Mercator, Fortaleza, v. 23, e23028, 2024. ISSN:1984-2201

EXPLORING CLIMATIC PATTERNS AMONG BRAZILIAN CAPITALS

https://doi.org/10.4215/rm2024.e23028

Souza, Amaury [a]* - Medeiros, E.S. [b]

(a) PhD in Environmental Technologies
ORCID: http://orcid.org/0000-0001-8168-1482. **LATTES:** http://lattes.cnpq.br/6113404589622533.
(b) PhD in Statistics and Agricultural Experimentation
ORCID: https://orcid.org/0000-0002-9694-4019. **LATTES:** http://lattes.cnpq.br/5052910988468793.

Article history:
Received 29 March, 2024
Accepted 04 October, 2024
Published 10 December, 2024

(*) CORRESPONDING AUTHOR
Address: UFMS. Cidade Universitaria s/nº, CEP: 79070-900, Campo Grande (MS), Brazil. Phone: (+55 67) 3457186
E-mail: amaury.souza@ufms.br

ARTICLE

Abstract

This study aims to examine the patterns of climatic variables and delineate climatically homogeneous regions among Brazilian capitals using Principal Component Analysis (PCA) and clustering techniques. Annual climate data from 26 state capitals and the Federal District of Brazil were analyzed, covering the period from 1958 to 2022, representing different regions of the country. The PCA results revealed that the first three principal components accounted for 91.8% of the total data variability. Four climatically homogeneous regions were identified: the East region, the Northeast region, the North region and another region that comprises the rest of the country. Cluster Analysis (CA) also confirmed the formation of four homogeneous groups among municipalities based on climatic variables. The analysis of spatial distribution showed significant variations in annual climate variables between Brazilian capitals, with maximum values observed for maximum and minimum temperatures, precipitation, wind speed, relative humidity, evapotranspiration and soil moisture. These findings contribute to a deeper understanding of climate variability in Brazilian capitals and provide a basis for future studies on climate dynamics and regional impacts.

Keywords: Principal Component Analysis, Cluster Analysis, Multivariate Analysis; Biplot

Resumo / Resumen

EXPLORANDO PADRÕES CLIMÁTICOS ENTRE CAPITAIS BRASILEIRAS

Este estudo tem como objetivo examinar os padrões das variáveis climáticas e delinear regiões climaticamente homogêneas entre as capitais brasileiras utilizando Análise de Componentes Principais (ACP) e técnicas de agrupamento. Foram analisados dados climáticos anuais de 26 capitais e do Distrito Federal do Brasil, abrangendo o período de 1958 a 2022, representando diferentes regiões do país. Os resultados do PCA revelaram que os três primeiros componentes principais representaram 91,8% da variabilidade total dos dados. Foram identificadas quatro regiões climaticamente homogêneas: a região Leste, a região Nordeste, a região Norte e outra região que compreende o restante do país. A Análise de Cluster (AC) também confirmou a formação de quatro grupos homogêneos entre os municípios com base nas variáveis climáticas. A análise da distribuição espacial mostrou variações significativas nas variáveis climáticas anuais entre as capitais brasileiras, sendo observados valores máximos para temperaturas máximas e mínimas, precipitação, velocidade do vento, umidade relativa, evapotranspiração e umidade do solo. Essas descobertas contribuem para uma compreensão mais profunda da variabilidade climática nas capitais brasileiras e fornecem base para estudos futuros sobre dinâmica climática e impactos regionais.

Palavras-chave: Análise de Componentes Principais, Análise de Cluster, Análise Multivariada

EXPLORANDO LOS PATRONES CLIMÁTICOS ENTRE LAS CAPITALES BRASILEÑAS

Este estudio tiene como objetivo examinar los patrones de las variables climáticas y delinear regiones climáticamente homogéneas entre las capitales brasileñas utilizando el Análisis de Componentes Principales (PCA) y técnicas de agrupamiento. Se analizaron datos climáticos anuales de 26 capitales de estados y del Distrito Federal de Brasil, cubriendo el período de 1958 a 2022, representando diferentes regiones del país. Los resultados del PCA revelaron que los primeros tres componentes principales representaron el 91,8% de la variabilidad total de los datos. Se identificaron cuatro regiones climáticamente homogéneas: la región Este, la región Noreste, la región Norte y otra región que comprende el resto del país. El Análisis de Conglomerados (AC) también confirmó la formación de cuatro grupos homogéneos entre municipios en función de variables climáticas. El análisis de la distribución espacial mostró variaciones significativas en las variables climáticas anuales entre las capitales brasileñas, observándose valores máximos para las temperaturas máximas y mínimas, las precipitaciones, la velocidad del viento, la humedad relativa, la evapotranspiración y la humedad del suelo. Estos hallazgos contribuyen a una comprensión más profunda de la variabilidad climática en las capitales brasileñas y proporcionan una base para futuros estudios sobre la dinámica climática y los impactos regionales.

Palabras-clave: Análisis de Componentes Principales, Análisis de Conglomerados, Análisis Multivariado; Biplot.

INTRODUCTION

The patterns of climatic variables and the delineation of climatically homogeneous regions between Brazilian capitals involve a complex and detailed analysis of climatic conditions throughout the country (Hänsel et al., 2016). Brazilian capitals represent a diversity of climates due to Brazil's vast territorial extension and its variation, which includes plains, mountains, tropical forests, and coastal areas (Fritzsons et al., 2017). To understand the patterns of climatic variables among Brazilian capitals, it is necessary to consider factors such as temperature, precipitation, humidity, atmospheric pressure, winds, and other relevant climatic elements.

Brazil has a great diversity of temperatures, ranging from the hot and humid equatorial climate in the Amazon to the cooler subtropical climate in the south of the country (Dubreuil et al., 2018). Capital cities located in the north and northeast generally have higher temperatures throughout the year, while those in the south and southeast may experience more pronounced seasonal variations. Rainfall also varies significantly between Brazilian capitals. Capitals such as Manaus and Belém experience high rainfall throughout the year, while others, such as Brasília and São Paulo, have distinct periods of rain and drought (Rao et al., 2016). Relative humidity can vary significantly between different regions of Brazil. Coastal areas tend to have a higher humidity, while more continental regions may have lower humidity. Wind patterns and atmospheric pressure also influence the climate of different regions. Capitals located in coastal areas may be influenced by offshore winds, while others may be subject to specific regional high-pressure systems or winds (Fritzsons et al., 2017).

The delineation of climatically homogeneous regions among Brazilian capitals involves the identification of consistent patterns in the climatic variables mentioned above. This can be done through statistical analyses such as cluster analysis, principal component analysis (PCA), and other multivariate analysis techniques.

Precipitation is considered the most important meteorological variable for agricultural production and regional climate. Knowledge of the behavior of this variable is essential for the planning of activities in many sectors of society, such as agriculture, the energy sector, water resources, and the hydrological cycle. Agriculture has the rainfall pattern as the main source of water, which sometimes compromises agricultural production due to its variability, sometimes with long periods of drought, sometimes with extreme events that exceed the water retention capacity of the soil, causing flooding (da Silva et al., 2007; Soccol et al., 2010; Vieira et al., 2010), and in some cases flooding entire communities or even large plains. In addition to influencing agriculture, severe droughts and droughts affect the water level of water sources and hydroelectric power plant reservoirs, causing problems for urban water supply and power generation (Rodrigues et al., 2013; Oliveira Junior et al., 2014; Teodoro et al., 2015).

The total area of Brazil is approximately 8,515,767 square kilometers. Regarding the area destined to agriculture, about 30% of the Brazilian territory is used for agricultural activities, which corresponds to approximately 2.55 million square kilometers. This number can vary over time due to changes in farming practices, urban sprawl, and other factors. the most important crops are soybeans, corn, cotton, sugarcane and irrigated rice, followed by beef cattle, favored by extensive flat areas of fields and, finally, initiatives for the construction of hydroelectric and thermoelectric plants that contribute to the degradation of environmental conditions (CONAB, 2014). On the other hand, in this region, there are several meteorological phenomena, from micro-scale to large-scale, which require a greater understanding, for example, soil-vegetation-atmosphere interaction processes during droughts and floods, which have local and climatic effects and, especially, on precipitation patterns (Domingues et al., 2004; Mesquita et al., 2013).

There are several climatological studies that use cluster analysis to investigate meteorological systems and climatic variables in Brazil that include:Identification of Precipitation Patterns, studies have applied cluster analysis to identify precipitation patterns in different regions of Brazil over different seasonal and annual periods (Souza et al 2022; Chou et al, 2014; Santos et al, 2019), to understand the distribution of temperature in different areas of the country and identify warming or cooling trends in certain regions (Santos, et al, 2019; Marengo et al, 2001), studies on wind patterns and atmospheric circulation through cluster techniques to understand atmospheric dynamics and their effects on local and regional climate (Guarienti et al 2020; Grimm et al, 2003; Carvalho et al, 2004), cluster analysis has been used to investigate climate variability at different temporal and spatial scales (Amorim et al, 2020), including extreme events such as droughts, heavy rainfall, and heat waves, climate modeling, cluster analysis can be used to evaluate the performance of climate models in reproducing observed patterns of climate variables, studies have applied cluster analysis to assess the impacts of climate change in different regions of Brazil, identifying patterns of variability and possible future trends (Silva et al, 2022).

These are just a few examples of how cluster analysis has been used in climatological studies on meteorological systems and climate variables in Brazil. The results of these studies contribute to a better understanding of the Brazilian climate and assist in decision-making in several areas, including agriculture, water resources, urban planning, and environmental management.

In an attempt to establish spatial patterns for meteorological variables, the cluster analysis (CA) tool is the most widely used tool in the literature on classification processes (Jackson and Weinand, 1995). The AC technique consists of determining the level of similarity or dissimilarity between individuals by applying a clustering function to a given variable (Macedo et al., 2010; Oliveira et al., 2010). Research in various regions of the world has applied multivariate techniques to investigate the spatial, seasonal, and annual variability of climatic variables.

Unal et al. (2003) suggest the use of the AC technique because it is an effective statistical algorithm that can use observed meteorological elements. Among the main clustering methods, Ward's method is widely used in climatological studies with satisfactory results (Lyra et al., 2014). Ramos (2001) analyzed rainfall distribution patterns by the k-means methods and Ward's algorithm for the northeast (NE) of Spain, where the two methods showed similar evaluations, but Ward's method identified the best patterns. Unal et al. (2003) used the AC technique in air temperature data and obtained seven homogeneous regions in Turkey, highlighting the a priori hypothesis of divergence between them. These results support the effective use of the AC technique.

Thus, the identification of similar regions in the spatio-temporal variability of rainfall provides the basis for new data that can be used in climate study programs with greater precision.

Thus, this study can perform a cluster analysis (K-Means) in the context of Brazil. The study also applies this method to precipitation and wind temperature and velocity (CRU) data to create spatially homogeneous groups of these climate variables in Brazil.

Material and Methods

Monthly climate data, considering from 1958 to 2022, obtained from the TerraClimate dataset (https://www.climatologylab.org/terraclimate.html), which is a high-resolution global dataset that provides information on various climate variables on a monthly scale.

1. tmax (Maximum Temperature): This variable represents the monthly maximum temperature. It is the highest temperature recorded during the month in question. The unit is degrees Celsius (°C).

2. tmin (Minimum Temperature): This variable represents the monthly minimum temperature. It is the lowest temperature recorded during the month in question. The unit of measurement is also in degrees Celsius (°C).

3.ppt (Precipitation): This variable represents the monthly precipitation, i.e., the amount of rain that fell during the month. The unit of measurement is millimeters (mm).

4. ws (Wind Speed): This variable represents the monthly average wind speed. It is a measure of the average wind speed during the month, measured in meters per second (m/s).

5. q (Runoff): This variable represents monthly runoff, which is the amount of water that flows surfacely, often resulting from excessive precipitation. The unit of measurement is in millimeters (mm).

6. aet (Current Evapotranspiration): This variable represents the actual monthly evapotranspiration, which is the amount of water that evaporates from the soil surface and is transpired by plants during the month. The unit of measurement is millimeters (mm).

7. soil (Soil Moisture): This variable represents the monthly soil moisture. It indicates the amount of water present in the soil at the end of the month. The unit is in millimeters (mm).

These climate variables are essential for understanding the climatic conditions in a given region over time and can be used in a variety of applications, such as climate studies, environmental monitoring, hydrological modeling, agriculture, and more. TerraClimate provides high-quality, spatial-resolution data for these variables on a global scale.

Consolidating the monthly data for annual averages, adding precipitation (ppt), runoff (q), real evapotranspiration (aet) and Soil Moisture (soil) throughout the year, and calculating the average for the other variables. We then averaged over the 65 years to get the annual average for each city.

Due to the difference in magnitude between the scales of the variables, we applied standardization to the data for each variable by means of the following expression:

$$z_i = \frac{x_i - \bar{x}}{s} \tag{1}$$

In Equation 1, xi represents the observed value of the variable in the -th capital, while x and s represent the mean and standard deviation, respectively. Therefore, all multivariate analyses will be conducted using the standardized data.

Principal component analysis (PCA), using the structure of covariances, was performed to determine the linear combinations of climatic variables. Mathematically, PCA represents linear combinations of random variables, and the number of climatic variables used in our study. The -th principal component is expressed as follows:k

$$Y_k = e_{k,1}Z_1 + e_{k,2}Z_2 + \cdots + e_{k,7}Z_7. \tag{2}$$

In Equation (2), we have the random vector $\mathbf{Z}^t = [Z_1, Z_2, \cdots, Z_7]$ sampled from a population with covariances Σ, whose eigenvalues are $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_7 \geq 0$ and $e_{i,p}$ represent the eigenvectors of the matrix Σ [Ferreira, 2018]. The percentage of total variance explained by the k-th principal component is given by: $\%VarExp(Y_k) = \frac{\lambda_k}{\sum_{i=1}^{p} \lambda_i} \times 100.$

Determining the optimal number of clusters for a K-Means analysis is a major challenge. There are several techniques that can help you choose the appropriate number of clusters. In our study, we applied the Elbow method and the method proposed by Pham et al. (2005).

The method proposed by Pham et al. (2005) is an approach to selecting the optimal number of clusters (K) when applying the K-Means algorithm to a dataset. The fundamental idea behind this method is to evaluate the quality of the resulting clustering for different K values and, based on that evaluation, determine which K value is most appropriate for the data set. The method uses a function called f(K) (Eq. 3) to evaluate the quality of the clusters. This function is a metric that measures how well the data has been clustered for a given value of K. The metric f(K) is defined in a way that reflects the internal cohesion of the clusters (the proximity of the points within the same clusters) and the separation between the clusters (the distance between the clusters).

$$f(K) = \begin{cases} 1, & \text{if } K = 1 \\ \dfrac{S_K}{\alpha_K S_{k-1}}, & \text{if } S_{K-1} \neq 0, \forall\, K > 1 \\ 1, & \text{if } S_{K-1} = 0, \forall\, K > 1 \end{cases} \tag{3}$$

MERCATOR

ARTICLE

In Equation (3), SK it represents the sum of the distortion of all clusters and αK is a weighting factor expressed as follows:

$$f(K) = \begin{cases} 1 - \dfrac{3}{4N_d}, & \text{if } K = 2 \text{ and } N_d > 1 \\ \alpha_{K-1} + \dfrac{1 - \alpha_{K-1}}{6}, & \text{if } K > 2 \text{ and } N_d > 1 \end{cases} \quad (4)$$

In Equation (4), Nd represents the number of dimensions in the dataset. The value of K at which the metric f(K) begins to stabilize or decrease is considered the optimal number of clusters for the dataset. In this study, we considered a k-threshold of 0.80.

Elbow Method: This is one of the most common methods for choosing the number of clusters. It runs K-Means for different numbers of clusters, and for each number it calculates the variance within the clusters (total within sum of square - wss). It then plots the wss relative to the number of clusters and looks for an "elbow" point on the curve where the variance reduction stabilizes. This point is generally considered the optimal number of clusters.

RESULTS

Figure 1 shows the data set used in the statistical analysis

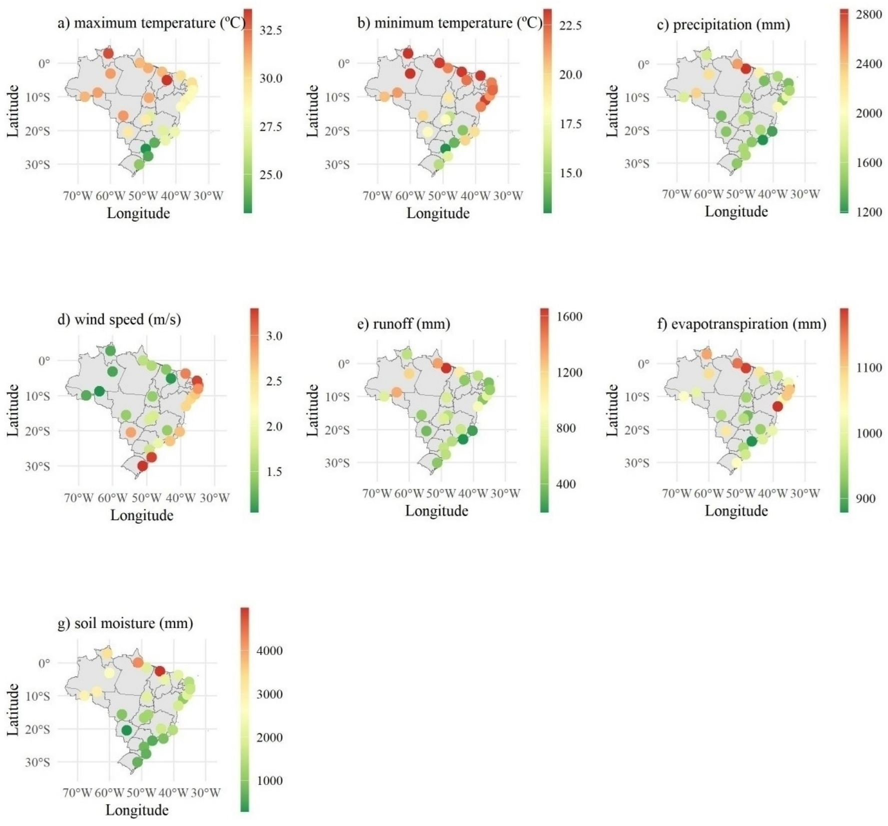

Figure 1 - Annual averages of the variables maximum temperature (a), minimum temperature (b), precpitation (c), wind speed (d), runoff (e), evapotranspiration (f) and soil moisture (g), referring to the 27 Brazilian capitals for the period from 1958 to 2022.

By examining the spatial distribution of annual climate variables, we observed the following maximum annual values: the annual average of the maximum temperature was 33.5°C in Teresina, while the average temperature average was 23 °C in Curitiba. For the the minimum annual average in Curitiba, with 12.5 °C. Regarding precipitation, the maximum annual average was 2513 mm to Macapá, while the minimum annual average was 1188 mm to Rio de Janeiro. The maximum average wind speeds were recorded in Porto Alegre, with 3.29 m/s, while the minimum average was in Teresina, with 1.06 m/s. The maximum average relative humidity index was in Macapá, with 1356 mm, and the minimum in Rio de Janeiro, with 196 mm. The maximum average evapotranspiration was in Salvador, with 1189 mm, and the minimum in São Paulo, with 879 mm. Soil presented its maximum value in São Luís, with 4972 mm, and the minimum in São Paulo, with 586 mm (Figure 1).

Figure 2 shows the correlation of the annual averages of the climatic variables. The data showed strong correlations, both negative and positive, among the climatic variables examined in this study, thus providing justification for a detailed analysis using multivariate statistical techniques.

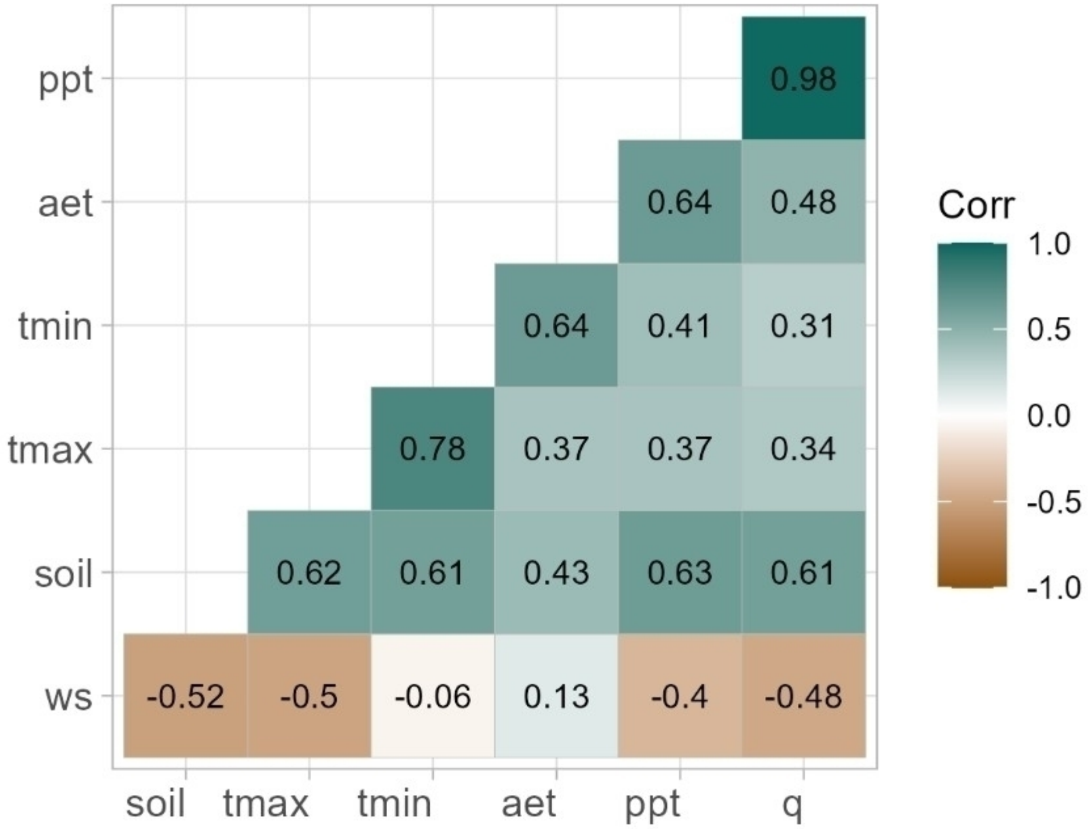

Figure 2 – Correlogram on the annual averages of the variables maximum temperature (tmax), minimum temperature (tmin), precipitation (ppt), wind speed (ws), runoff (q), evapotranspiration (aet) and soil moisture (soil), referring to the 27 Brazilian capitals for the period from 1958 to 2022.

Table 1 shows the results of the multivariate analysis, highlighting the eigenvalues in the PCA and the respective percentages of the total variation explained by each component. In addition, the eigenvectors that represent the linear combination of climatic variables are presented. The first two principal components explained 75.6% of the total variability, with the first principal component accounting for 56.5% of this variability and showing all positive coefficients, except for the variable wind speed.

Components	**PC1**	**PC2**	**PC3**	**PC4**	**PC5**	**PC6**	**PC7**
Standard deviation ($\sqrt{\lambda_k}$)	1.988	1.156	1.064	0.564	0.441	0.254	0.001
Proportion of Variance	56.5%	19.1%	16.2%	4.5%	2.8%	0.9%	0.0%
Cumulative Proportion	56.5%	75.6%	91.8%	96.3%	99.1%	100.0%	100.0%
Variables	**Autovectors**						
Maximum Temperature	0.375	0.071	0.560	-0.434	0.040	0.592	-0.001
Minimum Temperature	0.369	0.458	0.331	0.010	0.398	-0.621	0.000
Precipitation	0.435	-0.116	-0.444	-0.140	0.170	0.043	-0.742
Wind speed	-0.258	0.665	-0.281	0.184	0.414	0.456	0.000
Runoff	0.412	-0.254	-0.432	-0.131	0.365	0.059	0.652
Evapotranspiration	0.339	0.506	-0.299	-0.126	-0.705	-0.038	0.158
Soil moisture	0.428	-0.096	0.158	0.851	-0.094	0.222	0.002

Table 1– Results of the main component analysis applied to climatic variables in the capitals of Brazil.

Therefore, we will construct a Biplot plot presenting the results of the two main components, in order to investigate the relationships between the variables and the capitals when observing the climatic variables (Figure 3).

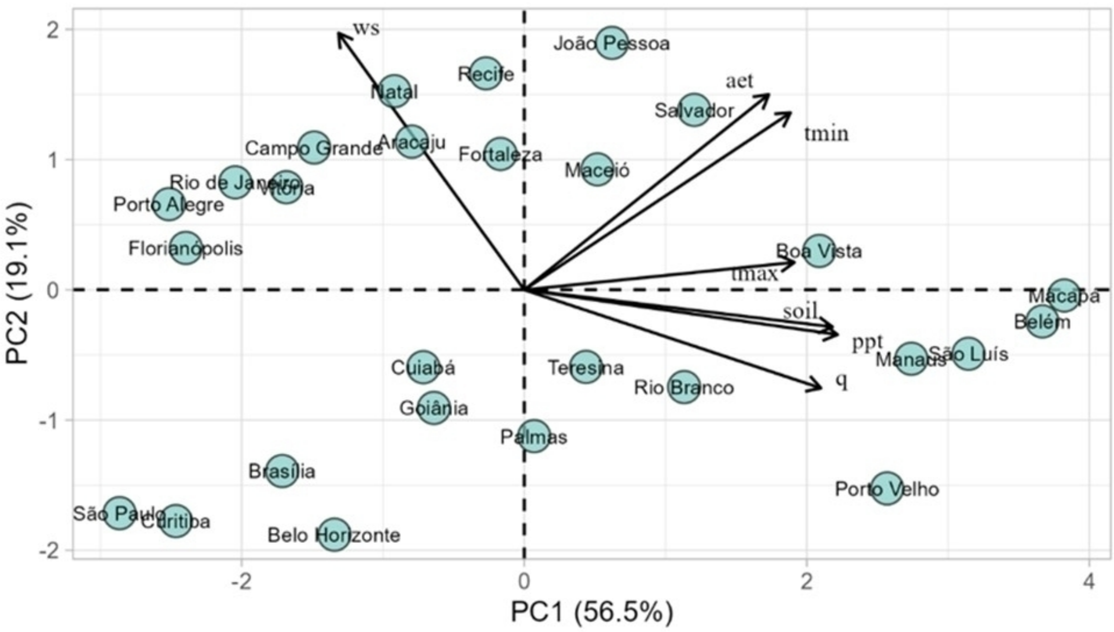

Figure 3 – Biplot of the two main components on the annual averages of the variables maximum temperature (tmax), minimum temperature (tmin), precpitation (ppt), wind speed (ws), runoff (q), evapotranspiration (aet) and sun moisture (soil), referring to the 27 Brazilian capitals for the period from 1958 to 2022.

Figure 4 presents the results for identifying the number of groups for the cluster analysis. Thus, this study proposed the creation of four groups.

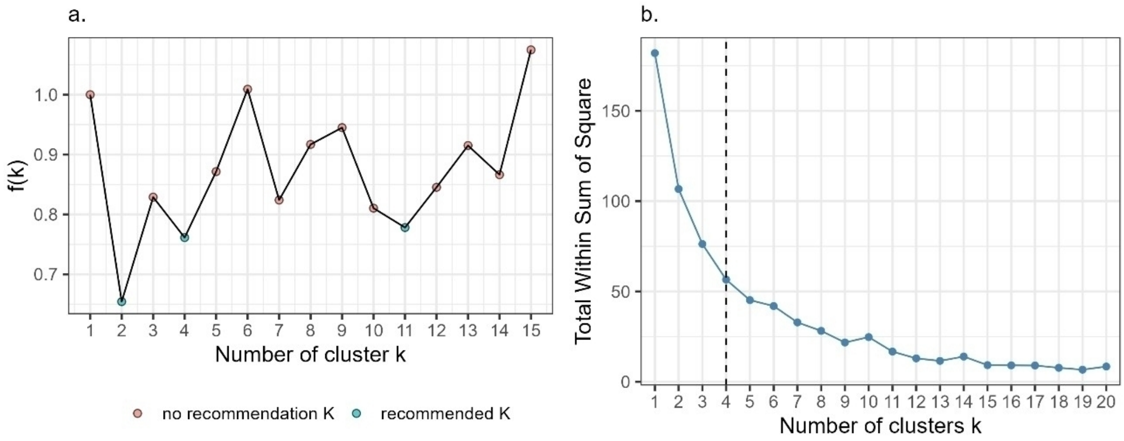

Figure 4 – Determination of the optimal number of clusters for a K-Means analysis using the method proposed by Pham et al. (2005) (a) and Elbow Method (b)

Similar to Figure 3, Figure 5 consists of a Biplot graph, but with the presentation of the cluster analysis, identifying the capitals within each of the four groups.

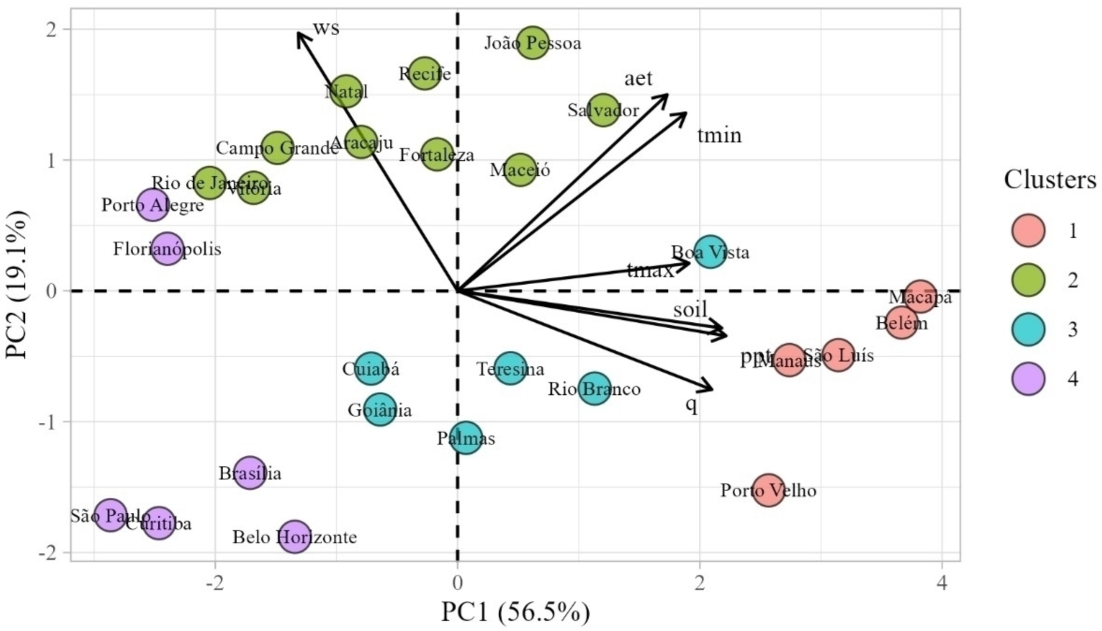

Figure 5 – Biplot with identification of the clusters on the annual averages of the variables maximum temperature (tmax), minimum temperature (tmin), precpitation (ppt), wind speed (ws), runoff (q), evapotranspiration (aet) and soil moisture (soil), referring to the 27 Brazilian capitals for the period from 1958 to 2022.

Figure 6 – Map of the clusters in the capitals

Based on Grouping Analysis (CA), we identified the formation of four homogeneous groups (Fig. 6). Group 1 (G1) covers the municipalities of Porto Velho, Manaus, São Luís, Belém and Macapá. Group 2 (G2) includes the municipalities of Campo Grande, Rio de Janeiro, Vitória, Salvador, Aracaju, Maceió, Recife, João Pessoa, Natal and Fortaleza. Group 3 (G3) is composed of the municipalities of Rio Branco, Porto Velho, Macapá, Cuiabá, Goiânia, Tocantins and Teresina. Group 4 (G4) encompasses the municipalities of Porto Alegre, Florianópolis, Curitiba, São Paulo, Belo Horizonte and Brasília.

The municipalities that make up the G1 groups are located in the North and Northeast regions, while G2 is in the Center-South and Northeast region, G3 covers the North and Midwest, and G4 is in the South and Southeast region.

The regions of Brazil are influenced by different climatic systems that determine their standards of temperature, precipitation and other climate characteristics (Rao et al., 2016; Luiz-Silva et al., 2020). Northern Region: characterized by equatorial climate, with high temperatures and abundant rainfall throughout the year, and small seasonal variations; Northeast Region: presents a great climate diversity, with predominance of semi-arid climate in the backlands and tropical climate in the coastal areas; Midwest Region: It has tropical and tropical climate of altitude, with two well-defined stations: rainy in summer and dry in winter; Southern Region: It has a subtropical and tempered climate, with well-defined stations, hot summers and cold winters, with frost; Center-South Region: A combination of climatic systems from the South, Southeast and Midwest, predominantly of a subtropical and tropical climate, well-defined stations and summer concentrated rains.

These descriptions provide an overview of climatic systems that influence each region. However, it is important to emphasize that there are variations within each of them due to factors such as altitude, proximity to the ocean, relief and local influences.

DISCUSSIONS

The analysis of Principal Components (PCA) applied to the monthly average data of climate variables has yielded significant insights into the patterns of variability and distribution across different regions of Brazil. The retention of the first three Principal Components (PCs), explaining 91.8% of total data variation, is a standard practice in PCA analysis. These PCs encapsulate the most prominent patterns of variability in climate data, effectively capturing the underlying structure and serving as representative indicators of the main standards of climate variability in Brazil.

Correlations between retained PCs and original variables highlight the relationship between identified weather patterns and specific climate variables. The substantial correlation (0.98) observed for PC1 signifies a significant association, suggesting that PC1 represents a distinctive climate pattern.

Analysis of precipitation patterns across different PCs reveals substantial spatial and temporal variations in the distribution of rainfall throughout Brazil. The gradual transition between climate patterns from one region to another, especially between coastal and southwest regions, underscores the intricate interactions between climatic and geographical factors.

In comparison with prior studies in other Brazilian regions such as Pará, Paraná, and Maranhão, the applicability and effectiveness of the main component analysis in identifying distinct climate patterns in diverse geographical contexts (Nery 2005; Santos et al, 2014; Nascimento et al, 2015; Araújo et al, 2015; Pansera et al, 2015; Gebert et al 2018; Souza, et al, 2022) are underscored. The variation in the most correlated months with each PC across different regions suggests significant differences in precipitation patterns, even within the same country.

The findings of this study establish a robust foundation for future investigations into climate patterns and precipitation trends in Brazil. The application of diverse grouping and analysis techniques, including K-Means, hierarchical, and hybrid methods, in precipitation regionalization studies can aid in identifying homogeneous regions and enhance our understanding of spatial rainfall distribution on a regional scale. In summary, principal component analysis furnishes valuable insights into climate variability standards and precipitation distribution in Brazil, contributing to a nuanced understanding of climate processes and offering crucial information for environmental planning and management.

The Principal Component Analysis Method (PCA) emerges as an innovative alternative to conventional cluster analysis approaches (AC) for identifying homogeneous groups. In PCA, variables are categorized into groups based on their load factor values. It is noteworthy that the "PCA regions" exhibit an intriguing feature, being potentially "confusing" due to the possibility of overlapping solutions, where some variables may be included in more than one group (Gong and Richman, 1995). The analyses performed indicate the presence of four distinct regions (or groups) of seasonal climatic environments on the Brazilian territory.

Numerous climatological studies focused on Brazil, utilizing parameters provided by national agencies, cover a broad spectrum of analyses, including rainfall characterization, trends in extreme rainfall events, vegetation dynamics, as well as trends in maximum and minimum temperatures and wind speed (Grimm, 2003; Nobre & Borma, 2009; Nobre et al 2016; Nobre, 2014; Klink & Machado, 2005; Klink & Moreira 2002; Marengo and Nobre & Tomasella 2001; Marengo, 2009; Artaxo et al, 2002).

These studies provide a deeper understanding of temperature extremes in Brazil and their implications for regional climate and natural systems. They are pivotal for formulating adaptation and mitigation policies to address climate change at local, regional, and national levels.

There exists a gap in information regarding the homogeneity of climatological data series in Brazil, as exemplified by Roziane's et al 2012, study for the state of Minas Gerais. Detecting heterogeneities necessitates a combination of statistical and metadata-based methods, and the absence of such information renders the assessment of homogeneity in climatic time series more challenging.

Brazil is recognized for its significant spatiotemporal variability in precipitation, characterized by the highest rainfall levels in coastal and forested regions such as the Atlantic Forest and the Amazon biome, which gradually decrease towards the interior. This variability is closely linked to the presence and dynamics of various meteorological systems, including the Intertropical Convergence Zone (ITCZ), cold fronts, and intense cyclones, among others (Molion and Bernardo, 2002). Moreover, phenomena like the El Niño Southern Oscillation (ENSO) and the Atlantic Ocean dipole play pivotal roles in shaping the quantity and distribution of rainfall (Kayano and Andreoli 2006; Andreoli and Kayano, 2009).

During El Niño events, for instance, the warming of the Equatorial Pacific Ocean often leads to extended dry periods in Northeast Brazil (NEB), while La Niña events typically bring about high rainfall variability in this region. The favorable dipole of the Atlantic Ocean also exerts influence on precipitation patterns, where temperature increases are associated with drier conditions due to alterations in atmospheric circulation (Berlato and Fontana, 2003).

Various studies have delved into the interrelationships between climatic anomalies in the tropical Atlantic and the El Niño phenomenon in the Pacific, proposing an atmospheric connection that modulates wind and temperature patterns, consequently affecting rainfall distribution across Brazil (Hastenrath, 2006). Although the impacts of ENSO and the Atlantic Ocean dipole on temperatures are relatively smaller compared to their effects on rainfall, they nonetheless remain significant factors in observed climate patterns (Halpert, and Ropelewski 1992).

Cluster analysis with total precipitation and average, minimum, and maximum temperatures poses significant challenges when evaluating data homogeneity and detecting signs of climatic variability. Total precipitation and temperature data are often affected by homogeneity issues, especially due to changes in measurement practices, location of meteorological stations, instrumentation, and surrounding environment over time. Data heterogeneity can lead to incorrect or distorted cluster groupings, as different meteorological stations may exhibit different behaviors over time due to non-climatic factors.

The temporal resolution of data can be excessively high for certain analytical purposes. Cluster analysis with data can result in groupings overly sensitive to short-term variations that may not represent significant climatic variations.

Detecting signals of climatic variability in data can be complex due to the presence of noise and random fluctuations. Climatic variability patterns, such as El Niño and La Niña, may manifest on temporal scales broader than the data resolution. Therefore, special care is needed when interpreting cluster analysis results regarding these patterns.

The representativeness of meteorological stations used in cluster analysis may be questionable. In many cases, stations are not distributed homogeneously or representatively of the study region, leading to inadequate generalizations about local climate patterns.

The presence of outliers and data failures can significantly affect the results of cluster analysis, distorting the identified groupings and impairing the interpretation of climatic patterns.

In summary, while cluster analysis with total precipitation and temperature data can offer insights into climatic variability, it is essential to critically address the limitations inherent in this data and carefully consider challenges related to data homogeneity, temporal resolution, interpretation of climatic variability signals, and the representativeness of meteorological stations used.

CONCLUSION

In this study, we conducted an extensive analysis of climate variables among Brazilian capitals using Principal Component Analysis (PCA) and clustering techniques. Analysis of annual climate data spanning 1958 to 2022 provided insights into the spatial and temporal variability of weather patterns across the country.

The PCA results revealed that the first three principal components accounted for a significant portion (91.8%) of the total variability in the climatic data, effectively capturing distinct patterns among Brazilian capitals. Through PCA, we identified changes in climate patterns occurring in a southwest-to-northeast direction, highlighting the analysis's ability to discern known climatic patterns in Brazil.

Furthermore, the grouping analysis based on climatic variables delineated Brazil into four distinct regions: east, northeast, north, and another region encompassing the remainder of the country. These findings offer valuable insights into the spatial distribution of climate patterns and provide a foundation for understanding regional climate dynamics and their impacts.

The application of PCA to monthly climate data proved effective in revealing climate variability and precipitation patterns throughout Brazil. By retaining the first three principal components, PCA captured significant variations in climate across different regions. Moreover, the correlation between retained principal components and original variables underscored the relationships between weather patterns and climate variables.

While PCA emerged as an innovative approach for identifying homogeneous climate groups, challenges such as data homogeneity, temporal resolution, and representativeness of weather stations must be addressed. Additionally, factors like El Niño, La Niña, and the Atlantic Ocean dipole influence Brazil's climate, complicating the interpretation of climate variability patterns.

Despite these challenges, understanding climate variability in Brazil is crucial for informing adaptation and mitigation policies at various levels. Overcoming these obstacles will enhance our comprehension of Brazil's climate processes and support informed decision-making in environmental planning and management.

In conclusion, this study contributes to the broader understanding of Brazil's climate dynamics and lays the groundwork for future research aimed at addressing the complex challenges posed by climate variability in the region.

REFERENCES

Amorim A C B; Lucio P S, Silva C M S, Rodrigues D T, Brito Neto F A. 2020. Regionalização e Análise da Tendência da Precipitação do Rio Grande do Norte Associados a Padrões de TSM. Revista Brasileira de Meteorologia, v. 35, n. 2, 269- 280, DOI: http://dx.doi.org/10.1590/0102-7786352024

Andreoli, R. V., and Kayano, M. T. (2009). "O clima da região Nordeste do Brasil," in Tempo e Clima do Brasil, Vol. 1, eds I. F. A. Cavalcanti, N. J. Ferreira, M. G. A. J. Silva, M. A. F. Dias (São Paulo: Oficina de Textos), 213–233.

Araújo, J. M. S., Maia, A. G., Neves, J. A. (2015) Uso dos índices PCP e PCD na determinação de regiões com precipitação pluvial homogênea. Revista Brasileira de Climatologia, v.17, p.119-139.

Artaxo, P., Martins, J.V., Yamasoe, M.A., Procopio, A.S., Pauliquevis, T.M., Andreae, M.O., Guyon, P., Gatti, L.V., & Leal, A.M.C. (2002). "Physical and chemical properties of aerosols in the wet and dry seasons in Rondônia, Amazonia." Journal of Geophysical Research: Atmospheres, 107(D20), ACH 8-1 - ACH 8-12.

Berlato, M. A., and Fontana, D. C. (2003). El Niño e La Niña: Impactos no Clima, na Vegetação e na Agricultura do Rio Grande do Sul; Aplicações de Previsões Climáticas na Agricultura. Porto Alegre: Ed. da UFRGS. 1 ed, 110 pag.

Carvalho, L. M. V., Jones, C., & Liebmann, B. 2004. "Climatology and interannual variability of storm-tracks in southeastern South America". Journal of Climate, Page(s): 88–108.

Chou, S.C., Lyra, A., Mourão, C., Dereczynski, C., Pilotto, I., Gomes, J., Bustamante, J., Tavares, P., Silva, A., Rodrigues, D., Campos, D., Chagas, D., Sueiro, G., Siqueira, G., Marengo, J., 2014. Assessment of Climate Change over South America under RCP 4.5 and 8.5 Downscaling Scenarios. American Journal of Climate Change 3, 512-525.

CONAB—Companhia Nacional de Abastecimento (2014) Acompanhamento de safra brasileira: grãos, décimo levantamento. Disponible in: www.conab.gov.br/. Access in January of 2024

da Silva JC, Heldwein AB, Martins FB, Trentin G, Grimm EL (2007) Análise de distribuição de chuva para Santa Maria, RS. Revista Brasileira de Engenharia Agrícola e Ambiental 11(1):67–72. doi:10.1590/S1415-43662007000100009

Domingues MO, Mendes Junior O, Chou SC, Abreu Sá LD, Manzi AO (2004) Análise das Condições Atmosféricas Durante a 2ª Campanha do Experimento Interdisciplinar do Pantanal Sul Mato-Grossense. Rev Bras Meteorol 19(1):73–88

Dubreuil, V., Fante, K. P., Planchon, O., & Sant'anna Neto, J. L. (2019). Climate change evidence in Brazil from Köppen's climate annual types frequency. International Journal of Climatology, 39(3), 1446-1456. doi: 10.1002/joc.5893

Ferreira, D.F. (2018) Multivariate statistics. 3rd edition, Lavras: Editora UFLA.

Fritzsons, E., Mantovani, L. E., & Wrege, M. S. (2017). Os biomas e o clima das capitais do Brasil. doi: 10.5935/1984-2295.20170075

Gebert, D.M.P., Kist, A., Virgens Filho, J.S. (2018) Determinação De Regiões Homogêneas De Precipitação Pluviométrica No Estado Do Paraná. Por Meio De Técnicas De Análise Multivariada E Geoestatística. Revista Brasileira de Climatologia, Ano 14 – Vol. 23 – JUL/DEZ 2018

Gong, X. and Richman, M. B. (1995) On the application of cluster analysis to growing season precipitation data in North America East of the Rockies, J. Clim., 8, p. 897–931.

Grimm AM (2003) The El Niño impact on the summer monsoon in Brazil: regional processes versus remote influences. J Clim 16(1):263–280

Guarienti, J. A., Almeida, A. K., Neto, A. M., de Oliveira Ferreira, A. R., Ottonelli, J. P., & de Almeida, I. K. (2020). Performance analysis of numerical methods for determining Weibull distribution parameters applied to wind speed in Mato Grosso do Sul, Brazil. Sustainable Energy Technologies and Assessments, 42, 100854.

Halpert, M. S., and Ropelewski, C. F. (1992). Surface temperature patterns associated with the Southern Oscillation. J. Clim. 5, 577–593.

Hänsel, S., Medeiros, D. M., Matschullat, J., Petta, R. A., & de Mendonça Silva, I. (2016). Assessing homogeneity and climate variability of temperature and precipitation series in the capitals of North-Eastern Brazil. Frontiers in Earth Science, 4, 29. doi: 10.3389/feart.2016.00029

Hastenrath, S. (2006). Circulation and teleconnection mechanisms of Northeast Brazil droughts. Prog. Oceanogr. 70, 407–415. doi: 10.1016/j.pocean.2005.07.004

Jackson IJ, Weinand H (1995) Classification of tropical rainfall stations: a comparison of clustering techiques. Int J Climatol 15(9):985–994. doi:10.1002/joc.3370150905

Kayano, M. T., and Andreoli, R. V. (2006). Relationships between rainfall anomalies over northeastern Brazil and the El Niño–Southern Oscillation. J. Geophys. Res. 111, D13101. doi: 10.1029/2005JD006142

Klink, C.A., & Machado, R.B. (2005). "Conservation of the Brazilian Cerrado." Conservation Biology, 19(3), 707-713.

Klink, C.A., & Moreira, A.G. (2002). "Past and current human occupation, and land use." In: Oliveira, P.S., & Marquis, R.J. (Eds.), The Cerrados of Brazil: Ecology and Natural History of a Neotropical Savanna (pp. 69-88). Columbia University Press.

Lyra GB, Oliveira-Júnior JF, Zeri M (2014) Cluster analysis applied to the spatial and temporal variability of monthly rainfall in Alagoas state, Northeast of Brazil. Int J Climatol 34:3546–3558. doi:10.1002/joc.3926

Macedo MJH, Guedes RVS, Sousa FAS, Dantas FRC (2010) Análise do índice padronizado de precipitação para o estado da Paraíba. Brazil. Revista Ambiente & Água 5(1):204–214. doi:10.4136/ambi-agua.130

Marengo, J.A. (2009). "Long-term trends and cycles in the hydrometeorology of the Amazon Basin since the late 1920s." Hydrological Processes, 23(22), 3236-3244.

Marengo, J.A., Nobre, C.A., & Tomasella, J. (2001). "The drought of northeastern Brazil: climatological diagnostics and predictability." Journal of Climate, 14(9), 4595-4608.

ARTICLE

Mesquita FLL, Marques Filho EP, Karam HA, Álvala RCS (2013) Balanço de radiação no pantanal sul Mato-Grossense durante a estação seca. Revista Brasileira de Meteorologia 28(1):65–74.

Molion, L.C.B. and Bernardo, S.O. (2002) Uma revisão da dinamica das chuvas no nordeste brasileiro. Revista Brasileira de Meteorologia, 17, 1-10.

Nascimento, F. C. A.; Araújo, F. R. C. D.; Braga, C. C.; Costa, E. V.S. (2015) Análise de Padrões Espaciais e Temporais da Precipitação no Estado do Maranhão – Brasil. Revista Brasileira de Geografia e Física, v.8, n.2, p.422-430.

Nery, J. T. (2005). Dinâmica climática da região sul do Brasil. Revista Brasileira de Climatologia, v. 1, n. 1. http://dx.doi.org/10.5380/abclima.v1i1.25233

Nobre, A.D. (2014). "The Future Climate of Amazonia." In: Keller M., Bustamante M., Gash J. (eds) Amazonia and Global Change. Geophysical Monograph Series, vol 186. American Geophysical Union, Washington, D.C.

Nobre, C.A., & Borma, L.S. (2009). "Tipping points for the Amazon forest." Current Opinion in Environmental Sustainability, 1(1), 28-36.

Nobre, C.A., Sampaio, G., Borma, L., Castilla-Rubio, J.C., Silva, J.S., Cardoso, M. (2016). "Land-use and climate change risks in the Amazon and the need of a novel sustainable development paradigm." Proceedings of the National Academy of Sciences, 113(39), 10759-10768.

Oliveira Júnior JF, Delgado RC, Gois G, Lannes A, Dias FO, Souza JCS, Souza M (2014) Análise da Precipitação e sua Relação com Sistemas Meteorológicos em Seropédica. Rio de Janeiro. Floresta e Ambiente 21(2):140–149. doi:10.4322/floram.2014.030

Oliveira LFC, Fioreze AP, Medeiros AMM, Silva MAS (2010) Comparação de Metodologias de Preenchimento de Falhas de Séries Históricas de Precipitação Pluvial Anual. Revista Brasileira de Engenharia Agrícola e Ambiental 14(11):1186–1192. doi:10.1590/S1415-43662010001100008

Pansera, W.L., Gomes, B.M., Vilas Boas, M.A., Sampaio, S.C., Melo, E.L., Queiroz, M.M.F. (2015) Regionalization of monthly precipitation values in the state of Paraná (Brazil) by using multivariate clustering algorithms. Irriga, Botucatu, v.20, n.3, p.473-489.

Pham, D. T., Dimov, S. S., & Nguyen, C. D. (2005). Selection of K in K-means clustering. Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science, 219(1), 103-119. https://doi.org/10.1243/095440605X8298

Ramos MC (2001) Divisive and hierarchical clustering techiniques to analysis variability of rainfall distribuition in a Mediterranean region. Atmos Res 57(2):123–138. doi:10.1016/S0169-8095(01)00065-5

Rao, V. B., Franchito, S. H., Santo, C. M., & Gan, M. A. (2016). An update on the rainfall characteristics of Brazil: seasonal variations and trends in 1979-2011. International Journal of Climatology, 36(1), 291-302. doi: 10.1002/joc.4345

Rodrigues JA, Santos Filho J, Chaves LM (2013) Funções densidade de probabilidade para a estimativa da precipitação mensal. Semina: Ciências Exatas e Tecnológicas 34(1):03–08. doi:10.5433/1679-0375.2013v34n1p03

Roziane, S. S., Gilbeto, C. S., Robson, A. O., and Gabriel, M. A. (2012). Homogeneidade de séries climatológicas em Minas Gerais. Revista Brasileira Engenharia Agrícola e Ambiental 16, 1338–1345. doi: 10.1590/S1415-43662012001200011

Santos CAG, Freire PKMM, Silva RM, Akrami SA (2019) Hybrid wavelet neural network approach for daily inflow forecasting using tropical rainfall measuring mission data. J Hydrol Eng 24:04018062. https://doi.org/10.1061/(asce)he.1943-5584.0001725

Santos, S. R. Q.; Braga, C. C.; Campos, T. L. O. B.; Brito, J. I. B.; Santos, A. P. P. (2014) Variabilidade da precipitação no Estado do Pará por meio de análise em componentes principais. Revista Brasileira de

Geografia Física, v.7, n.3, p.615- 627.

Silva I D M, Medeiros DM, Sakamoto MS, Leal JBV, Mendes D and Ambrizzi T (2022) Evaluating homogeneity and trends in extreme daily precipitation indices in a semiarid region of Brazil. Front. Earth Sci. 10:1071128. doi: 10.3389/feart.2022.1071128

Soccol OJ, Cardoso CO, Miquelluti DJ (2010) Análise da precipitação mensal provável para o município de Lages, SC. Revista Brasileira de Engenharia Agrícola e Ambiental 14(6):569–574. doi:10.1590/S1415-43662010000600001

Souza, A., Abreu, M.C,, De Oliveira-Júnior, J. F., Aristone, F., Fernandes, W. A., Aviv-Sharon, E., Graf, R. (2022) Climate Regionalization in Mato Grosso do Sul: a Combination of Hierarchical and Non-hierarchical Clustering Analyses Based on Precipitation and Temperature. Brazilian Archives of Biology and Technology (Online), v. 65, p. 1-10.

Teodoro PE, Corrêa CCG, Torres FE, Oliveira-Junior JF, Silva Junior CA, Gois G, Delgado RC (2015) Analysis of the occurrence of wet and drought periods using standardized precipitation index in Mato Grosso do Sul State, Brazil. J Agron 14(2):80–86. doi:10.3923/ja.2015.80.86

Unal Y, Kindap T, Karaca M (2003) Redefining the climate zones of Turkey using cluster analysis. Int J Climatol 23(9):1045–1055. doi:10.1002/joc.910

Vieira JP, Souza MJH, Teixeira JM, Carvalho FP (2010) Estudo da precipitação mensal durante a estação chuvosa em Diamantina, Minas Gerais. Revista Brasileira de Engenharia Agrícola e Ambiental 14(7):762–767. doi: 10.1590/S1415-43662010000700012

Luiz-Silva, W., Oscar-Júnior, A. C., Cavalcanti, I. F. A., & Treistman, F. (2021). An overview of precipitation climatology in Brazil: space-time variability of frequency and intensity associated with atmospheric systems. Hydrological Sciences Journal, 66(2), 289-308. doi: 10.1080/02626667.2020.1863969

Author's Affiliation

Souza, Amaury - Federal University of Mato Grosso do Sul, Campo Grande (MS), Brazil.
Medeiros, E.S. - Federal University of Lavras, Lavras (MG), Brazil

Authors' Contribution

Souza, Amaury - The author contributed to the elaboration, realization and manipulation of data and writing.
Medeiros, E.S. - The author contributed to the elaboration, realization and manipulation of data and writing.

Editors in Charge

Alexandra Maria Oliveira
Alexandre Queiroz Pereira