Revista Brasileira de Ciências Ambientais
Brazilian Journal of Environmental Sciences

Seasonal forecasting of inflows to the Carnaubal Reservoir in the Brazilian semi-arid region using artificial neural networks

Previsão sazonal de vazões afluentes ao açude Carnaubal no Semiárido Brasileiro utilizando redes neurais artificiais

Tatiane Lima Batista[1], *Alan Michell Barros Alexandre*[2], *José Kerlly Soares de Araújo*[1], *Francisco de Assis de Souza Filho*[1] *

ABSTRACT

Given the challenges posed by the climatic conditions of semi-arid regions that rely on reservoir-based water management policies to ensure water availability, machine learning methods have been increasingly used as tools for efficient water resources management. This study aimed to calibrate a seasonal streamflow forecasting model for the Carnaubal Reservoir, located in the semi-arid region of Ceará, Brazil, using artificial neural networks (ANN), specifically the Multilayer Perceptron (MLP) architecture. Two models were developed using different sets of sea surface temperature (SST) indices representing oceanic conditions in the tropical Atlantic and equatorial Pacific, including Niño 3, Niño 1+2, TAN, and TAS, for the period from 1912 to 2015. The forecasts were evaluated using deterministic, probabilistic, and classification-based performance metrics. The model based on Atlantic predictors showed superior performance, with a Nash–Sutcliffe efficiency (NSE) of 0.58 and a correlation coefficient (R) of 0.78 in the test dataset, and it also outperformed climatology (Performance index of 1.16). In classifying dry, normal, and wet years, the model achieved 57% accuracy and performed better at predicting wet years (F1-score of 0.74). The results highlight the relevance of tropical Atlantic variability for streamflow predictability in the region and demonstrate the potential of neural networks as decision-support tools for water resources management in semi-arid environments. Future studies may explore multimodel approaches to further improve forecasting accuracy.

Keywords: sea surface temperature indices; machine learning; multilayer perceptron.

RESUMO

Diante dos desafios impostos pelas condições climáticas de regiões semiáridas que dependem da política de açudagem para garantia de disponibilidade hídrica, métodos de aprendizado de máquina estão sendo usados como ferramentas para a gestão eficiente dos recursos hídricos. Este trabalho teve como objetivo calibrar um modelo de previsão sazonal de vazões para o açude Carnaubal, no Semiárido cearense, utilizando redes neurais artificiais (RNA) do tipo *Multilayer Perceptron* (MLP). Foram desenvolvidos dois modelos com diferentes conjuntos de índices de temperatura da superfície do mar (TSM), incluindo Niño 3, Niño 1+2, TAN, TAS, para o período de 1912 a 2015, representando condições dos oceanos Atlântico tropical e Pacífico equatorial. As previsões foram avaliadas por meio de métricas determinísticas, probabilísticas e de classificação. O modelo baseado em preditores do Atlântico apresentou melhor desempenho, com coeficiente de *Nash-Sutcliffe* (NSE) de 0,58 e coeficiente de correlação (R) de 0,78 no conjunto de teste, além de desempenho superior à climatologia (índice de desempenho de 1,16). Na classificação de anos secos, normais e chuvosos, o modelo obteve acurácia de 57% e melhor desempenho na previsão de anos chuvosos (F1-score de 0,74). Os resultados evidenciam a relevância da variabilidade do Atlântico tropical na previsibilidade das vazões na região e demonstram o potencial das redes neurais como ferramenta de apoio à decisão na gestão de recursos hídricos em ambientes semiáridos. Estudos futuros podem explorar abordagens multimodelo para aprimorar a acurácia das previsões.

Palavras-chave: índices de temperatura da superfície do mar; aprendizado de máquina; *multilayer perceptron*.

[1]Universidade Federal do Ceará – Fortaleza (CE), Brazil.
[2]Universidade Federal do Ceará – Crateús (CE), Brazil.
**In Memoriam*
Corresponding author: Tatiane Lima Batista. Universidade Federal do Ceará, Departamento de Engenharia Hidráulica e Ambiental, Campus do Pici – Bloco 713. Cep: 60400-900. Fortaleza – Ceará, Brazil. E-mail: tatiane@crateus.ufc.br

Received on: 07/30/2025. Accepted on: 06/08/2026.
https://doi.org/10.67790/rbciamb.026.2716

Introduction

Water scarcity is a critical global issue, particularly in arid and semi-arid regions (Herrera-Franco et al., 2024; Pacheco-Treviño and Manzano-Camarillo, 2024; Wang et al., 2024). Climate change and population growth pose additional challenges, further increasing the complexity of this issue and requiring efficient and innovative management solutions (Araújo et al., 2024; Bernabé-Crespo and Loáiciga, 2024; Liao, 2024). In this context, efficient water management helps ensure water availability for present and future generations, aligning with Sustainable Development Goal 6 (SDG 6) of the United Nations (UN), which seeks to ensure the availability and sustainable management of water and sanitation for all (UN, 2015).

Northeastern Brazil (NEB) is predominantly characterized by a semi-arid climate, marked by recurrent and prolonged drought events (Marengo et al., 2016; Marengo et al., 2022), which trigger a cascade of socioeconomic and environmental impacts (Cavalcante et al., 2025). The high variability of rainfall both within and between years, combined with a short rainy season, is mainly influenced by Atlantic and Pacific Ocean temperatures (Medeiros and Oliveira, 2021). In the state of Ceará, the subsurface is predominantly crystalline and associated with shallow soils, favoring the formation of intermittent rivers and limiting groundwater production (IPECE, 2021). Reservoir construction policies in these regions are essential to ensure access to drinking water for local populations, especially during drought years. Therefore, forecasting inflows to reservoirs is fundamental for the efficient management of water resources in these regions, enabling decision-makers to implement more effective reservoir operation and water allocation processes (Souza Filho and Lall, 2004), as well as to anticipate future drought and flood events by adopting strategies better suited to the socio-hydrological realities of the system.

Artificial intelligence-based streamflow forecasting models, particularly machine learning techniques, have been widely applied in hydrology (Deo and Şahin, 2016; Cheng et al., 2020; Ahmed et al., 2021; Gurbuz et al., 2024), demonstrating the ability to capture nonlinearities present in historical series (Ghobadi and Kang, 2022). These models are capable of learning relationships among variables from previously available datasets. Local hydrological variables, such as precipitation and river discharge, are strongly influenced by large-scale global climate factors. Consequently, historical series of climate indices are frequently employed as inputs in these models (Dantas et al., 2020; Zhao et al., 2022; Lappicy and Lima, 2023; Oad et al., 2023).

In this context, semi-arid regions represent challenging environments for hydrological modeling, requiring robust streamflow forecasting models (Al-Juboori, 2023). In the state of Ceará, advanced streamflow forecasting contributes to water resources management planning and to negotiated water allocation processes conducted by river basin committees, involving institutions such as the Ceará State Water Resources Management Company, water users, and municipal administrations. Such tools improve decision-making processes by providing estimates of future reservoir storage volumes and the projected water demands to be met (Pereira et al., 2023). During recurrent drought events, these instruments become particularly relevant, since such events are associated with reduced reservoir storage, agricultural production losses, and increased socioeconomic vulnerability of the population (Walker et al., 2024).

Within this scenario, the Carnaubal Hydrosystem, located in the state of Ceará, recently experienced conflicts over water use during the latest drought period (2012–2018), requiring an inter-basin water transfer between different hydrographic regions to meet regional demands, according to the Proactive Drought Management Plan for the Carnaubal Hydrosystem (COGERH, 2024). Therefore, this region highlights the importance of efficient water resources management to ensure coexistence with drought conditions while meeting the demands of multiple water uses.

In NEB, recent studies have identified consistent relationships between sea surface temperature (SST) indices associated with the Atlantic and Pacific Oceans and rainfall variability in the region, highlighting the role of large-scale climate forcings in regional hydrological dynamics (Rodrigues et al., 2024; Vitorio et al., 2025; Rolim and Souza Filho, 2026).

More recent studies in Brazil have expanded the application of machine learning techniques, such as neural networks and hybrid approaches, to reservoir inflow forecasting, as demonstrated in studies of the Sobradinho Reservoir (Saraiva et al., 2021; Marcos Junior et al., 2024). However, these approaches primarily focus on local variables (streamflow and precipitation) and neglect large-scale climate indices, whose potential for rainfall forecasting has already been demonstrated (Pinheiro and Ouarda, 2023).

Regarding streamflow forecasting using SST indices in NEB, Araújo et al. (2020) modeled seasonal inflows to the Orós Reservoir (CE) using different combinations of SST indices and evaluated performance through deterministic and probabilistic metrics (Nash-Sutcliffe efficiency and maximum likelihood). Sousa et al. (2021) applied neural networks at the Boa Esperança Hydropower Plant (PI), emphasizing calibration and statistical performance analysis. While these studies focused primarily on predictive performance, the present research advances the field by incorporating classification metrics (confusion matrix and F1-score) to distinguish hydrological regimes and by analyzing the influence of climate predictors on semi-arid water management.

This study aims to develop and evaluate a seasonal inflow forecasting model for the Carnaubal Reservoir, located in the semi-arid region of Ceará, using Multilayer Perceptron (MLP) neural networks based on SST indices. The proposed approach integrates deterministic, probabilistic, and classification metrics to assess model performance from multiple perspectives, evaluate the influence of climatic variables, and discuss the potential application of the results to decision-making processes in semi-arid environments.

Methodological procedures

The methodology adopted in this study was divided into two main stages (Figure 1). The first stage involved collecting and organizing the input data for the model, consisting of historical series of inflows to the Carnaubal Reservoir and SST indices. This stage also included selecting predictor variables and preprocessing the data. The second stage consisted of calibrating and validating the MLP neural network model, followed by evaluating its performance using deterministic and probabilistic metrics.

Study area

The study area for this research was the Carnaubal Reservoir, located in the municipality of Crateús, in the state of Ceará, Brazil. The Carnaubal Reservoir is the main reservoir of the Carnaubal Hydrosystem, supplying the municipality of Crateús together with the smaller Batalhão Reservoir, and is part of the Sertões de Crateús Hydrographic Region (RHSC). According to the Water Resources Plan for the Sertões de Crateús Hydrographic Region, the RHSC is located in the western portion of the state of Ceará, between the coordinates 4°36’44”S – 6°07’40”S latitude and 40°00’00”W – 41°09’16”W longitude (COGERH, 2021a). It borders the Serra da Ibiapaba and Acaraú hydrographic regions to the north, the Banabuiú Hydrographic Region to the east, the Upper Jaguaribe Hydrographic Region to the south, and the state of Piauí to the west, covering an area of 10,793.80 km^2. Figure 2 shows the location of the Carnaubal Reservoir and the RHSC.

Data collection

Time series of indices calculated from SST anomalies were used in this study. Initially, the seven indices described in Table 1 were evaluated, covering the period from 1912 to 2015. Monthly data were obtained from the International Research Institute for Climate and Society (IRI) website and from the Earth System Research Laboratory (ESRL) portal, a division of the National Oceanic and Atmospheric Administration (NOAA).

The inflow series, used as the dependent variable in this study, was obtained from the Alocar Project (COGERH, 2021b). Figure 3 presents the observed mean monthly inflows and mean annual inflows calculated for the period from 1913 to 2016, which was considered in this research.

March, April, and May have the highest mean inflow values. During the remaining months, the mean values decrease substantially, approaching zero during the second half of the year. It is also possible to observe interannual variability in streamflow, with some years showing inflows close to zero and others exhibiting peaks ranging from 6 to 14 m^3/s, as shown in Figure 3b.

Selection of model input variables

The SST index data series were organized as quarterly averages in order to filter out monthly oscillations that could introduce noise into the data and to account for seasonality (Souza Filho and Lall, 2004). Thus, 70 data series were obtained, corresponding to 10 series for each variable, representing the 10 predictor quarters: JFM (January–February–March), FMA (February–March–April), MAM (March–April–May), AMJ (April–May–June), MJJ (May–June–July), JJA (June–July–August), JAS (July–August–September), ASO (August–September–October), SON (September–October–November), and OND (October–November–December). Subsequently, correlations were calculated between each series and the mean annual inflow series.

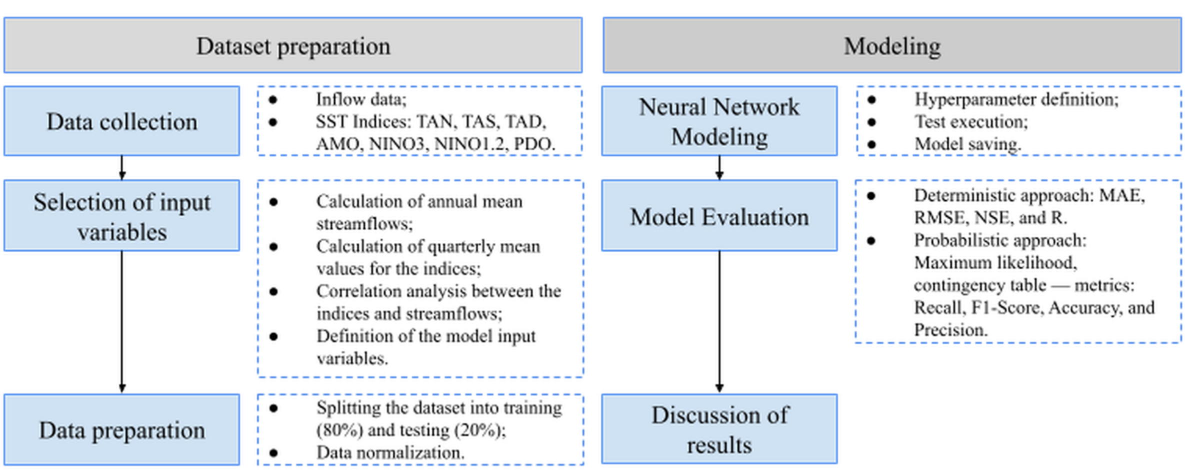

Figure 1 – Research framework.

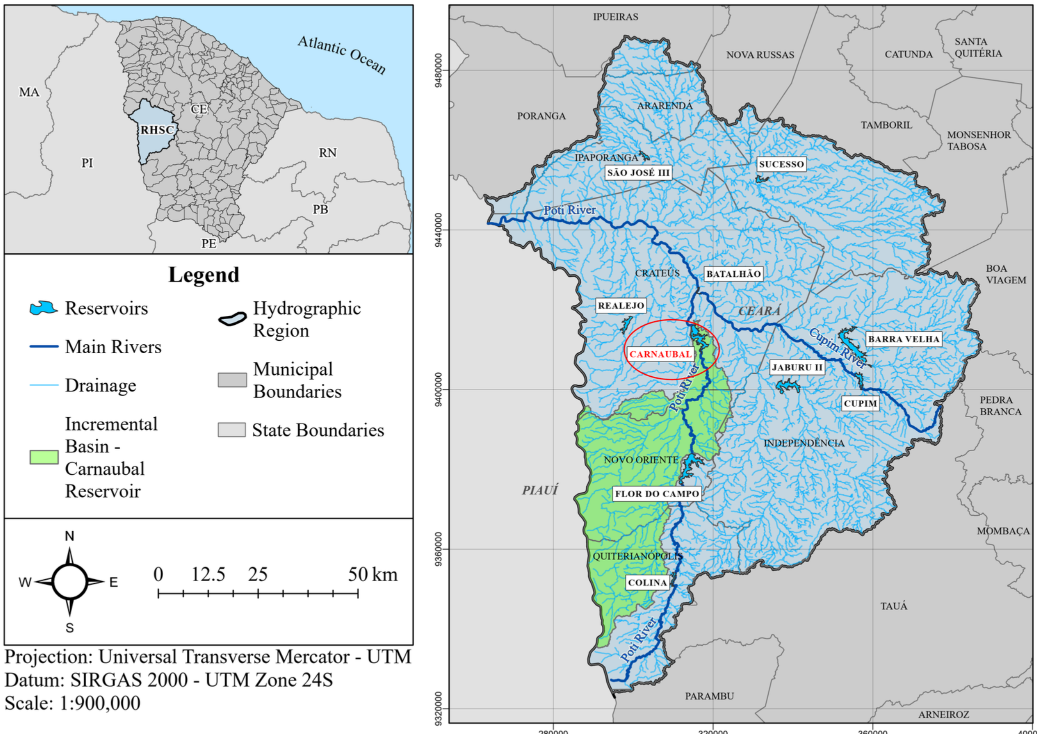

Figure 2 – Location of the Carnaubal Reservoir and the Sertões de Crateús Hydrographic Region.

Table 1 – Description of the SST indices initially evaluated.

Index	Description	Source
NINO3	Niño 3: Eastern Tropical Pacific SST (5°N–5°S and 150°W–90°W)	Developed by Kaplan et al. (1998), available at: http://iridl.ldeo.columbia.edu/SOURCES/.KAPLAN/.
NINO1.2	Niño 1+2 (0-10S and 90W-80W)	
TAS	Tropical South Atlantic – SST anomaly (0°–20°S and 10°E–30°W)	
TAN	Tropical North Atlantic – SST anomaly (5.5°N–23.5°N and 15°W–57.5°W)	
TAD	Tropical Atlantic Dipole	TAN - TAS
AMO	Atlantic Multidecadal Oscillation	NOAA – available at: http://www.esrl.noaa.gov/psd/data/climateindices/list.
PDO	Pacific Decadal Oscillation	

For each index, the predictor period with the highest correlation coefficient was selected, yielding seven series. A correlation matrix was then calculated for the dataset composed of the seven previously selected indices and the mean annual inflow series. The model input variables were subsequently defined based on the analysis of these results.

Data preparation for modeling

The data used in this study were initially loaded into a DataFrame from a CSV file using the pandas library (Pandas Development Team, 2024). The training (80%) and testing (20%) datasets were randomly split using the train_test_split function with random_state = 5, ensuring sampling reproducibility. The variables were then normalized to the range 0.15-0.85 using the MinMaxScaler, following the methodology of Araújo et al. (2020). The normalization parameters were fitted exclusively on the training dataset and subsequently applied to the testing dataset (Equation 1).

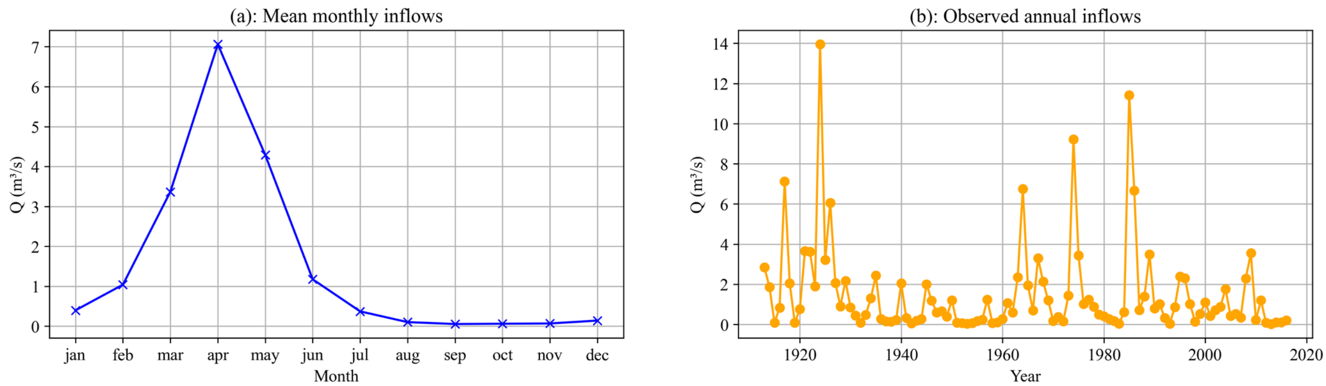

Figure 3 – (a) Mean monthly inflows (m^3/s) to the Carnaubal Reservoir and (b) mean annual inflows (m^3/s) to the Carnaubal Reservoir for the period from 1913 to 2016. Source: COGERH (2021b).

$$x_{normalized} = \left(\frac{x - \min(x)}{\max(x) - \min(x)}\right)(b - a) + a \quad (1)$$

Where: variable a was set to 0.15 and variable b to 0.85 in this study; min(x) represents the minimum value of the training dataset for each variable; max(x) represents the maximum value of the training dataset for each variable; and x is the original data value.

Development of the predictive model

Artificial neural networks are composed of interconnected neurons that transmit output signals from one neuron as inputs to others in the subsequent layer. Their parametric weights control the influence of each input signal on the target neuron. During training, these weights are iteratively adjusted to minimize the difference between the network outputs and the desired results. This weight adaptation process is the mechanism through which neural networks learn complex relationships within the data. Therefore, an artificial neural network architecture consists of simple interconnected processing units whose synaptic weight configuration encodes the learned knowledge (James et al., 2023).

A network containing multiple layers with connections among all neurons is referred to as an MLP. An MLP is composed of an input layer, hidden layers, and an output layer (Géron, 2019). The MLP neural network model was implemented using the Keras library. The adopted architecture consisted of an input layer with n predictor variables, one fully connected hidden layer with a ReLU activation function, and an output layer with a single neuron and a Sigmoid activation function.

The Sigmoid activation function was employed in the output layer to constrain the predicted values to the range between 0 and 1, consistent with the normalization applied to the streamflow variable. Hyperparameter definitions were determined through a systematic search evaluating different parameter combinations. The following were tested: (i) the number of neurons in the hidden layer, ranging from 2 to 25; (ii) learning rates of 0.001, 0.01, and 0.1; (iii) δ parameter values for the Huber Loss function ranging from 0.1 to 0.9; (iv) the number of epochs between 100 and 800; and (v) batch sizes of 10 and 20.

Training was performed using the Adam optimizer, and the final model configuration was selected based on the performance achieved on the validation dataset, with a focus on minimizing prediction errors.

The hyperparameter combinations that resulted in the best performance were as follows: the final MLP model architecture consisted of one hidden layer with 20 neurons and a ReLU activation function, and one output layer with a single neuron and a Sigmoid activation function. Training was performed using the Adam optimizer (learning rate = 0.01) and the Huber Loss function (δ = 0.5), with a batch size of 20. A total of 600 epochs were used for Model 1 and 700 epochs for Model 2.

To ensure reproducibility of the results, a random seed was fixed during model training.

Model evaluation

The model was initially evaluated using the following performance metrics (Equations 2 to 5): Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Pearson Correlation Coefficient (R), and Nash–Sutcliffe Efficiency (NSE), using the Python programming language version 3.11 in the Spyder software environment.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(O_i - P_i)^2} \quad (2)$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|O_i - P_i| \quad (3)$$

$$R = \frac{\sum_{i=1}^{n} (O_i - \bar{O})(P_i - \bar{P})}{\sqrt{\sum_{i=1}^{n} (O_i - \bar{O})^2} \sqrt{\sum_{i=1}^{n} (P_i - \bar{P})^2}} \quad (4)$$

$$NSE = 1 - \frac{\sum_{i=1}^{n} (O_i - P_i)^2}{\sum_{i=1}^{n} (O_i - \bar{O})^2} \quad (5)$$

Where: n is the number of observations; O_i denotes the observed values of mean annual inflows; P_i denotes the values predicted by the models; $\overline{O}$ is the mean value of the observed inflows and $\overline{P}$ is the mean value of the inflows predicted by the model.

To capture the uncertainty associated with the forecasts, the errors (observed minus predicted values) were fitted to a Student's t-distribution. Adjusted forecasts were then generated by incorporating uncertainty through simulating random samples from the t-distribution. A 95% confidence interval was calculated for the forecasts, providing a measure of the expected variability in the results. It is noteworthy that a normal distribution was initially tested for the error series; however, it did not present a satisfactory fit.

The observed inflow series was classified into three categories (dry, normal, and wet years) by dividing the data distribution into terciles. The thresholds were defined using the 33rd and 66th percentiles of the historical series, such that values below the 33rd percentile were classified as dry years, values between the 33rd and 66th percentiles as normal years, and values above the 66th percentile as wet years. For each year, the category of the predicted inflow was also defined based on the largest area under the probability distribution curve of the adjusted forecasts.

Model performance was then calculated using the maximum likelihood ratio, with climatology adopted as the reference baseline, as described by Silveira (2014) and Araújo et al. (2020). Likelihood was defined as the nth root of the product of the probabilities assigned by the model to the same category as the observed data for each year. Performance values greater than 1 indicate that the model performs better than climatology.

The predicted categories were compared with the observed categories, enabling evaluation of model accuracy in terms of classification capability. Based on the predicted and observed category data, a contingency table, or confusion matrix (Grus, 2016), was constructed, and the following performance metrics were calculated: accuracy, precision, recall, and F1-score.

Accuracy measures the proportion of correct predictions among all predictions made. Precision represents the proportion of correctly predicted positive cases among all positive predictions, indicating how reliable positive predictions are. Recall, or sensitivity, measures the proportion of true positive cases correctly identified among all actual positive cases. The F1-score is the harmonic mean of precision and recall and is commonly used to evaluate the balance between these metrics (Grus, 2016).

Results and discussion

Analysis of input variables for the model

The correlation analysis between the SST indices and inflows to the Carnaubal Reservoir, presented in Figure 4, identified SON as the best predictor period for the TAD, TAN, AMO, and NINO1.2 indices. For TAS, the selected period was JJA; for NINO3, AMJ; and for PDO, MJJ. It is noteworthy that recent studies have also used these indices for streamflow and rainfall forecasting in NEB watersheds (Araújo et al., 2020; Sousa et al., 2021).

Based on Figure 5, the indices TAD_SON, TAN_SON, AMO_SON, and TAS_JJA showed the best performance, with correlation values above 25% relative to streamflow. High correlation values (above 50%) were also observed among TAD_SON, TAN_SON, and AMO_SON, as well as among NINO3_AMJ, NINO1.2_SON, and PDO_MJJ.

Based on this analysis, two models were defined for testing: the first including the four best-performing indices and the second excluding indices with high intercorrelations. The tested models were: Model 1: Q (TAD_SON, TAN_SON, AMO_SON, and TAS_JJA) and Model 2: Q (TAD_SON, TAS_JJA, NINO3_AMJ, and PDO_MJJ).

Neural network training

The tests indicated that the use of the Adam optimizer combined with the Huber Loss function provided the best modeling performance. It was observed that increasing the number of epochs improved model fitting up to a stabilization point. For Model 1, the best performance was achieved with 600 epochs, whereas for Model 2, optimal performance was obtained with 700 epochs. These configurations yielded the best results on the validation dataset and were therefore adopted in the final simulations.

Model evaluation

The calculated evaluation metrics (Table 2) showed that the results of both evaluated models were very similar. However, based on an initial analysis of these results, Model 1 was selected because it presented better Performance index values than Model 2 during both the training and testing stages. Therefore, the analyses presented below refer only to Model 1. It is important to note that a normal distribution was initially tested for the error series; however, it did not present a satisfactory fit.

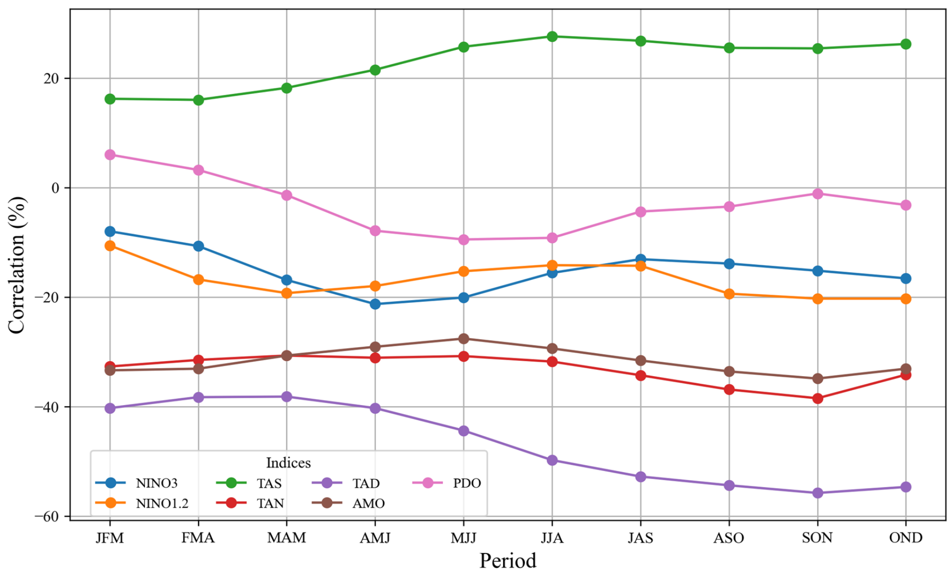

Figure 4 – Correlation between streamflow and SST indices for each predictor period.

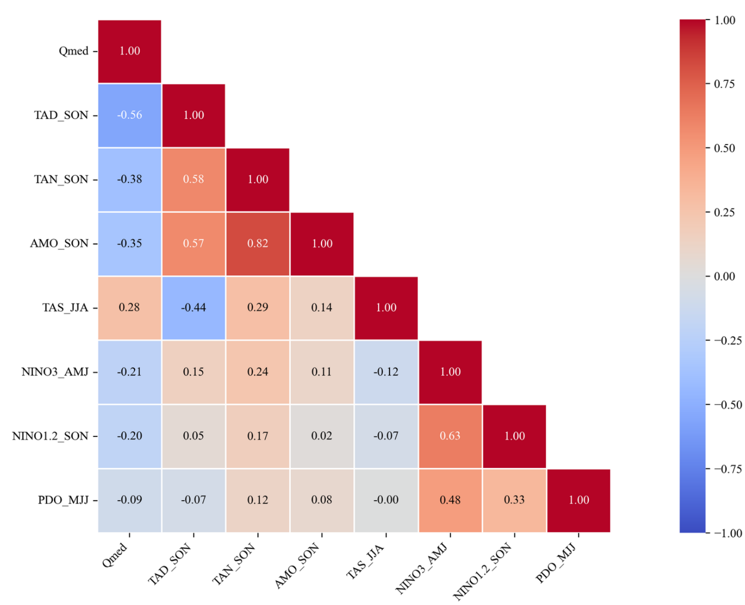

Figure 5 – Correlation matrix among SST indices, considering the predictor period with the highest correlation value with streamflow for each index.

Table 2 – Evaluation metrics of the artificial neural network model.

Model	Dataset	R	RMSE	MAE	NSE	Performance
1	Training	0.83	1.26	0.89	0.69	1.17
	Testing	0.78	1.52	1.08	0.58	1.16
2	Training	0.81	1.35	0.88	0.64	1.14
	Testing	0.78	1.48	1.05	0.60	1.08

When plotting the observed and predicted values (Figure 6), a certain dispersion around the y=x line was observed, which represents the equality between predicted and observed values. For peak flows (greater than 6 m^3/s), the points were predominantly located below the line (with only one exception), indicating that the predicted values were lower than the observed values in these cases. For streamflows below 6 m^3/s, the opposite pattern was observed: approximately 70% of the points were located above the y=x line, indicating that predicted values were higher than observed values in most cases.

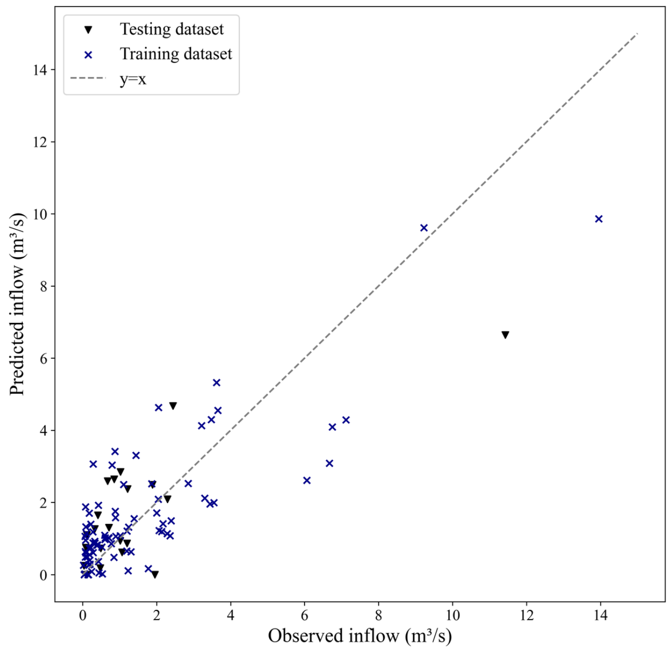

Figure 6 – Correlation between streamflow values predicted by Model 1 and observed inflows to the Carnaubal Reservoir during the training and testing phases.

The model showed better performance during training than during testing, which was expected. However, the metric values obtained for the testing dataset were very close to those observed for the training dataset, indicating that model overfitting did not occur. The Pearson correlation coefficient (R) indicated a high correlation between observed and predicted values (0.78 for the testing phase). The Nash–Sutcliffe Efficiency coefficient (NSE) indicated satisfactory model performance (Moriasi et al., 2007), with a value of 0.58 during testing.

Compared with previous studies, the values obtained for these indices can be considered satisfactory. Araújo et al. (2020) obtained an NSE value of 0.5 for seasonal inflow forecasting to the Orós Reservoir in Ceará using an MLP neural network. Sousa et al. (2021) obtained NSE values ranging from 0.56 to 0.64 for seasonal streamflow forecasting in the Parnaíba River, located on the border between the states of Piauí and Maranhão, also using an MLP neural network.

Analysis of the prediction errors indicated an MAE of approximately 1.0 m^3/s. RMSE values were higher than MAE values. As previously observed, the streamflow series presents peak events, during which prediction errors were predominantly larger in absolute terms, as shown in Figure 6, contributing to the increase in RMSE relative to MAE.

When evaluating the model using the Performance index, it was observed that the model predicts annual inflows to the Carnaubal Reservoir better than climatology, as the values obtained were greater than 1.0. This result is consistent with Araújo et al. (2020), who found a Performance value of 1.3 for an MLP neural network model.

Based on the analysis of the forecast graph with uncertainty bars considering a 95% confidence interval for both the training and testing phases (Figure 7), it can be observed that the uncertainty bars tend to increase during peak flow events, indicating that the model is capable of capturing increases in streamflow during those years. This behavior may be interpreted in light of the high hydrological variability characteristic of NEB, where streamflow dynamics are marked by strong nonstationarity and alternation between dry periods and high-magnitude events. Recent studies have shown that this variability is associated with the influence of climatic forcings and the recurrence

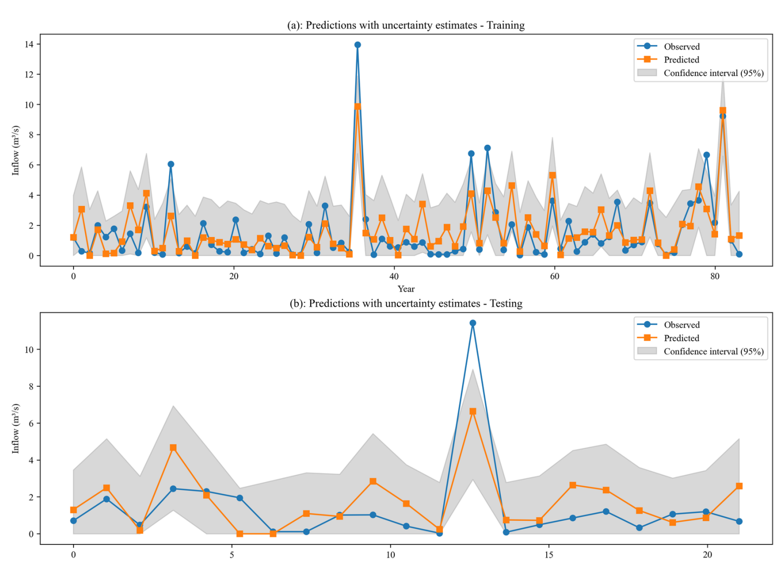

Figure 7 – Forecasts with uncertainty bars considering a 95% confidence interval for model errors during the training and testing phases.

of hydrological extremes in the region (Gesualdo et al., 2024; Rolim and Souza Filho, 2026).

Tables 3 and 4 present, respectively, the confusion matrix (or contingency table) and the model performance metrics for forecasting dry, normal, and wet years.

Table 3 – Contingency table for evaluating model forecasts during the training and testing phases.

Training		Model			Testing		Model		
		D	N	W			D	N	W
Observed	D	18	3	7	Observed	D	4	0	0
	N	6	11	6		N	4	1	1
	W	4	2	26		W	4	0	7

The confusion matrix provides an initial assessment of model performance in classifying hydrological regimes, highlighting patterns of correct and incorrect classifications across categories. The model exhibited varying classification capabilities across categories, with higher concentrations of correct predictions in some classes and greater confusion in others. However, a more detailed evaluation requires the analysis of classification metrics such as precision, recall, and F1-score, presented below, which provide a more robust quantification of the model's ability to represent each hydrological regime (Sathyanarayanan and Tantri, 2024).

Table 4 – Model performance metrics for forecasting dry, normal, and wet years.

Metrics		Training	Testing
Accuracy		0.66	0.57
Precision	D	0.64	0.33
	N	0.69	1.00
	W	0.67	0.88
Recall	D	0.64	1.00
	N	0.48	0.17
	W	0.81	0.64
F1-Score	D	0.64	0.50
	N	0.56	0.29
	W	0.73	0.74

Analysis of Tables 3 and 4 indicates that, overall, the model achieved higher accuracy during training than during testing. During training, 66% of the categories were correctly classified, whereas during testing this value was 57%.

Considering the Dry (D) class separately, precision was equal to recall and, consequently, equal to the F1-score for the training dataset. This indicates that 64% of the years predicted by the model as dry were indeed dry, while 4 years (14%) were actually wet years. Regarding recall, 64% of the years that were actually dry were correctly predicted as dry, whereas 7 years (25%) were predicted as wet. In the testing phase, the results differed: only 33% of the years predicted as dry were actually dry, while 4 years (33%) were wet. However, recall was equal to 1, meaning that all dry years were correctly identified by the model. This indicates that, when generalizing to new data, the model does not present difficulty in identifying dry years, although it may incorrectly classify non-dry years as dry.

For the Normal (N) class, 69% of the years predicted as normal by the model were indeed normal during training. The remaining years were either dry or wet. In the case of the Normal class, the implications of misclassification toward either Dry or Wet are equivalent, since this class lies between the two extremes. Regarding actual normal years, 48% were correctly classified as normal, while the remaining cases were equally classified as dry or wet. During testing, the model predicted only one year as normal, which was indeed normal, resulting in a precision of 100%. However, the other four normal years were predicted as dry, resulting in a recall of 17%.

For the Wet (W) class, 67% of the years predicted as wet by the model during training were indeed wet. Among the remaining cases, 7 years (18%) were actually dry. In terms of recall, 81% of the wet years were correctly classified as wet, while only 13% were classified as dry. During testing, 88% of the years predicted as wet were indeed wet. The only year predicted as wet that was not actually wet was classified as normal. Regarding recall, 64% of wet years were correctly identified, while the remaining four years were predicted as dry.

Considering the F1-score, which summarizes precision and recall, the model achieved better performance for wet years during both training and testing. Based on these results and analyses, it can be highlighted that if the model predicts a wet year, there is a high probability that the year will indeed be wet. However, when analyzing the model's generalization capability on the testing dataset, it can be observed that normal or wet years may still be classified as dry, indicating a tendency of the model to predict dry years in the testing phase. This behavior is consistent with recent studies showing that the performance of machine learning models in streamflow classification problems depends on class definition and separation (AlDahoul et al., 2023), suggesting that adjustments in class delimitation may contribute to improved model performance. These results reinforce the importance of classification metrics, allowing a more detailed assessment of model performance across different hydrological regimes.

It should be noted that the relatively small number of years allocated to testing may have contributed to these results. The use of a longer time series could provide better conditions for evaluating model performance. Similar data scarcity limitations were also highlighted by Dong et al. (2025), who applied deep neural networks with climate forecast data to estimate streamflows, even under scenarios with limited observational records.

The obtained results indicate that modeling based on SST indices presents potential for application in water resources management in the semi-arid region of Ceará, especially in the context of reservoir operation and water allocation processes. The model's ability to represent streamflow variability and more consistently identify wet years suggests that its application may contribute to improving currently adopted forecasting approaches. These findings are consistent with recent studies highlighting the potential of neural networks and climate indices for improving hydrological forecasting (Lappicy and Lima, 2023; Pokharel and Roy, 2024).

On the other hand, the limitations observed in distinguishing between dry and normal years, as well as the reduced size of the testing sample, indicate that operational application of the model still requires caution. Future advances may include the extension of historical series, the use of multimodel approaches, and the incorporation of additional relevant variables to improve forecast robustness and reliability for supporting decision-making processes.

Conclusions

The results indicate that models based on Atlantic Ocean SST indices presented slightly superior performance compared with those incorporating Pacific indices, although the differences among the evaluation metrics were not substantial. In addition, the SON predictor period showed the strongest relationship with inflows, highlighting the importance of seasonality in model development.

Overall, the MLP model demonstrated satisfactory performance in forecasting inflows to the Carnaubal Reservoir, outperforming climatology and yielding results consistent with those reported in the literature. The model showed better capability in representing wet years, indicating possible limitations in forecasting drier conditions.

These findings reinforce the potential of using large-scale climate indices as predictors in seasonal streamflow forecasting models for semi-arid regions, contributing to water resources management under conditions of high climatic variability. As a future perspective, the adoption of multimodel approaches is highlighted as a promising alternative for improving forecast robustness and accuracy.

Acknowledgments

The authors would like to thank the Laboratório de Gerenciamento do Risco Climático para a Sustentabilidade Hídrica (GRC-UFC) for providing the infrastructure necessary for the development of this study, and Companhia de Gestão dos Recursos Hídricos, within the scope of the ALOCAR Project, for providing the data used in this research.

Authors' Contributions: Batista, T.L.: conceptualization, data curation, formal analysis, investigation, methodology, writing – original draft, writing – review and editing; Alexandre, A.M.B.: conceptualization, data curation, formal analysis, methodology, supervision, validation, visualization, writing – review and editing; Araújo, J.K.S.: conceptualization, data curation, formal analysis, methodology, validation, visualization, writing – review and editing; Souza Filho, F.A.: conceptualization, data curation, formal analysis, funding acquisition, methodology, project administration, resources, supervision, validation, writing – review and editing.

Conflicts of interest: the authors declare no conflicts of interest.

Funding: This study was carried out with the support of the Fundação Cearense de Apoio ao Desenvolvimento Científico e Tecnológico (FUNCAP), Grant Numbers: 29012.006446/2024-76 and 31052.000039/2024-21. This study was also carried out with the support of the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001.

Declaration on the use of generative artificial intelligence: During the preparation of this manuscript, the authors used ChatGPT, developed by OpenAI, for translation, grammar correction, and language editing. After using this tool, the authors reviewed and edited the content as necessary and take full responsibility for the content of the published article.

References

Ahmed, A.A.M.; Deo, R.C.; Feng, Q.; Ghahramani, A.; Raj, N.; Yin, Z.; Yang, L., 2021. Deep learning hybrid model with Boruta-Random forest optimiser algorithm for streamflow forecasting with climate mode indices, rainfall, and periodicity. Journal of Hydrology, v. 599, 126350. https://doi.org/10.1016/j.jhydrol.2021.126350.

AlDahoul, N.; Momo, M.A.; Chong, K.L.; Ahmed, A.N.; Huang, Y.F.; Sherif, M.; El-Shafie, A., 2023. Streamflow classification by employing various machine learning models for peninsular Malaysia. Scientific Reports, v. 13 (1). https://doi.org/10.1038/s41598-023-41735-9

Al-Juboori, A.M., 2023. Prediction of Hydrological Drought in Semi-arid Regions Using a Novel Hybrid Model. Water Resources Management, v. 37 (9), 3657-3669. https://doi.org/10.1007/s11269-023-03520-1

Araújo, C.B.C.; Souza Filho, F.A.; Araújo Júnior, L.M.; Silveira, C.S., 2020. Previsão Sazonal de Vazões para a Bacia do Orós (Ceará, Brasil) Utilizando Redes Neurais e a Técnica de Reamostragem dos K-vizinhos. Revista Brasileira de Meteorologia, v. 35 (2), 197-207. https://doi.org/10.1590/0102-7786351015.

Araújo, D.C.S.; Montenegro, S.M.G.L.; Silva, S.F.; Farias, V.E.M.; Rodrigues, A.B., 2024. Analysis of climate change scenarios using CMIP6 models in Pernambuco, Brazil. Revista Brasileira de Ciências Ambientais (RBCIAMB), v. 59, e1868. http://dx.doi.org/10.5327/z2176-94781868.

Bernabé-Crespo, M.B.; Loáiciga, H., 2024. Managing potable water in Southeastern Spain, Los Angeles, and Sydney: transcontinental approaches to overcome water scarcity. Water Resources Management, v. 38 (4), 1299-1313. http://dx.doi.org/10.1007/s11269-023-03721-8.

Cavalcante, L.; Walker, D.W.; Kchouk, S.; Ribeiro Neto, G.; Carvalho, T.M.N.; Brito, M.M.; Pot, W.; Dewulf, A.; van Oel, P.R., 2025. From insufficient rainfall to livelihoods: Understanding the cascade of drought impacts and policy implications. Natural Hazards and Earth System Sciences, v. 25 (6), 1993-2005. https://doi.org/10.5194/nhess-25-1993-2025

Cheng, M.; Fang, F.; Kinouchi, T.; Navon, I.M.; Pain, C.C., 2020. Long lead-time daily and monthly streamflow forecasting using machine learning methods. Journal of Hydrology, v. 590, 125376. http://dx.doi.org/10.1016/j.jhydrol.2020.125376.

Companhia de Gestão dos Recursos Hídricos (COGERH), 2021a. Plano de Recursos Hídricos da Região Hidrográfica dos Sertões de Crateús.

Companhia de Gestão dos Recursos Hídricos (COGERH), 2021b. Relatório de cálculo das afluências aos reservatórios estratégicos do Ceará: definição das vazões oficiais. Projeto ALOCAR.

Companhia de Gestão dos Recursos Hídricos. (COGERH), 2024. Plano de Gestão Proativa de Secas – Hidrossistema Carnaubal.

Dantas, L.G.; Santos, C.A.C.; Olinda, R.A.; Brito, J.I.B.; Santos, C.A.G.; Martins, E.S.P.R.; Oliveira, G.; Brunsell, N.A., 2020. Rainfall prediction in the state of Paraíba, Northeastern Brazil using generalized additive models. Water, v. 12 (9), 2478. http://dx.doi.org/10.3390/w12092478.

Deo, R.C.; Şahin, M., 2016. An extreme learning machine model for the simulation of monthly mean streamflow water level in eastern Queensland. Environmental Monitoring and Assessment, v. 188 (2), 1-24. http://dx.doi.org/10.1007/s10661-016-5094-9.

Dong, N.; Hao, H.; Yang, M.; Wei, J.; Xu, S.; Kunstmann, H., 2025. Deep-learning-based sub-seasonal precipitation and streamflow ensemble forecasting over the source region of the Yangtze River. Hydrology and Earth System Sciences, v. 29 (8), 2023-2042. https://doi.org/10.5194/hess-29-2023-2025.

Géron, A., 2019. Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow. 2nd ed. O'Reilly Media, Sebastopol, 850 p.

Gesualdo, G.C.; Benso, M.R.; Mendiondo, E.M.; Brunner, M.I., 2024. Spatially Compounding Drought Events in Brazil. Water Resources Research, v. 60 (11). https://doi.org/10.1029/2023WR036629.

Ghobadi, F.; Kang, D., 2022. Improving long-term streamflow prediction in a poorly gauged basin using geo-spatiotemporal mesoscale data and attention-based deep learning: a comparative study. Journal of Hydrology, v. 615, 128608. http://dx.doi.org/10.1016/j.jhydrol.2022.128608.

Grus, J., 2016. Data Science do Zero. Starlin Alta Editora e Consultoria Eireli, São Paulo.

Gurbuz, F.; Mudireddy, A.; Mantilla, R.; Xiao, S., 2024. Using a physics-based hydrological model and storm transposition to investigate machine-learning algorithms for streamflow prediction. Journal of Hydrology, v. 628, 130504. http://dx.doi.org/10.1016/j.jhydrol.2023.130504.

Herrera-Franco, G.; Morante-Carballo, F.; Bravo-Montero, L.; Valencia-Robles, J.; Aguilar-Aguilar, M.; Martos-Rosillo, S.; Carrión-Mero, P., 2024. Water sowing and harvesting (WS&H) for sustainable management in Ecuador: a review. Heritage, v. 7 (7), 3696-3718. https://doi.org/10.3390/heritage7070175

Instituto de Pesquisa e Estratégia Econômica do Ceará (IPECE), 2021. Ceará em números 2021. http://www2.ipece.ce.gov.br/publicacoes/ceara_em_numeros/2021/territorial/index.htm

James, G.; Witten, D.; Hastie, T.; Tibshirani, R., 2023. An Introduction to Statistical Learning: With Applications in R. Springer, New York, 615 p.

Kaplan, A.; Cane, M.A.; Kushnir, Y.; Clement, A.C.; Blumenthal, M.B.; Rajagopalan, B., 1998. Analyses of global sea surface temperature 1856–1991. Journal Geophysical Research, v. 103 (C9), 18567-18589. https://doi.org/10.1029/97JC01736

Lappicy, T.; Lima, C.H.R., 2023. Enhancing monthly streamflow forecasting for Brazilian hydropower plants through climate index integration with stochastic methods. Revista Brasileira de Recursos Hídricos, v. 28, e48. http://dx.doi.org/10.1590/2318-0331.282320230118.

Liao, R., 2024. Water scarcity assessment index from the realistic perspective of human basic water requirements. Environmental and Sustainability Indicators, v. 22, 100404. http://dx.doi.org/10.1016/j.indic.2024.100404.

Marcos Junior, A.D.; Silveira, C.S.; Costa, J.M.F.; Gonçalves, S.T.N., 2024. Combining traditional hydrological models and machine learning for streamflow prediction. Revista Brasileira de Recursos Hídricos (RBRH), v. 29, e11. https://doi.org/10.1590/2318-0331.292420230105

Marengo, J.A.; Galdos, M.V.; Challinor, A.; Cunha, A.P.; Marin, F.R.; Vianna, M. dos S.; Alvala, R.C.S.; Alves, L.M.; Moraes, O.L.; Bender, F., 2022. Drought in Northeast Brazil: A review of agricultural and policy adaptation options for food security. Climate Resilience and Sustainability, v. 1 (1), e17. https://doi.org/10.1002/cli2.17

Marengo, J.A.; Torres, R.R.; Alves, L.M., 2016. Drought in Northeast Brazil—past, present, and future. Theoretical and Applied Climatology, v. 129 (3-4), 1189-1200. http://dx.doi.org/10.1007/s00704-016-1840-8.

Medeiros, F.J.; Oliveira, C.P., 2021. Dynamical aspects of the recent strong El Niño events and its climate impacts in Northeast Brazil. Pure and Applied Geophysics, v. 178 (6), 2315-2332. http://dx.doi.org/10.1007/s00024-021-02758-3.

Moriasi, D.N.; Arnold, J.G.; Van Liew, M.W.; Bingner, R.L.; Harmel, R.D.; Veith, T.L., 2007. Model Evaluation Guidelines for Systematic Quantification of Accuracy in Watershed Simulations. Transactions of the ASABE, v. 50 (3), 885–900. https://doi.org/10.13031/2013.23153.

Oad, S.; Imteaz, M.A.; Mekanik, F., 2023. Artificial neural network (ANN) – based long-term streamflow forecasting models using climate indices for three tributaries of Goulburn River, Australia. Climate, v. 11 (7), 152. http://dx.doi.org/10.3390/cli11070152.

Pacheco-Treviño, S.; Manzano-Camarillo, M.G., 2024. The socioeconomic dimensions of water scarcity in urban and rural Mexico: a comprehensive assessment of sustainable development. Sustainability, v. 16 (3), 1011. https://doi.org/10.3390/su16031011.

Pandas Development Team, 2024. Pandas: Python Data Analysis Library. https://pandas.pydata.org.

Pereira, J.M.R.; Raimundo, C.C.; Reis, D.S.; Vasconcelos, F.C.; Martins, E.S.P.R., 2023. Use of Climate Information in Water Allocation: A Case of Study in a Semiarid Region. Water (Switzerland), v. 15 (13). https://doi.org/10.3390/w15132460

Pinheiro, E.; Ouarda, T.B.M.J., 2023. Short-lead seasonal precipitation forecast in northeastern Brazil using an ensemble of artificial neural networks. Scientific Reports, v. 13 (1). https://doi.org/10.1038/s41598-023-47841-y

Pokharel, S.; Roy, T., 2024. A parsimonious setup for streamflow forecasting using CNN-LSTM. Journal of Hydroinformatics, v. 26 (11), 2751-2761. https://doi.org/10.2166/hydro.2024.114.

Rodrigues, B.D.; Silveira, C.S.; Vasconcelos Júnior, F.C.; Brito Neto, F.A.; Silva, I.A. e.; Sakamoto, M.S.; Martins, E.S.P.R., 2024. Atmospheric and oceanic mechanisms in precipitation in March 2018 in Ceará, Brazil. Theoretical and Applied Climatology, v. 155 (9), 8633–8650. https://doi.org/10.1007/s00704-024-05143-x

Rolim, L.Z.R.; Souza Filho, F.A., 2026. Analyzing rainfall and streamflow variability in semi-arid Ceará: the role of climate drivers. Theoretical and Applied Climatology, v. 157 (3), 149. https://doi.org/10.1007/s00704-026-06058-5

Saraiva, S.V.; Carvalho, F.O.; Santos, C.A.G.; Barreto, L.C.; Freire, P.K.M.M., 2021. Daily streamflow forecasting in Sobradinho Reservoir using machine learning models coupled with wavelet transform and bootstrapping. Applied Soft Computing, v. 102. https://doi.org/10.1016/j.asoc.2021.107081

Sathyanarayanan, S.; Tantri, B.R., 2024. Confusion Matrix-Based Performance Evaluation Metrics. African Journal of Biomedical Research, v. 27, 4023-4031. https://doi.org/10.53555/ajbr.v27i4s.4345

Silveira, C.S., 2014. Modelagem Integrada de Meteorologia e Recursos Hídricos em Múltiplas Escalas Temporais e Espaciais: Aplicação no Ceará e no Setor Hidroelétrico Brasileiro. Doctoral Thesis, Universidade Federal do Ceará, Fortaleza. https://repositorio.ufc.br/handle/riufc/11327

Sousa, D.L.; Souza Filho, F.A.; Araujo, R.B.A.; Farias, G.M.; Castro, M.A.H., 2021. Previsão Sazonal de Vazões para a Usina Hidrelétrica de Boa Esperança-PI utilizando Redes Neurais Artificiais. Revista Brasileira de Geografia Física, v. 14 (5), 2629–2645. https://doi.org/ 10.26848/rbgf.v14.5.p2629-2645.

Souza Filho, F.A.; Lall, U., 2004. Modelo de previsão de vazões sazonais e interanuais. Revista Brasileira de Recursos Hídricos (RBRH), v. 9 (2), 61-74.

United Nations (UN), 2015. Transforming our world: the 2030 agenda for sustainable development (A/RES/70/1). United Nations. https://undocs.org/en/A/RES/70/1.

Vitorio, E.L.; Silva, T.; Vilela, I.; Santos, E.; Veleda, D., 2025. Interannual rainfall variability in Northeast Brazil influenced by Pacific and Atlantic climate modes. Dynamics of Atmospheres and Oceans, v. 112. https://doi.org/10.1016/j.dynatmoce.2025.101596

Walker, D.W.; Oliveira, J.L.; Cavalcante, L.; Kchouk, S.; Ribeiro Neto, G.; Melsen, L.A.; Fernandes, F.B.P.; Mitroi, V.; Gondim, R.S.; Martins, E.S.P.R.; van Oel, P.R., 2024. It’s not all about drought: What “drought impacts” monitoring can reveal. International Journal of Disaster Risk Reduction, v. 103. https://doi.org/10.1016/j.ijdrr.2024.104338

Wang, C.; Shuai, C.; Chen, X.; Sun, J.; Zhao, B., 2024. Linking local and global: assessing water scarcity risk through nested trade networks. Sustainable Development, v. 33 (1), 1-18. http://dx.doi.org/10.1002/sd.3103.

Zhao, S.; Fu, R.; Anderson, M.L.; Chakraborty, S.; Jiang, J.H.; Su, H.; Gu, Y., 2022. Extended seasonal prediction of spring precipitation over the Upper Colorado River Basin. Climate Dynamics, v. 60 (5-6), 1815-1829. http://dx.doi.org/10.1007/s00382-022-06422-x.