriging maps were provided for strategic water resource management applications across the entire region. Spatial interpolation through kriging and statistical modeling through multiple regression provided valuable insights into the relationship between precipitation and its geographical coordinates. Identifying more suitable variogram models for different months allowed for a better understanding of the spatial distribution of rainfall. The relationship between geographical coordinates and precipitation indicated the influence of topography on rainfall distribution.

The arid scenario projected for the São Francisco River Basin (SFRB) necessitates adaptive measures and effective management policies to address future challenges. Therefore, the novelty and originality of this study lie in the application of geostatistics and pluviometric maps to the São Francisco River Basin, a critical region facing water resource challenges. Thus, this study can contribute to a comprehensive understanding of rainfall distribution and variability in the basin, providing valuable information for decision-making and strategy development.

In summary, this study investigated the spatial distribution of rainfall for water management and planning in the SFRB. The conducted analyses provided a holistic view of rainfall in the region, enabling an understanding of spatial patterns and shedding light on the challenges faced in a climate change scenario. As the region becomes more vulnerable to climate change, implementing adaptation and mitigation strategies is crucial to ensure the sustainability of local populations and ecosystems.

Several limitations and weaknesses can be highlighted in this research. The inclusion of additional explanatory variables in trend modeling, such as altitude and the effect of seasonality, may enhance the accuracy of the results. The 341 pluviometric stations could be considered a relatively low quantity given the extensive territorial scope of the SFRB, and future research could benefit from including rainfall estimates derived from rain gauges and satellite observations. Moreover, the irregular distribution of spatial stations leads to increased prediction errors during kriging.

References

BARROS, T.H.S.; BENDER, F.D.; SILVA, F.R.B.; JOSÉ, J.V.; COSTA, J.O.; *et al.* Geostatistics as a tool to estimate rainfall variability in Pernambuco. **Agrarian**, v. 13, n. 50, p. 513-520, 2020. doi

BHATTACHARJEE, S.; GHOSH, S.K.; CHEN, J. **Semantic Kriging for Spatio-temporal Prediction**. Singapore: Springer, 2019. doi

CAMBARDELLA, C.A.; MOORMAN, T.B.; NOVAK, J.M.; PARKIN, T.B.; KARLEN, D.L.; *et al.* Field-scale variability of soil properties in Central Iowa soils. **Soil Science Society of America Journal**, v. 58, n. 5, p. 1501-1511, 1994. doi

CBHSF – Comitê da Bacia Hidrográfica do Rio São Francisco. **Plano de Recursos Hídricos 2016/2025**. Belo Horizonte: CBHSF, 2016. Available in https://cbhsaofrancisco.org.br/documentacao/plano-de-recursos-hidricos-2016-2025/, accessed on August 22, 2023.

CHARLES, T.S.; LOPES, T.R.; DUARTE, S.N.; NASCIMENTO, J.G.; RICARDO, H.C.; *et al.* Estimating average annual rainfall by ordinary kriging and TRMM precipitation products in midwestern Brazil. **Journal of South American Earth Sciences**, v. 118, 103937, 2022. doi

COMISSO, H.S.; MEDEIROS, E.S. Mapeamento da precipitação no Estado de Alagoas por meio de técnicas geoestatísticas. **Revista Univap**, v. 27, n. 55, p. 1-12, 2021. doi

FERREIRA, D.B.; BARROSO, G.R.; DANTAS, M.S.; OLIVEIRA, K.L.D.; CHRISTOFARO, C.; *et al.* Pluviometric patterns in the São Francisco River Basin in Minas Gerais, Brazil. **Revista Brasileira de Recursos Hídricos**, v. 26, e27, 2021. doi

GOIS, G.; DELGADO, R.C.; OLIVEIRA JÚNIOR, J.F. Modelos teóricos transitivos aplicados na interpolação espacial do standardized precipitation index (SPI) para os episódios de El Niño Forte no Estado do Tocantins, Brasil. **Irriga**, v. 20, n. 2, p. 371-387, 2015. doi

GRÄLER, B.; PEBESMA, E.J.; HEUVELINK, G.B. Spatio-temporal interpolation using gstat. **The R Journal**, v. 8, n. 1, p. 204-208, 2016. doi

LUCAS, M.C.; KUBLIK, N.; RODRIGUES, D.B.B.; MEIRA NETO, A.A.; ALMAGRO, A.; *et al.* Significant baseflow reduction in the São Francisco River Basin. **Water**, v. 13, n. 1, p. 1-17, 2020. doi

LYRA, G.B.; OLIVEIRA JÚNIOR, J.F.; ZERI, M. Cluster analysis applied to the spatial and temporal variability of monthly rainfall in Alagoas state, Northeast of Brazil. **International Journal of Climatology**, v. 34, n. 13, p. 3546-3558, 2014. doi

LYRA, G.B.; OLIVEIRA-JÚNIOR, J.F.; GOIS, G.; CUNHA-ZERI, G.; ZERI, M. Rainfall variability over Alagoas under the influences of SST anomalies. **Meteorology and Atmospheric Physics**, v. 129, p. 157-171, 2017. doi

MAGERSKI, J.M; VIRGENS FILHO, J.S. Avaliação da técnica de krigagem ordinária utilizando o modelo geoestatístico estável no preenchimento de falhas de séries de precipitação pluviométrica nas sub-bacias hidrográficas localizadas em regiões de classificação climática distintas no estado do Paraná. **Revista Brasileira de Geografia Física**, v. 14, n. 4, p. 2149-2171, 2021. doi

MARENGO, J.A.; CUNHA, A.P.M.; NOBRE, C.A.; RIBEIRO NETO, G.G.; MAGALHAES, A.R.; *et al*. Assessing drought in the drylands of northeast Brazil under regional warming exceeding 4 °C. **Natural Hazards**, v. 103, p. 2589-2611, 2020. doi

MARENGO, J.A.; TORRES, R.R.; ALVES, L.M. Drought in Northeast Brazil - Past, present, and future. **Theoretical and Applied Climatology**, v. 129, p. 1189-1200, 2017. doi

MEDEIROS, E.S.; LIMA, R.R.; OLINDA, R.A.; DANTAS, L.G.; SANTOS, C.A.C. Space-time kriging of precipitation: Modeling the large-scale variation with model GAMLSS. **Water**, v. 11, n. 11, p. 2368, 2019. doi

MEDEIROS, E.S.; LIMA, R.R.; OLINDA, R.A.; DANTAS, L.G.; SANTOS, C.A.C. Performance assessment of spatio-temporal regression kriging with GAMLSS models as trends. **Anais da Academia Brasileira de Ciências**, v. 94, n. 3, e20211241, 2022. doi

OLIVEIRA, W.S.; MEDEIROS, E.S.; SILVA, A.Q.; OLIVEIRA, L.A. Spatial mapping of annual rainfall in the São Francisco River Basin. **Revista Ambiente & Água**, v. 17, e2762, 2022. doi

PARKER, S.R.; ADAMS, S.K.; LAMMERS, R.W.; STEIN, E.D.; BLEDSOE, B.P. Targeted hydrologic model calibration to improve prediction of ecologically-relevant flow metrics. **Journal of Hydrology**, v. 573, p. 546-556, 2019. doi

PEREIRA, S.B.; PRUSKI, F.F.; SILVA, D.D.; RAMOS, M.M. Study of the hydrological behavior of São Francisco River and its main tributaries. **Revista Brasileira de Engenharia Agrícola e Ambiental**, v. 11, n. 6, p. 615-622, 2007. doi

PRIZIBISCZKI, C. **O Brasil no Relatório do IPCC - O que Podemos Esperar do Futuro no Cenário de Mudanças Climáticas. 2022**. Available in https://oeco.org.br/reportagens/o-brasil-no-relatorio-do-ipcc-o-que-podemos-esperar-do-futuro-no-cenario-de-mudancas-climaticas, accessed on August 22, 2023.

R CORE TEAM. **R: A Language and Environment for Statistical Computing**. Vienna: R Foundation for Statistical Computing, 2023. Available at https://www.R-project.org/.

ROCHA, L.H.D.S.; SILVA, D.F.D.; KAYANO, M.T.; BONFIM, O.E.T. Homogeneity, extreme events and climatic causes: São Francisco River Basin. **Revista Brasileira de Meteorologia**, v. 37, p. 199-212, 2022. doi

RUFFATO-FERREIRA, V.; BARRETO, R.C.; OSCAR JÚNIOR, A.; SILVA, W.L.; VIANA, D.B.; *et al.* A foundation for the strategic long-term planning of the renewable energy sector in Brazil: hydroelectricity and wind energy in the face of climate change scenarios. **Renewable and Sustainable Energy Reviews**, v. 72, p. 1124-1137, 2017. doi

SANTOS, G.R.; OLIVEIRA, M.S.; LOUZADA, J.M.; SANTOS, A.M.R.T. Krigagem simples versus krigagem universal: Qual o preditor mais preciso? **Energia na Agricultura**, v. 26, n. 2, p. 49-55, 2011. doi

SANTOS, A.; LOPES, P.M.O.; SILVA, M.V.; JARDIM, A.M.R.F.; MOURA, G.B.A.; *et al.* Causes and consequences of seasonal changes in the water flow of the São Francisco river in the semiarid of Brazil. **Environmental and Sustainability Indicators**, v. 8, p. 100084, 2020. doi

SILVA, D.F.; BRITO, J.I.B. Variabilidade do vento na bacia hidrográfica do rio São Francisco durante a ocorrência da ZCAS. **Ambiência**, v. 4, n. 2, p. 221-235, 2008.

SOBRAL, M.D.C.M.; DE ASSIS, J.M.O.; DE OLIVEIRA, C.R.; DA SILVA, G.M.N.; MORAIS, M.; *et al.* Impact of climate change on water resources in the subdom of the São Francisco River Basin - Brazil. **REDE - Revista Eletrônica do PRODEMA**, v. 12, n. 3, p. 95-106, 2018. doi

VIEIRA, R.M.D.S.P.; TOMASELLA, J.; BARBOSA, A.A.; POLIZEL, S.P.; OMETTO, J.P.H.B.; *et al.* Land degradation mapping in the MATOPIBA region (Brazil) using remote sensing data and decision-tree analysis. **Science of The Total Environment**, v. 782, p. 146900, 2021. doi

VILAR, R.D.A.A.; CORREIA, M.D.F.; NÓBREGA, R.S.; ARAGÃO, M.R.D.S.; BARBIERI, L.F.P.; *et al.* Impact evaluation of dried dryes in the Northeast Brazil on electric power generation through the newave model: Projection of

afluent and stored energies. **Revista Brasileira de Meteorologia**, v. 35, n. 1, p. 89-98, 2020. doi
WEBSTER, R.; OLIVER, M.A. **Geostatistics for Environmental Scientists**. 2. ed. Chester: Wiley, 2007. doi
WICKHAM, H. **ggplot2: Elegant Graphics for Data Analysis**. New York: Springer, 2009.

Internet Resources

HydroWeb Portal, https://www.snirh.gov.br/hidroweb/apresentacao.