Revista Brasileira de Recursos Hídricos
Brazilian Journal of Water Resources
Versão On-line ISSN 2318-0331
RBRH, Porto Alegre, v. 27, e25, 2022
Scientific/Technical Article

https://doi.org/10.1590/2318-0331.272220220051

Use of self organizing map to identify precipitation patterns and assess their impact on hydrographic basins in Brazil

Uso de mapa auto-organizado para identificar padrões de precipitação e avaliar seu impacto em bacias hidrográficas no Brasil

Yoshiaki Sakagami[1], Vinicius Nunes Folganes[1], Cesar Alberto Penz[1], Murilo Reolon Scuzziato[1] & Fabrício Yutaka Kuwabata Takigawa[1]

[1]Instituto Federal de Santa Catarina, Florianópolis, SC, Brasil
E-mails: yoshi@ifsc.edu.br (YS), vinicius.nf@aluno.ifsc.edu.br (VNF), cesar.penz@ifsc.edu.br (CAP), murilo.scuzziato@ifsc.edu.br (MRS), takigawa@ifsc.edu.br (FYKT)

Received: June 16, 2022 – Revised: August 26, 2022 – Accepted: September 21, 2022

ABSTRACT

In this study, we used neural networks known as self-organizing maps (SOMs) to identify clusters of spatial synoptic precipitation patterns. These clusters were compared with the precipitation regime of the ten main hydrographic sub-basins in Brazil. Sixty years of daily precipitation data obtained from over 389 weather station in Brazil were used as input data for the SOMs, with a number of six clusters being prescribed as the optimal number according to the elbow and silhouette methods. The six precipitation patterns identified by the SOMs reflect the typical synoptic conditions associated mainly with the cold frontal systems (CF), South American Monsoon System (SAMS) and Inter-tropical Convergence Zone (ITCZ). In conclusion, SOMs perform well using interpolated precipitation data as the input data to identify synoptic precipitation patterns, which could be used to monitor the spatial distribution of precipitation, which affects the hydrographic basins in Brazil and hence hydropower plant performance.

Keywords: Self-organizing map; Precipitation regime; Spatial-Temporal Pattern.

RESUMO

Neste estudo, utilizamos uma técnica de redes neurais conhecida como Mapa Auto Organizável (SOM, em inglês), para identificar e agrupar padrões espaço-temporal de precipitação. Esses grupos foram comparados com o regime de precipitação de dez principais sub-bacias hidrográficas do Brasil. Sessenta anos de dados de precipitação diária foram obtidos de uma rede de 389 estações meteorológicas no Brasil onde foram utilizados como entrada de dados para o SOM, com um número de seis grupos pré-definidos através dos métodos de Elbow e Silhoette. Os seis padrões espaço-temporal de precipitação foram identificados pelo SOM o qual conseguiu identificar condições sinóticas típicas como o Sistema de Frentes Frias (FF) Sistema de Monção da América do Sul (SMAS) e a Zona de Convergência Intertropical (ZCIT). Assim, concluímos que o SOM teve um bom desempenho utilizando dados interpolados de precipitação para identificar padrões espaço-temporal de precipitação, sendo que essa técnica de SOM pode ser promissora para monitorar os padrões espaço-temporal de precipitação que causam consideráveis impactos nas bacias hidrográficas e usinas hidroelétricas no Brasil.

Palavras chave: Mapa auto organizável; Regime de precipitação; Padrões Espaço-Temporal.

INTRODUCTION

Water resources play a key role in the electrical power sector in Brazil, since hydropower plants account for 62.5% of the energy generation capacity (Agência Nacional de Energia Elétrica, 2022). The Brazilian Independent System Operator (ISO) and Market Operator (MO) started to obtain hourly energy dispatch and price data in 2020 and 2021, respectively, improving the representation of energy supply and demand (Câmara de Comercialização de Energia Elétrica, 2022). In this regard, power generation planning and energy prices are sensitive to the amount of precipitation and thus a good understanding of spatial and temporal precipitation distribution in hydrographic basins is required. Hence, precipitation is one of the most important variables in the energy sector in Brazil, which is usually monitored by satellites, radar, rain gauges and numerical atmospheric models (Centro de Previsão de Tempo e Estudos Climáticos, 2022). Efficient precipitation forecasting is also important to market players, aiding planning and decision-making in the energy spot market. However, all data from weather monitoring networks need to be processed and analyzed using several approaches, including numerical map models, satellite images and synoptic charts. In this way, it is possible to identify the spatial and temporal precipitation patterns and associate them with the weather systems. Several studies on the spatial and temporal precipitation regimes over South America (SA) have been investigated from interdecadal to inter-annual climate variability, such as Pacific Decadal Oscillation (PDO) (Oliveira-Júnior et al., 2018; Prado et al., 2021), El Nino and the South America Oscillation (ENSO) (Cavalcanti et al., 2009), as well as intra-annual variability, for instance, the Madden-Julian Oscillation (MJO) (Grimm, 2019). Multivariate statistical techniques, e.g., principal component analysis (PCA), are commonly used in climatology to identify patterns in atmospheric processes (Cavalcanti et al.,2009; Grimm, 2011), but they are not practical for identifying short-term weather patterns, as they are statistical tools used for climate analysis. Objective analysis is used to automatically identify meteorological systems, such as cold fronts (Lagerquist et al., 2020), the Inter-tropical convergence zone (ITCZ) (Hohenegger & Jakob, 2020) and the South Atlantic Convergence Zone (SACZ) (Zilli & Carvalho, 2021). However, this approach is limited to identifying a specific weather system in a strict domain, which is dependent on thresholds that may not be valid for all cases (Grotjahn et al., 2016). In addition, a combination of objective analysis processes would be necessary since the hydrographic basins are located within different climate precipitation regimes in SA (Reboita et al., 2010). Methods based on unsupervised neural networks have been used as an alternative approach to identify the multi-dimensional patterns, based on cluster algorithms (Xu & Tian, 2015), such as self-organizing maps (Kohonen, 2013), the Ward method (Ward, 1963) and K-means (Lloyd, 1982). Cluster analysis has been applied in meteorological studies to classify regions with similar climate conditions (Netzel & Stepinski, 2016) and identify weather patterns from two-dimensional maps of surface air pressure (Sheridan & Lee, 2011), sea surface temperature (Johnson, 2013), precipitation (Santos et al., 2019) and geopotential height (Liu et al., 2016). In Brazil, studies involving clustering-based procedures are usually related to improving our general understanding of the weather and climate conditions (Anunciação et al., 2014; Lyra et al., 2014; Brito et al., 2017; Ferreira & Reboita, 2022) and more applications focused on the Brazilian energy sector are needed. In this study, we used the type of neural network known as the self-organizing map (SOM) to identify precipitation patterns and compare them with the precipitation regime in the ten main hydrographic basins in Brazil. Firstly, 60 years of daily precipitation data from the Brazilian network was processed and interpolated to sub-basins and climatological reference stations. Secondly, two-dimensional maps of interpolated data were grouped using the SOM technique applying an optimal number of six clusters. Lastly, the precipitation anomaly for each sub-basins was correlated with the total annual number of events for each cluster.

METHODS

Precipitation data

We processed 60 years of daily accumulated precipitation data from conventional network weather stations from 1961 to 2020. The database can be found at the Banco Nacional de Dados Meteorologicos (BNDMET) of the Instituto Nacional de Meteorologia (INMET) and Departamento de Controle do Espaco Aereo (DECEA) (Departamento de Controle do Espaço Aéreo, 2020). There is a total of 784 conventional weather stations and 633 automatic weather stations in the network, but the data available is not constant over the whole period, as seen in Figure 1.

In the year 1961, the conventional network increased from around 40 to 300 weather stations (Figure 1). The data available daily varied from 108 to 389 in this period, and although there are many gaps in this time series, we filled and homogenized using interpolation method as described in Section 2.2. The automatic weather stations started operating in 2000 and presented many inconsistencies in the data, which were not used in this study. There are also several precipitation datasets from satellite and reanalysis data that were not considered in this work, due to discrepancies between the datasets regarding the variation in annual precipitation (Sun et al., 2018; Reis et al., 2020).

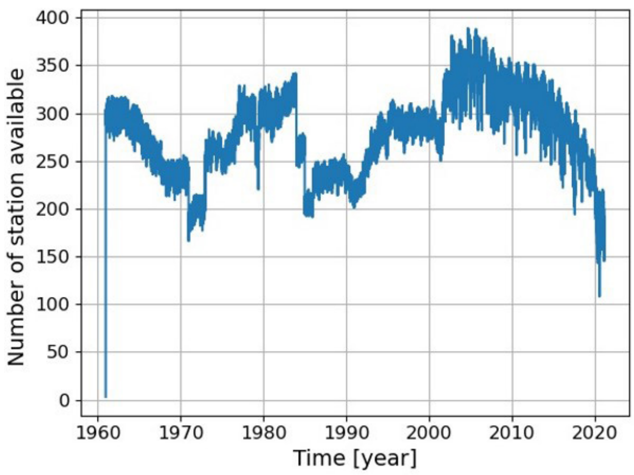


Figure 1. Number of conventional precipitation measurements in the BNDMET database from 1961 to 2020 (Departamento de Controle do Espaço Aéreo, 2020).

Interpolation method

The SOM algorithm requires the same number of lines and rows as inputs to cluster the precipitation data, but the number of weather stations available for each day in the network is not the same over the 60-year period (Figure 1). Thus, the data were spatially extrapolated and then interpolated at 317 selected positions, which are the positions of the convectional stations used for normal climates in Brazil, as shown in Figure 2a (black dots). There are several spatial interpolation methods in the literature and those most commonly used include the deterministic spatial interpolation methods and geostatistical approaches (Ly et al., 2013). In this study, we used the quadratic radial base functions (RBF) method (Buhmann, 2003), which has the advantage that it fits a surface (spatial precipitation patterns) through the daily precipitation data, while minimizing the total curvature of the surface (Wang et al., 2014). It also performed satisfactorily for days with reduced data available, through a smoothed extrapolation, and when difficulty is encountered when using geostatistical methods (Krige, 1951) to extrapolate for data-sparse regions, according to Ahrens (2006). Negative values sometimes occur when the data are not well distributed and negative values are forced to zero.

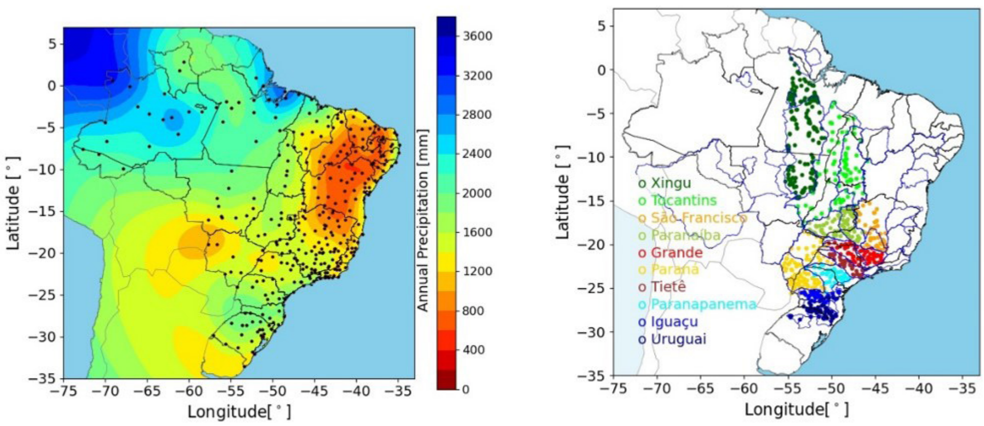
The RBF interpolation method was also used to estimate the spatial average of daily precipitation in the ten sub-basins. The rain gauge network is operated by Agência Nacional de Águas (ANA) and the stations are strategically installed to cover the ten main sub-basins of the largest hydro-power installations in Brazil. Figure 2b shows the rain gauge network, where the dots represent the set of rain gauges located in the following sub-basins: Amazonas/Xingu (darkgreen), Tocantins (green), Sao Francisco (orange), Paranaíba (yellow green), Grande (red), Parana (gold), Tiête (brown), Paranapanema (light blue), Iguassu (blue) and Uruguai (dark blue). The Operador Nacional do Sistema Elétrico (ONS) uses this network precipitation data as input data for the hydrological model (SMAP) to estimate the energy price (Operador Nacional do Sistema Elétrico, 2018).

Self-Organizing Map

Self-organizing map (SOM) is considered a special class of artificial neural network, based on unsupervised training using only two layers (Kohonen, 2013). They are commonly used to identify patterns of two-dimensional arrays of maps for a broad range of scientific fields (Oja et al., 2003). SOM allows the clusters to be separated and analyzed as a continuous pattern in space. Thus, transitions, as well as extreme patterns, can be identified rather than only the primary patterns, which is generally the case in principal component analysis. However, with SOM, it is difficult to define the best number of clusters (or neurons) and the size of the input array, which needs to be predetermined and there are no recommendations for this (Grotjahn et al., 2016). In this study, the input was the 60 years of daily interpolated data for 317 fixed positions at the locations of the convectional stations, as described above (Figure 2a). The 317 columns contain the precipitation values for each position and the rows contain the time series of daily precipitation from 1961 to 2020. The optimal number of neurons is determined by other methods, as described below.

Number of cluster

The optimum number of clusters is determined by the elbow method (Thorndike, 1953), where the Euclidian distance within each cluster is minimized, and also by the silhouette method (Rousseeuw, 1987), which considers the minimum distance within the cluster and the maximum distance between the clusters. Both methods involve the use of the non-supervised neural network K-means (Lloyd, 1982). Figure 3a shows the elbow curve, where the mean squared error within the cluster decreases inversely to the number of clusters. The elbow kink is not clear as the curve varies smoothly. The second method complements this analysis by way of the silhouette coefficient, where the coefficient decreases with the



(a) (b)

Figure 2. Annual average precipitation and the position of the conventional stations in Brazil from INMET (Departamento de Controle do Espaço Aéreo, 2020) (a) and ANA/ONS (Operador Nacional do Sistema Elétrico, 2018) rain gauge network with ten sub-basins, where most hydropower operations are located (b).

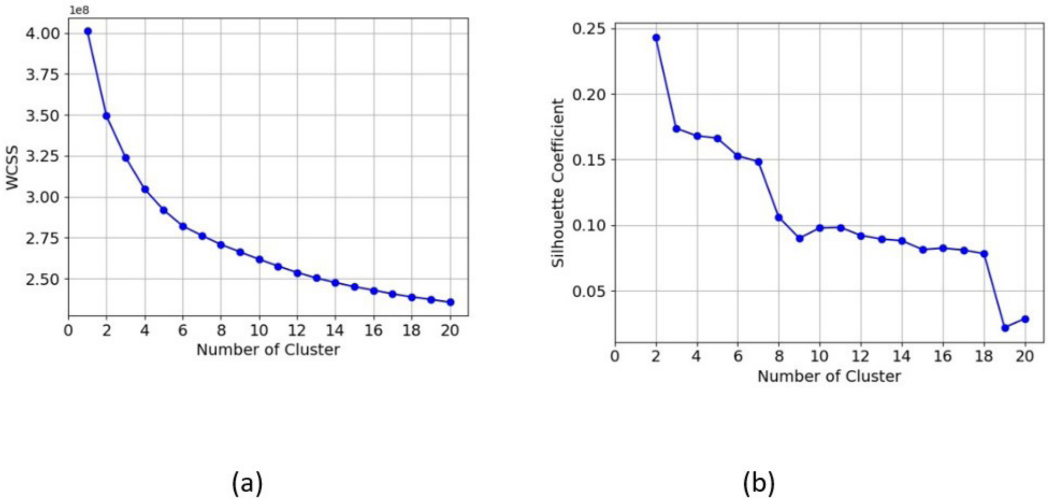


Figure 3. Elbow method: Within-Cluster-Sum-of-Squares (WCSS) (a) and silhouette coefficient (b).

number of clusters as the clusters approach each other. Figure 3b shows two significant variations in the silhouette coefficient, which suggests that a main group of clusters ranges from 3 to 7 clusters, and the second set of clusters varies from 9 to 18 clusters. This work was limited to analyzing only the main group, with six clusters, where the self-organizing map was fixed at 3x2 neurons.

RESULTS AND DISCUSSION

Precipitation patterns

The self-organizing map separated 60 years of daily interpolated precipitation data, obtained from 317 stations, into six clusters, as described in Section 2. The precipitation of each cluster was averaged only for the stations with non-zero precipitation, where the synoptic precipitation patterns can be viewed on the six maps in Figure 4a and the respective seasonal regimes in Figure 4b.

The first (C1) and second (C2) clusters reveal the spatial-temporal precipitation patterns associated with cold fronts that occur in the southern region of Brazil (Rodrigues et al., 2004). Although the first and second clusters are spatially similar, they differ from their seasonal regimes. Cluster C1 occurs mostly from spring to summer and is associated with mesoscale convective complexes (MMC) and instability lines (IL) developed ahead of the cold fronts (Figure 4b). The second cluster is related to the cold fronts that often occurs in the winter to early spring season and is associated with polar jets (PJ) and extra-tropical cyclones. These results are in agreement with reports in the literature that could fronts can cause precipitation when passing through the south of Brazil (C2) or induce instability lines that cause strong precipitation even when the cold fronts is at a distance (C1) (Reboita et al., 2010). The average number of C1 frontal systems occurring weekly in the south of Brazil is in agreement with Britto & Saraiva (2001) and Rodrigues et al. (2004). The cluster C3, seen in Figure 4a, is the non-precipitation condition that covers almost the whole of Brazil, characterized by the dry season from April to October (Figure 4b), and is the prevailing condition, with an annual average occurrence of 113 events during the 60-year period (Figure 5a). A few areas of precipitation can be observed along the northeast coast, caused by local sea breezes and trade winds (convergence zone), north Amazon and the south of Rio Grande do Sul state, caused by cold fronts in winter. One outlier of C3 was observed in 2009 with only 76 events during moderate El Niño conditions. The prevailing precipitation condition in summer, from October to March, is represented by cluster C4, which is associated with the South American Monsoon System (SAMS) that covers a large area of precipitation in central Brazil, with a transversal area oriented from northwest to southeast (shown in Figure 4a), as described by Carvalho et al. (2004). This area matches Region 5 of the adapted Climate Atlas of South America, for which the major meteorological systems that impact the precipitation regime in this area are described by Reboita et al. (2010). This is the pattern with the second most frequent precipitation events, with an annual average of 89 events per year (Figure 5a). Two outliers were observed in 1963 with only 63 events and in 1983 during moderate and strong El Nino conditions, respectively. Cluster C5, with the lowest annual occurrence (34 events per year), shows a major area of precipitation located in the southwest part of the Amazon region that occurs mostly in October, November and March. The absence of the cold fronts (possibly caused by blocking) in the south of Brazil and the absence of ITCZ in the north indicate that the precipitation is only associated with local convection caused by radiative heating in the Amazon region. The last precipitation pattern, represented by cluster C6, is clearly associated with the ITCZ (Carvalho & Oyama, 2013), where the center of the area of precipitation is located in the northeast part of the north region of Brazil (Figure 4a). This cluster is frequently present from March to May and has the highest number of events in April, with an average of 14 events per year (Figures 4b and 5b). There are no outliers but there is a high variability, with a standard deviation of ±12 days per year.

Sub-basin correlation

The annual spatial and seasonal average precipitation data for each sub-basin are shown in Figures 6a and 6b, respectively, where similar annual median precipitation ranges of 1397 to

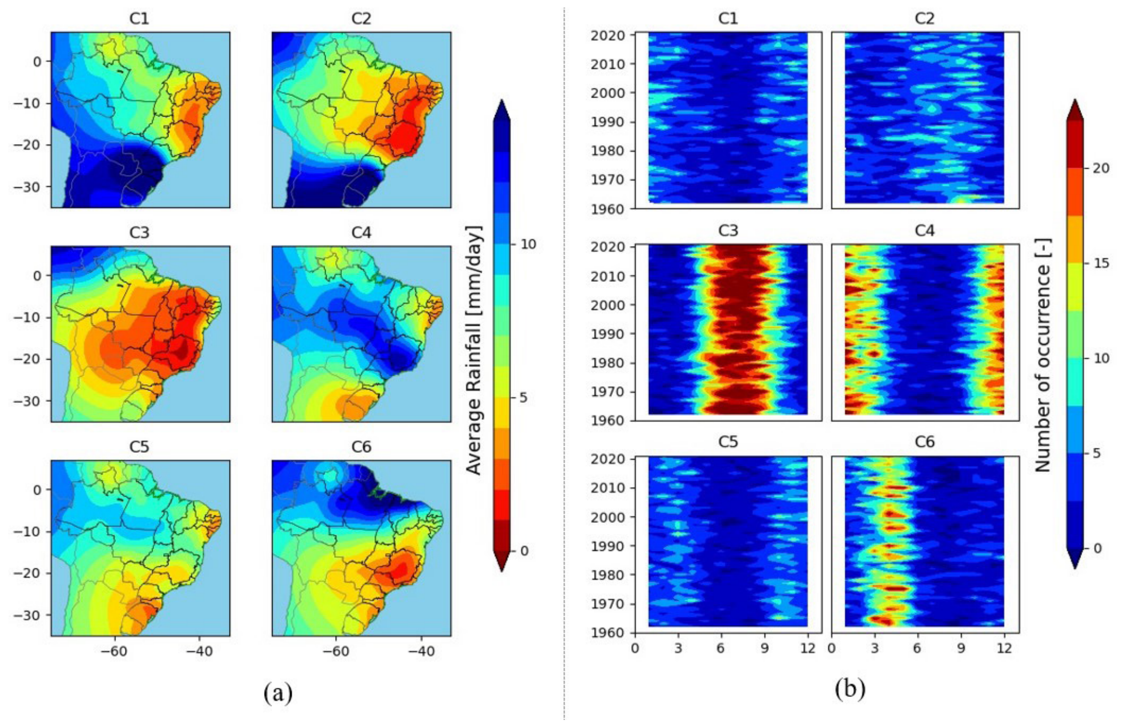


Figure 4. Average precipitation of each cluster (a) and the number of occurrences by year vs. month (b).

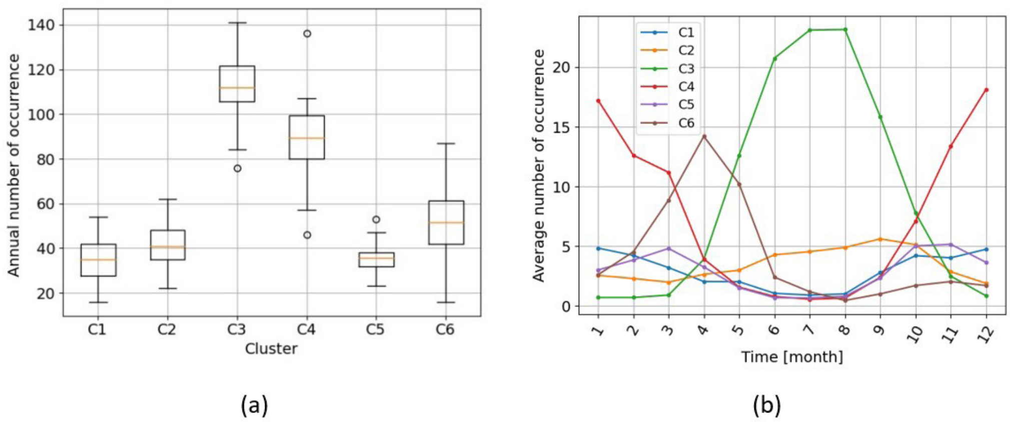


Figure 5. Annual number of occurrences of each cluster (a) and monthly average number of occurrences of each cluster (b).

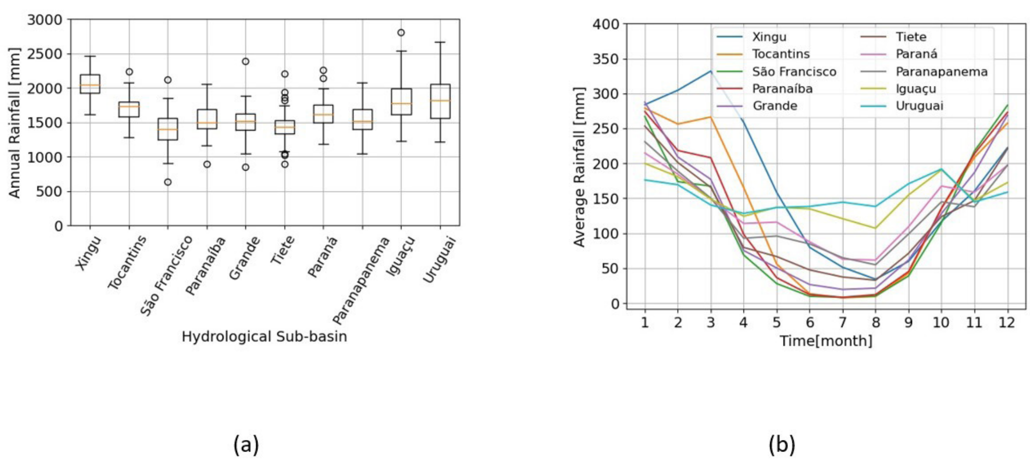


Figure 6. Annual average precipitation for each of the ten main sub-basins in Brazil (a) and the seasonal precipitation (b).

1825 mm/year can be observed. Iguassu and Uruguai present the strongest annual variability, where the difference between the upper and lower quartile reaches 379 mm and 492 mm respectively. Tietê has the lowest annual variability, with 192 mm, but there is a significant number of outliers (7 events).

The Tocantins, Sao Francisco, Paranaíba, Grande and Tietê sub-basins have a high season variability, with the rainy season in summer and dry season in winter, influenced by the seasonal regime of cluster C4 (Figure 5b). The Amazonas/Xingu sub-basin, which has a sensitive runoff to land use and cover (Cruz et al., 2022), has also a high season variability, but it is mostly affected by the cluster C6 associated with the ITCZ. The highest monthly average precipitation occurs in March/April and the lowest in August/September in agreement with Santos et al. (2019). The sub-basins located in the south are affected by clusters C1 and C2 (Figure 5b), where the Paraná and Paranapanema sub-basins have a moderate seasonal variability, and Iguassu and Uruguai have almost the same monthly average precipitation throughout the year. The impact of the cluster on the annual precipitation of each sub-basin is analyzed through the correlation between the variation of the number of occurrences of each cluster and the annual precipitation anomaly (Figure 7).

Clusters C1 and C2 associated with the cold fronts have a positive correlation with the sub-basins located in the south of Brazil. The correlations of C1, associated with the frontal system with summer instabilities, show the strongest relationship with

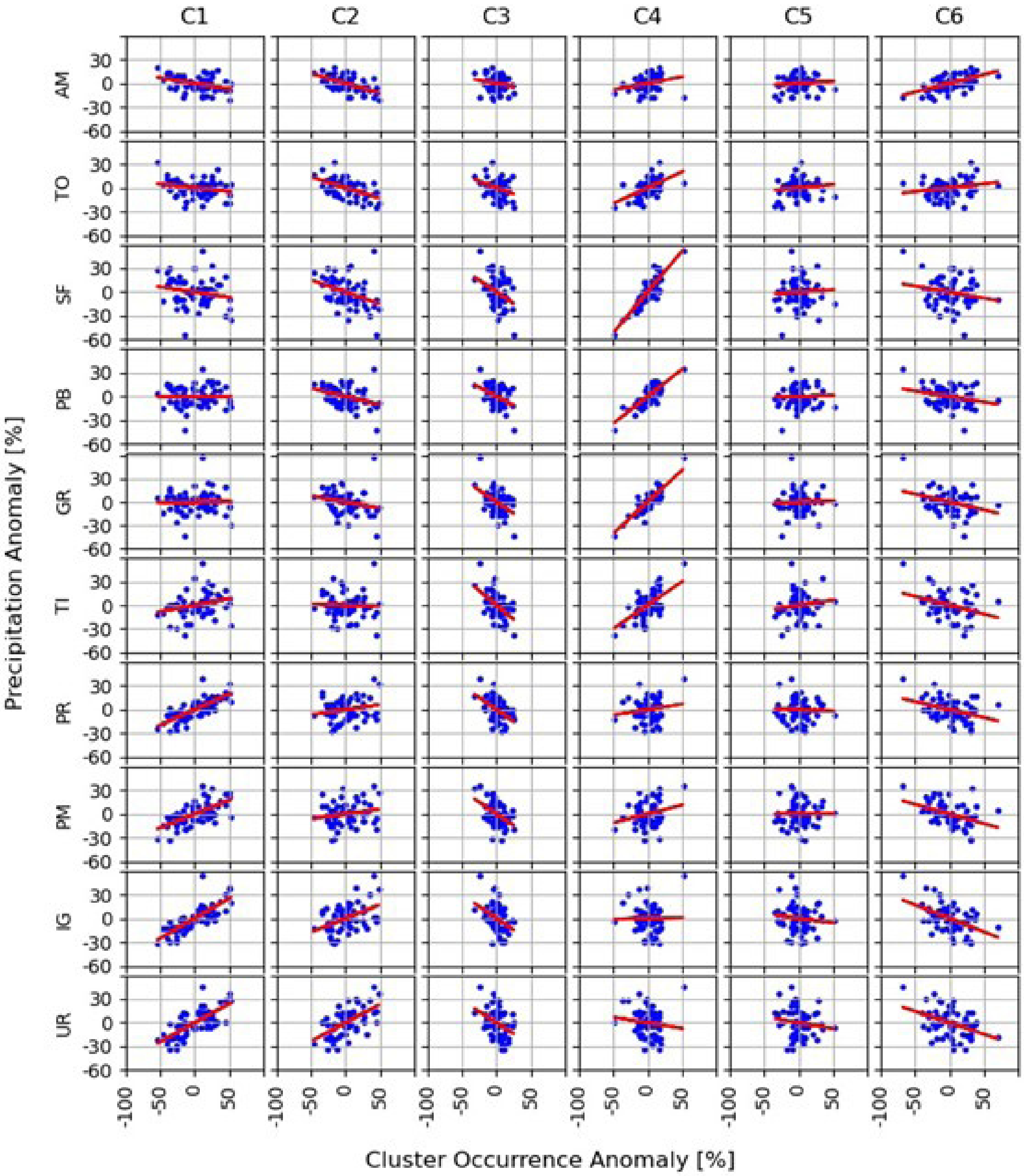


Figure 7. Correlation between the variation of cluster occurrence vs. annual precipitation anomaly of each sub-basin: Amazonas/Xingu (AM), Tocantins (TO), São Francisco (SF), Paranaíba (PB), Grande (GR), Tiete (TI), Paraná (PR), Paranapanema (PM), Iguaçu (IG) , Uruguai (UR).

Table 1. Correlation between annual precipitation and occurrence of each cluster.

Sub-basin	C1	C2	C3	C4	C5	C6
Xingu	-0.39	**-0.57**	-0.18	0.27	0.09	**0.57**
Tocantins	-0.23	**-0.54**	-0.34	**0.57**	0.13	0.22
São Francisco	-0.19	-0.37	-0.37	**0.88**	0.05	-0.21
Parnaíba	0.01	-0.36	-0.41	**0.84**	0.04	-0.27
Grande	0.02	-0.23	-0.50	**0.83**	0.04	-0.27
Tietê	0.28	-0.05	-0.53	**0.59**	0.14	-0.35
Paraná	**0.76**	0.20	-0.5	0.16	-0.04	-0.36
Paranapanema	**0.65**	0.19	-0.47	0.25	0.01	**-0.42**
Iguaçú	**0.79**	**0.46**	-0.40	0.02	-0.11	**-0.50**
Uruguai	**0.74**	**0.60**	-0.37	-0.13	-0.15	**-0.41**

sub-basins of Paraná, Paranapanema, Iguassu and Uruguai, as shown in Table 1.

Cluster C2, associated with the cold fronts in winter and spring, has a satisfactory correlation with the sub-basins of Iguassu (0.46) and Uruguai (0.60). The sub-basins located above the southeast region of Brazil show low correlations, except cluster C2, which has with a negative correlation of -0.54 with Tocantins. Cluster C3, which represents dry conditions, displays moderated negative correlations with all sub-basins (varying from -0.34 to -0.53). In contrast to C1 and C2, cluster C4 has a better correlation with the sub-basins in the southeast and central regions of Brazil, which are influenced by the South American Monsoon System. Table 1 shows strong positive correlations for Tocantins, Sao Francisco, Paranaíba and Grande and Tietê, and low correlations for the sub-basins located in the south. Clusters C5 and C6 (ITCZ) are associated with precipitation in the north region of Brazil and have very low impact on the precipitation anomaly in most of the sub-basins in this study, except for cluster C6 with positive correlation of 0.57 for the Amazonas/Xingu sub-basin, and the Paranapanema, Iguassu and Uruguai sub-basins, where a substantial negative correlation is observed.

CONCLUSION

The non-supervised neural network SOM was able to cluster six synoptic precipitation patterns in Brazil. The clusters are consistent with the main synoptic precipitation features associated with frontal systems (C1 and C2), dry seasonal conditions (C3), South American Monsoon System (C4), Bolivian High with convection in the Amazon region through radiative heating (C5) and the Inter-Tropical Convergence Zone (C6). The sub-basins in the south of Brazil present a strong correlation with clusters C1 and C2 associated with frontal systems, whereas the sub-basins above the southeast region have a strong correlation with cluster C4, associated with the South American Monsoon System. The breakdown of two clusters associated with frontal systems aids an understanding of the strong correlation between prefrontal instabilities in the summer (C1) and the annual precipitation anomaly in sub-basins in the south of Brazil. The negative correlations between cluster C2 and the Amazonas/Xingu/Tocantins sub-basins and between cluster C6 and the Paranapanema/Iguassu/Uruguai sub-basins shows the complexity of the weather behavior in Brazil. The variability of the annual precipitation in hydrological basins with spatial precipitation patterns in the north region can impact sub-basins in the south and vice-versa.

This study demonstrates how the non-supervised neural network algorithm known as the self-organizing map can aid the identification of synoptic precipitation patterns, and provide new look to analyze the climatology of precipitation based on machine learning instead of traditional statistics methods of monthly averages. For water resource applications, these machine learning tools can collaborate to yield new ways to monitor and predict the spatial distribution of precipitation, as we display here by investigating the impacts of the hydrographic basins in Brazil and consequently the performance of hydroelectric installations.

ACKNOWLEDGEMENTS

The authors wish to acknowledge the support of Embrapii-IFSC under the project number PFLN-1911-005, funded by Central Comercializadora de Energia LTDA Company and Embrapii-IFSC.

REFERENCES

Agência Nacional de Energia Elétrica – ANEEL. (2022). *Sistema de informações de geração da Aneel - SIGA*. Retrieved in 2022, January 16, from https://bit.ly/2IGf4Q0

Ahrens, B. (2006). Distance in spatial interpolation of daily rain gauge data. *Hydrology and Earth System Sciences*, *10*, 197-208.

Anunciação, Y. M. T., Walde, D. H. G., & da Rocha, R. P. (2014). Observed summer weather regimes and associated extreme precipitation over Distrito Federal, west-central Brazil. *Environmental Earth Sciences*, *72*, 4835-4848.

Brito, T. T., Oliveira-Júnior, J. F., Lyra, G. B., Gois, G., & Zeri, M. (2017). Multivariate analysis applied to monthly rainfall over Rio de Janeiro state, Brazil. *Meteorology and Atmospheric Physics*, *129*, 469-478.

Britto, F., & Saraiva, J. M. (2001). Estudo da precipitação associada aos sistemas frontais na cidade do Rio Grande na estação do inverno. *Geografia em Atos (Online)*, *1*, 111-123.

Buhmann, M. D. (2003). *Radial basis functions: theory and implementations. cambridge monographs on applied and computational mathematics*. Cambridge: Cambridge University Press.

Câmara de Comercialização de Energia Elétrica – CCEE, (2022). Retrieved in 2022, January 16, from https://www.ccee.org.br/

Carvalho, L. M. V., Jones, C., & Liebmann, B. (2004). The south Atlantic convergence zone: Intensity, form, persistence, and relationships with intraseasonal to interannual activity and extreme rainfall. *Journal of Climate*, *17*, 88-108.

Carvalho, M. A. V. d., & Oyama, M. D. (2013). Variabilidade da largura e intensidade da Zona de Convergência Intertropical Atlântica: aspectos observacionais. *Revista Brasileira de Meteorologia*, *28*, 305-316.

Cavalcanti, I. F. A., Ferreira, N. J., Dias, M. A. F., & Justi, M. G. A. (2009). *Tempo e Clima no Brasil*. São Paulo: Oficina de Textos.

Centro de Previsão de Tempo e Estudos Climáticos – CPTEC, & Instituto Nacional de Pesquisas Espaciais – INPE. (2022). Retrieved in 2022, January 16, from https://www.cptec.inpe.br/

Cruz, J. S., Blanco, C. J. C., & de Oliveira Júnior, J. F. (2022). Modeling of land use and land cover change dynamics for future projection of the Amazon number curve. *The Science of the Total Environment*, *811*, 152348.

Departamento de Controle do Espaço Aéreo – DECEA, & Instituto Nacional de Meteorologia – INMET. (2020). *Banco nacional de dados meteorológicos*. Retrieved in 2022, January 16, from https://bndmet.decea.gov.br/sobre

Ferreira, G. W. S., & Reboita, M. S. (2022). A New Look into the South America Precipitation Regimes: observation and Forecast. *Atmosphere*, *13*, 873.

Grimm, A. M. (2011). Interannual climate variability in South America: impacts on seasonal precipitation, extreme events, and possible effects of climate change. *Stochastic Environmental Research and Risk Assessment*, *25*, 537-554.

Grimm, A. M. (2019). Madden–Julian Oscillation impacts on South American summer monsoon season: precipitation anomalies, extreme events, teleconnections, and role in the MJO cycle. *Climate Dynamics*, *53*, 907-932.

Grotjahn, R., Black, R., Leung, R., Wehner, M.F., Barlow, M., Bosilovich, M., Gershunov A., Gutowski, W.J., Gyakum, J.R., Katz, R.W., Lee, Y.Y., Lim, Y.K., & Prabhat (2016). North american extreme temperature events and related large scale meteorological patterns: a review of statistical methods, dynamics, modeling, and trends. *Climate Dynamics*, *46*, 1151-1184.

Hohenegger, C., & Jakob, C. (2020). A relationship between ITCZ organization and subtropical humidity. *Geophysical Research Letters*, *47* (16), e2020GL088515.

Johnson, N. C. (2013). How many enso flavors can we distinguish? *Journal of Climate*, *26*, 4816-4827.

Kohonen, T. (2013). Essentials of the self-organizing map. *Neural Networks*, *37*, 52-65.

Krige, D. (1951). A statistical approach to some basic mine valuation problems on the Witwatersrand. *Journal of the Southern African Institute of Mining and Metallurgy*, *52*, 119-139.

Lagerquist, R., Allen, J. T., & McGovern, A. (2020). Climatology and variability of warm and cold fronts over North America from 1979 to 2018. *Journal of Climate*, *33*(15), 6531-6554.

Liu, W., Wang, L., Chen, D., Tu, K., Ruan, C., & Hu, Z. (2016). Large-scale circulation classification and its links to observed precipitation in the eastern and central Tibetan plateau. *Climate Dynamics*, *46*, 3481-3497.

Lloyd, S. (1982). Least squares quantization in pcm. *IEEE Transactions on Information Theory*, *28*, 129-137.

Ly, S., Charles, C., & Degré, A. (2013). Different methods for spatial interpolation of rainfall data for operational hydrology and hydrological modeling at watershed scale. a review. *Biotechnologie, Agronomie, Société et Environnement*, *17*, 392-406.

Lyra, G. B., Oliveira-Junior, J. F., & Zeri, M. (2014). Cluster analysis applied to the spatial and temporal variability of monthly rainfall in Alagoas state, northeast of Brazil. *International Journal of Climatology*, *34*, 3546-3558.

Netzel, P., & Stepinski, T. (2016). On using a clustering approach for global climate classification. *Journal of Climate*, *29*(9), 160222131641003.

Oja, M., Kaski, S., & Kohonen, T. (2003). Bibliography of self-organizing map som papers: 1998-2001 addendum. *Neural Computing Surveys*, *3*, 1-156.

Oliveira-Júnior, J. F., de Gois, G., de Bodas Terassi, P. M., da Silva Junior, C. A., Blanco, C. J. C., Sobral, B. S., & Gasparini, K. A. C. (2018). Drought severity based on the SPI index and its relation to the ENSO and PDO climatic variability modes in the regions North and Northwest of the State of Rio de Janeiro - Brazil. *Atmospheric Research*, *212*, 91-105.

Operador Nacional do Sistema Elétrico – ONS. (2018). *Generalização do modelo SMAP/ONS para previsão de vazões no âmbito do SIN* (Nota Técnica ONS 0056/2018). Brasília: ONS.

Prado, L. F., Wainer, I., Yokoyama, E., Khodri, M., & Garnier, J. (2021). Changes in summer precipitation variability in central Brazil over the past eight decades. *International Journal of Climatology*, *41*, 4171-4186.

Reboita, M. S., Gan, M. A., Rocha, R. P. d., & Ambrizzi, T. (2010). Regimes de precipitação na América do Sul: uma revisão bibliográfica. *Revista Brasileira de Meteorologia*, *25*, 185-204.

Reis, A. A., Fernandes, W. S., & Ramos, M. H. (2020). Assessing two precipitation data sources at basins of special interest to hydropower production in Brazil. *RBRH*, *25*, e14.

Rodrigues, M. L. G. A., Franco, D., & Sugahara, S. (2004). Climatologia de frentes frias no litoral de Santa Catarina. *Revista Brasileira de GeofAsica*, *22*, 135-151.

Rousseeuw, P. J. (1987). Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, *20*, 53-65.

Santos, V. C., Blanco, C., & Oliveira Júnior, J. F. (2019). Distribution of rainfall probability in the Tapajos River Basin, Amazonia, Brazil. *Revista Ambiente & Água*, *14*(3), e2284.

Sheridan, S. C., & Lee, C. C. (2011). The self-organizing map in synoptic climatological research. *Progress in Physical Geography: Earth and Environment*, *35*, 109-119.

Sun, Q., Miao, C., Duan, Q., Ashouri, H., Sorooshian, S., & Hsu, K. L. (2018). A review of global precipitation data sets: data sources, estimation, and intercomparisons. *Reviews of Geophysics*, *56*, 79-107.

Thorndike, R. L. (1953). Who belongs in the family? *Psychometrika*, *18*, 267-276.

Wang, S., Huang, G. H., Lin, Q. G., Li, Z., Zhang, H., & Fan, Y. R. (2014). Comparison of interpolation methods for estimating spatial distribution of precipitation in Ontario, Canada. *International Journal of Climatology*, *34*, 3745-3751.

Ward, J. H. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, *58*, 236-244.

Xu, D., & Tian, Y. (2015). A comprehensive survey of clustering algorithms. *Annals of Data Science*, *2*, 165-193.

Zilli, M. T., & Carvalho, L. M. V. (2021). Detection and attribution of precipitation trends associated with the poleward shift of the South Atlantic Convergence Zone using CMIP5 simulations. *International Journal of Climatology*, *41*, 3085-3106.

Authors contributions

Yoshiaki Sakagami: Investigation, Conceptualization, Methodology, Software, Formal analysis, Data curation, Writing - original draft.

Vinicius Nunes Folganes: Software, Validation, Formal analysis, Investigation, Writing - review & editing, Visualization.

César Alberto Penz: Investigation, Conceptualization, Methodology.

Murilo Reolon Scuzziatto: Project administration, Resources, Writing - review & editing, Funding acquisition.

Fabrício Yutaka Kuwabata Takigawa: Investigation, Writing - review & editing, Supervision, Project administration, Funding acquisition.

Editor in-Chief: Adilson Pinheiro

Associated Editor: Carlos Henrique Ribeiro Lima