ABES

Revista Brasileira de Ciências Ambientais
Brazilian Journal of Environmental Sciences

# Hybrid neural network for daily streamflow projection on a sub-seasonal timescale

## Rede neural híbrida para projeção diária de vazão em escala sub-sazonal

*Everton Santos Castro*[1], *Carlos Eduardo Sousa Lima*[1], *Antônio Duarte Marcos Junior*[1], *Cleiton da Silva Silveira*[1], *Francisco das Chagas Vasconcelos Junior*[2]

## ABSTRACT

Forecasts of hydrological variables are extremely important for water resource management in regions that are particularly vulnerable to the effects of climate change and climate variability. The aim of this study was to develop a hybrid neural network capable of producing daily streamflow forecasts on a sub-seasonal scale. This hybrid neural network consists of an artificial neural network with inputs preprocessed by the wavelet transform (WANN). The WANN was tested in the Três Marias, Sobradinho, and Retiro Baixo reservoirs, located in the São Francisco River Basin (SFRB). The obtained results show that WANN was highly accurate in short-term forecasts (7–28 days); however, for long-term forecasts (35 and 42 days), there was a significant drop in performance, especially during the transition periods to the rainy season and in the dry months. The comparison between the performance metrics of the WANN forecasts and the National Electricity System Operator (ONS) operational models for the Três Marias, Sobradinho, and Retiro Baixo basins showed that WANN outperformed all these models. The results obtained show that WANN is a valuable tool for addressing the complex and dynamic challenges of hydrology, making it essential for decision-making on water resource management.

**Keywords:** streamflow projection; artificial neural network; wavelet transform.

## RESUMO

Previsões de variáveis hidrológicas são de extrema importância para a gestão dos recursos hídricos em regiões particularmente vulneráveis aos efeitos das mudanças climáticas e variabilidade climática. Nesse contexto, o presente trabalho objetivou o desenvolvimento de uma rede neural híbrida capaz de produzir previsões diárias de vazão em escala subsazonal. Essa rede neural híbrida consiste em uma Rede Neural Artificial com entradas pré-processadas pela transformada Wavelet (RNAW). A RNAW foi testada nas bacias hidrográficas dos reservatórios de Três Marias, Sobradinho e Retiro Baixo, todas localizadas na Bacia Hidrográfica do rio São Francisco (BHSF). Os resultados obtidos mostram que a RNAW apresentou alta precisão nas previsões de curto prazo (7 a 28 dias); entretanto, para previsões de longo prazo (35 e 42 dias), houve uma queda significativa no desempenho, especialmente durante os períodos de transição para a estação chuvosa e nos meses secos. A comparação entre as métricas de desempenho das previsões da RNAW e dos modelos operacionais do Operador Nacional do Sistema Elétrico (ONS) para as bacias de Três Marias, Sobradinho e Retiro Baixo mostrou que a RNAW superou todos esses modelos. Os resultados obtidos indicam que a RNAW é uma ferramenta promissora para lidar com os complexos e dinâmicos desafios da hidrologia, conferindo-lhe grande valor no processo de tomada de decisão sobre a gestão dos recursos hídricos.

**Palavras-chave:** projeção de vazão; rede neural artificial, transformada wavelet.

[1]Universidade Federal do Ceará – Fortaleza (CE), Brazil.

[2]Fundação Cearense de Meteorologia e Recursos Hídricos – Fortaleza (CE), Brazil.

Corresponding author: Carlos Eduardo Sousa Lima – Hydraulic and Environmental Engineering Department (DEHA), Federal University of Ceará (UFC) – CEP: 60440-970 – Fortaleza (CE), Brazil. E-mail: eduardolima@alu.ufc.br

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

Funding: the present study was supported by the Fundação Cearense de Apoio ao Desenvolvimento Científica e Tecnológico (FUNCAP) and the Conselho Nacional de Desenvolvimento Científico e Tecnológico–Brasil (Project, No. 310286/2018-2: Uma análise sobre nexo: Clima, água e energia na bacia do rio São Francisco).

Received on: 03/08/2025. Accepted on: 12/18/2025.

https://doi.org/10.5327/Z2176-94782495

## Introduction

Understand that the climate variability in a given region plays a fundamental role in the context of climate change and in the search for effective public policies for adaptation and mitigation. It is through this kind of understanding that it becomes possible to take appropriate measures to reduce the impacts of these changes and guarantee the water and food security of local communities (Santana and Santos, 2020; Bevacqua et al., 2022).

In regions that are particularly vulnerable to the effects of climate change and climate variability, such as the Brazilian semiarid region, forecasting hydrological variables plays a crucial role in water resource management. Comprehending how climate fluctuations affect not only precipitation patterns but also water flows in reservoirs is essential to ensure water stability and promote sustainable development (Souza and Nascimento, 2022; Al-Juboori, 2023).

In this sense, the São Francisco River Basin (SFRB), with a drainage area corresponding to 8% of the Brazilian territory, exemplifies the need for such comprehension. Starting from Minas Gerais in Serra da Canastra, where the source of the São Francisco River is, to the Atlantic Ocean on the border of the states of Alagoas and Sergipe, this vast area integrates the Northeast and Southeast regions of the country. It runs through 507 municipalities in six states, constituting one of Brazil's 12 hydrographic regions (Bettencourt et al., 2016). The forecasting and monitoring of hydrological variables in the SFRB play a key role in ensuring the effective management of water resources and ensuring the sustainability and well-being of local populations.

Water resources management in the region encompasses a set of interconnected steps and strategies aimed at optimizing water use, preventing conflicts, and preserving the river ecosystem. This management is based on the collection, analysis, and continuous monitoring of hydrological and meteorological data throughout the basin. This includes measuring water levels and river flows, as well as assessing water quality and forecasting hydrological variables, allowing to understand water conditions and the identification of climate patterns and variations (Bettencourt et al., 2016; Ribeiro, 2017). In this context, improving the effectiveness of the decision support system is intrinsically linked to modeling meteorological, precipitation, and flow forecasts at different timescales (Al-Juboori, 2023; Nearing et al., 2024).

According to the 2021 report by the National Electric Energy Agency (ANEEL), seven large hydroelectric plants are installed in SFRB, which play a crucial role in supplying electricity to meet national demands, especially the Retiro Baixo, Três Marias, and Sobradinho plants. The National Electricity System Operator (ONS) uses flow projections made in SFRB as the basis for various strategic activities. Among these activities are the preparation of annual flood prevention planning, the formulation of guidelines for flood control operations, the generation of inflow scenarios, the continuous updating of technical data on hydroelectric projects, the improvement of information on hydraulic restrictions associated with these projects, and the evaluation of the reservoir filling process (ONS, 2015).

Despite the continuous progress observed in hydrological models, there is room for improvement in the forecasts. As new techniques and data emerge, the refinement of existing methodologies can offer significant gains in the accuracy and applicability of these essential tools for the management and preservation of water resources.

Among the various techniques used in hydrological and climatological modeling, artificial neural networks (ANNs) have stood out as a promising approach, constituting a valid mechanism to be applied in the prediction of hydrological variables. In this sense, this methodology has been widely discussed, and its results have been published in various studies (Freire et al., 2019; Höge et al., 2022; Li W et al., 2022; Li X et al., 2022; Zanial et al., 2023; Ekwueme, 2024). These networks have a high capacity for learning and generalizing from data input and are able to process information in a parallel and adaptive manner, making them a powerful tool for predicting hydrological and climatological variables (Freire et al., 2019; Sherstinsky, 2020).

ANNs are computational models inspired by the functioning of the human brain, composed of units called artificial neurons, organized into layers (input, hidden, and output). Each neuron processes the incoming signals, weighting them through adjustable parameters, and transmits the result to subsequent layers, allowing the model to learn complex patterns in the data. This learning process is carried out through optimization algorithms that iteratively adjust the weights to minimize the error between predictions and observed values. This characteristic gives ANNs a high capacity for learning and generalization from input data, enabling the recognition of nonlinear relationships and interdependencies among hydrological variables, which are frequently observed in natural processes (Hunt et al., 2022; Tian et al., 2024; Zhao et al., 2024).

Among ANN architectures, recurrent neural networks (RNNs), and particularly the long–short-term memory (LSTM) model, have gained increasing prominence in hydrological applications due to their ability to capture temporal dependencies and dynamics in historical series of streamflow, precipitation, and evapotranspiration. These networks incorporate memory mechanisms that allow them to retain information from relevant past states, preventing the "vanishing gradient" problem observed in traditional recurrent networks. As a result, LSTM networks have shown strong performance in hydrological time series forecasting tasks and are widely used in recent studies on streamflow modeling and flood prediction (Hunt et al., 2022; Santos et al., 2022; Arsenault et al., 2023).

The objective of this work is to develop a hybrid neural network capable of producing daily flow forecasts on a sub-seasonal scale. The studies related to streamflow forecasts conducted in this research include substantial modifications to a model previously proposed by Freire et al. (2019). The fundamental contributions of this study are the adoption of a different ANN architecture, the expansion of the wavelet options used to decompose the time series, the simultaneous use of all the decomposed series as input to the neural network, and the methodology used to evaluate the model's performance.

## Methodology

### Study area

The study area covers the basins of the Retiro Baixo, Três Marias, and Sobradinho reservoirs, all located within the SFRB. The SFRB is traditionally subdivided into four physiographic zones: Upper, Middle, Sub-middle, and Lower São Francisco (Bettencourt et al., 2016). The Três Marias basin, located in the Upper São Francisco, has an approximate area of 54,317 km². The Retiro Baixo basin, situated in the Middle São Francisco, covers about 13,390 km². The Sobradinho basin, with approximately 490,610 km², extends across the Middle and Sub-middle São Francisco, encompassing areas with distinct physiographic characteristics.

The climate regime in the region is marked by strong seasonality, with average annual precipitation around 1,400 mm and a rainy season concentrated between October and March. From May to September, monthly totals generally remain below 50 mm (Alves et al., 2005; Cavalcanti, 2016). The Três Marias and Retiro Baixo basins lie predominantly in tropical climate areas, characterized by wet summers and dry winters. The Sobradinho basin, in contrast, lies between two climatic regions: while its southern portion maintains tropical characteristics, its northern sector exhibits semiarid conditions, with greater irregularity in rainfall distribution throughout the year. Figure 1 shows the location of the reservoirs studied, as well as their respective river basins. The water bodies in the area of interest represent the São Francisco River and some of its most significant affluents.

### Hydrometeorological data

The study uses a data set made up of historical series of flows and rainfall in the aforementioned river basins. The rainfall series consists of daily data obtained from the National Agency for Water and Basic Sanitation (ANA). The streamflow series represents naturalized daily flows, acquired from the ONS. The series for the Três Marias and So-

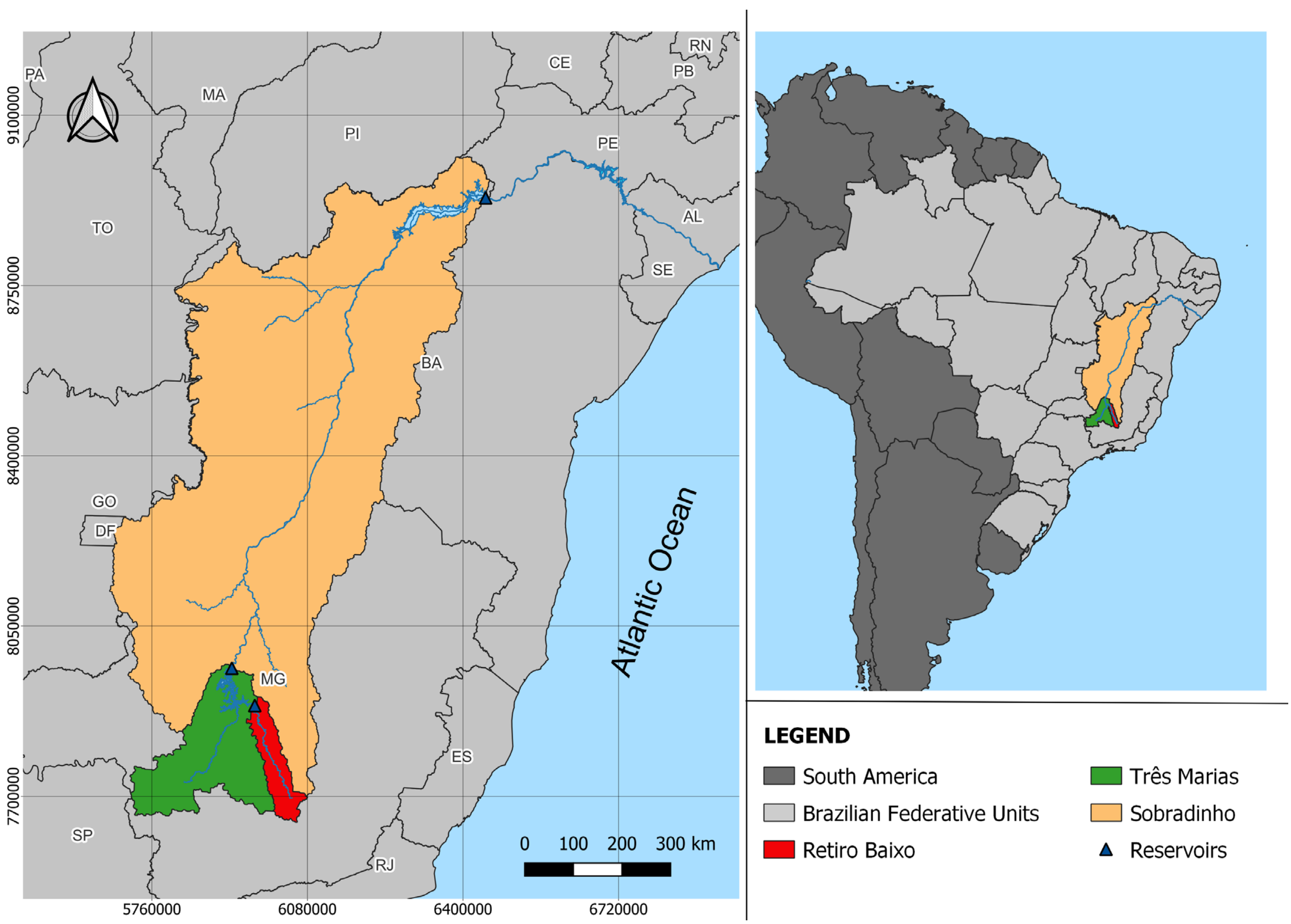


**Figure 1 – Studied reservoirs and their respective river basins.**

bradinho basins refer to the period from 1977 to 2015, while the series for the Retiro Baixo basin covers the period from 1987 to 2015.

## Streamflow forecast

To forecast streamflows in the Retiro Baixo, Três Marias, and Sobradinho river basins, a forecasting model based on neural networks was developed, with data preprocessing using the wavelet transform.

The data set was initially divided into two distinct groups for the training and testing phases. In the division, 90% of the data was destined for training, while the remaining 10% was reserved for the test phase. The training and test periods for each river basin area are shown in Table 1.

After this division, the streamflow and rainfall series from the training and test sets were decomposed separately using the wavelet transform. This decomposition process involves the use of a filter bank that separates the high- and low-frequency components of the signal.

The transformed wavelet can be defined as a sequence of applications of the high- and low-pass filters, giving rise to the wavelet decomposition tree generating approximation components and detail components, thus allowing a better visualization of the fundamental behavior of the series (Benedetto and Michael, 2021; Akujuobi, 2022).

Initially, the signal is subjected to a bank of high-pass and low-pass filters. The components resulting from high-pass filtering contain high frequencies and are called detail coefficients, commonly represented by $cD_1$. The components resulting from low-pass filtering are called approximation coefficients, represented by $cA_1$. These coefficients only provide a general view of these frequencies.

The approximation coefficients $cA_1$ are then passed through the high-pass and low-pass filters. In the same way as above, this procedure gives rise to two new coefficients, $cA_2$ and $cD_2$. The coefficients that have been subjected to the filter provide details, while the coefficients that have been subjected to the low-pass filter are subjected to the filters once again.

The process continues by complementing the approximation coefficients of the previous level, so that the signal is split into several low-resolution components. The decomposition process is applied until the last level of detail contains only one sample. However, the practical limit of decomposition is usually selected based on the nature of the signal (Freire et al., 2019; Benedetto and Michael, 2021; Akujuobi, 2022).

In this sense, the streamflow and rainfall series were subjected to the wavelet transform decomposition process, using the mother wavelets *haar*, *db10*, *sym8*, *coif5*, and *dmey*. The streamflow series was decomposed into five decomposition levels, while the rainfall series was decomposed into four levels for each of the aforementioned mother wavelets. The raw streamflow series, as well as the approximation coefficients of the streamflow and rainfall series, generated at each level of decomposition during the training period, were used as input for the WANN in order to calibrate the network and obtain the synaptic weight matrices. To demonstrate the effect of wavelet decomposition in extracting low- and high-frequency components present in hydrological time series, Figures 2 and 3 show the multiscale decomposition applied to the daily precipitation series (up to the fourth level) and to the daily streamflow series of the Retiro Baixo reservoir (up to the fifth level). In both cases, the Coiflet mother wavelet was used.

**Table 1 – Training and test periods for each river basins.**

| Basin | Training | Test |
|---|---|---|
| Retiro Baixo | 01/01/1987 - 02/04/2013 | 02/05/2013 - 12/31/2015 |
| Três Marias | 02/22/1971 - 07/06/2011 | 07/07/2011 - 12/31/2015 |
| Sobradinho | 02/22/1971 - 07/06/2011 | 07/07/2011 - 12/31/2015 |

Applying the transformed wavelet to the neural network's input data aims to identify the different frequency scales present in the signal, allowing the neural network to be exposed to the short-term and long-term patterns in the data. It should be noted that this study adopted a different approach to the input method for the ANN compared to the work by Freire et al. (2019). While in this model all the decomposed series were used simultaneously as input for the ANN, incorporating up to five mother wavelets, Freire et al. (2019) adopted a sequential approach, supplying one decomposed series at a time to the network. This resulted in different models for each wavelet. The approach adopted in this research seeks to comprehensively capture the different frequency scales present in the flow and precipitation data, thus enhancing the neural network's predictive capacity.

Forecasts were carried out for time horizons of 7, 14, 21, 28, 35, and 42 days. For all these time horizons, the forecasts were made considering:

1. streamflow and rainfall data for the current day;
2. streamflow and rainfall data for the previous 14 days; and
3. decomposed streamflow and rainfall series from item ii).

This information was used as input to the calibrated WANN to provide continuous streamflow forecasts over these time horizons.

## WANN architecture

The neural network architecture selected for this study is LSTM, recognized for its ability to capture long-term dependencies in sequential data.

Since there are no rigid guidelines for selecting the network's structural parameters, we opted for an iterative process of experimentation for this choice. Various configurations were tested and evaluated in terms of performance, and the configuration with the best results was selected. In this sense, the activation function chosen was the linear function, considering its proven benefits in similar problems (Freire et al., 2019; Sharma et al., 2022; Yilmaz et al., 2022). The other parameters used in the network architecture are shown in Table 2.

In the training process, the Adam optimizer was used together with the mean squared error (MSE) loss function to guide the adjustment of the network weights. To avoid possible overfitting problems,

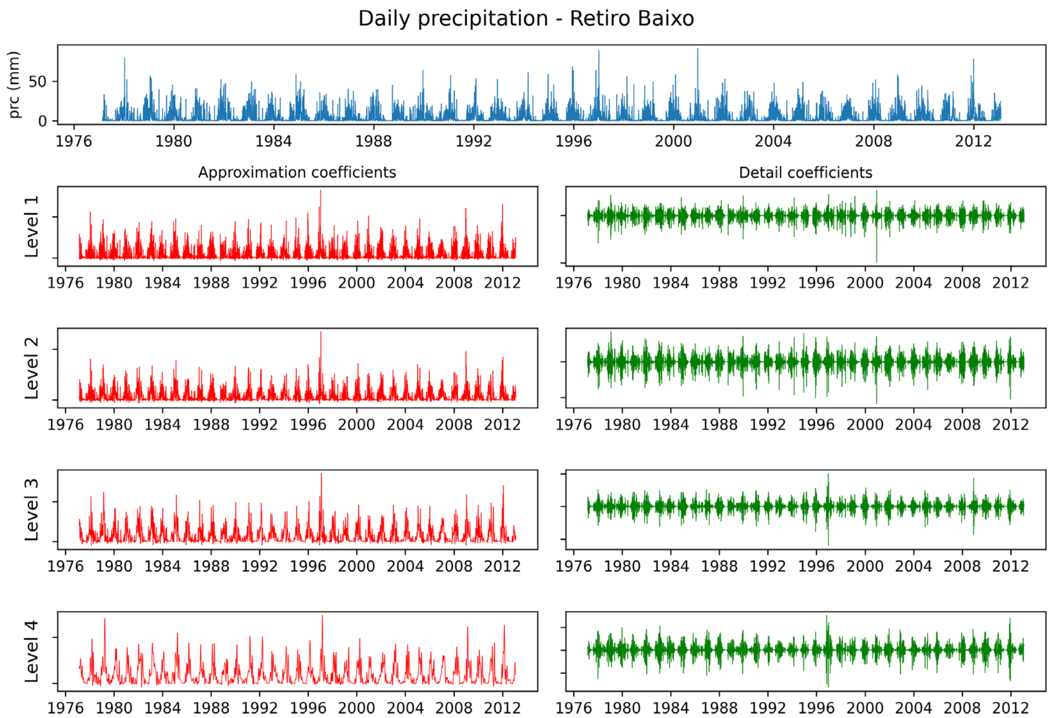


**Figure 2 – Decomposition of the rainfall series — wavelets (coif5).**

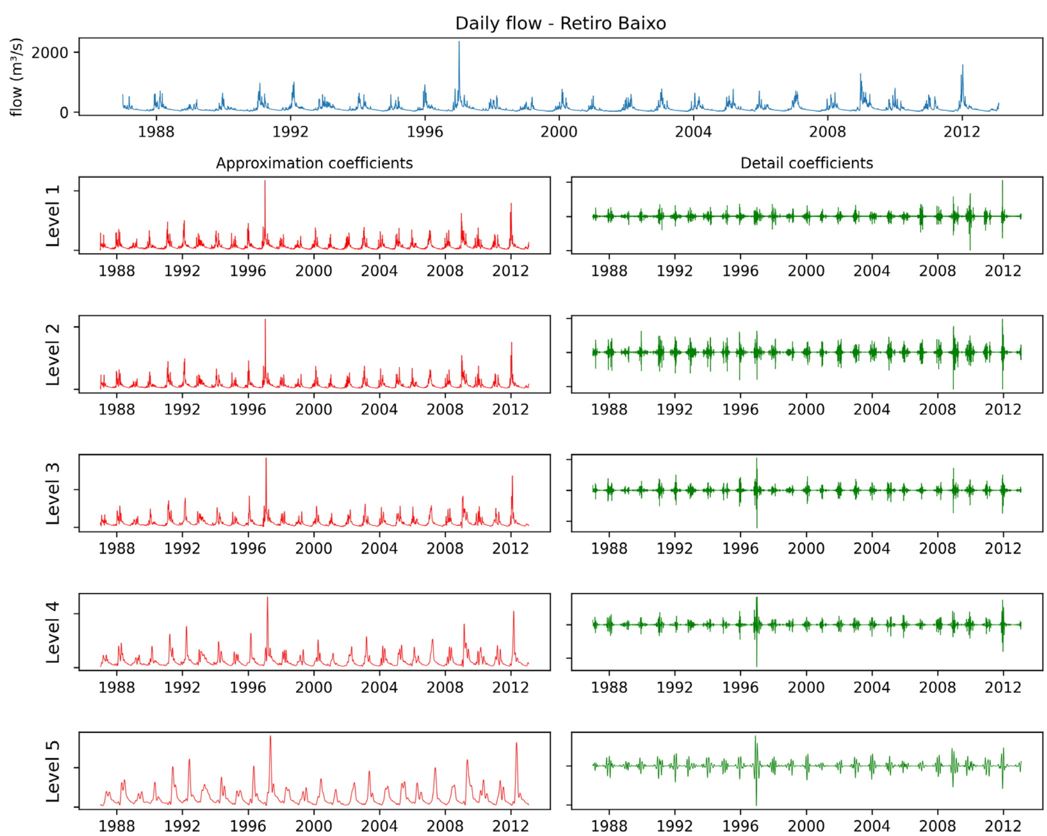


**Figure 3 – Decomposition of the flow series — wavelets (coif5).**

**Table 2 – WANN hyperparameters.**

| Parameters | WANN |
|---|---|
| Number of neurons in the input layer | 45 |
| Number of hidden layers | 4 |
| Number of neurons in each hidden layer | 200 |
| Loss function | MSE |
| Activation function | Linear |
| Optimizer | Adam |
| Maximum number of seasons | 200 |

a dropout layer was inserted after each hidden layer. These layers have the function of randomly deactivating a fraction of the neurons during training, which helps to avoid overfitting the training data.

### Forecast evaluation

Subsequently, after the calibration process, the model was evaluated using data from the test set. The WANN's performance was evaluated on a monthly scale, i.e., the daily streamflow forecasts were evaluated for each month of the year. To illustrate, the daily forecasts for January were compared only with the observed data for the same month, providing performance metrics for that month. We used three performance metrics to evaluate and quantify the accuracy and reliability of the daily streamflow forecast: the Nash–Sutcliffe efficiency (NSE), the Pearson correlation coefficient (r), and the mean absolute percentage error (MAPE). The NSE assesses the model's ability to reproduce the variability of the observed series, taking values between $-\infty$ and 1, with values closer to 1 indicating greater predictive skill. The correlation coefficient r measures the strength of the linear association between observed and predicted values, ranging from −1 to 1 and reflecting both the intensity and the direction of this relationship. The MAPE quantifies the average magnitude of relative errors in percentage terms, such that smaller values indicate greater estimation accuracy. More detailed descriptions of these metrics can be found in Wilks (2011).

In addition, the performance of the WANN projections was compared to the performance of the naturalized streamflow forecasts of the reservoirs that constitute the National Interconnected System (SIN), carried out by the ONS. These forecasts are especially used within the scope of the monthly operation program (MOP), with the aim of promoting the centralized dispatch of power plants in an optimized manner (ONS, 2015). Since January 2006, ONS has been using models that incorporate observed and forecast rainfall information to improve its streamflow forecasts. Among the models adopted are PREVIVAZ (Maceira et al., 1999) and CPINS (Paiva and Acioli, 2007).

The ONS forecasts for Sobradinho come from the CPINS model, while the forecasts for Retiro Baixo and Três Marias are made by the PREVIVAZ model. The estimates comprise weekly natural streamflow averages, projected one, two, and three weeks in advance, for the first quarter of 2014 and 2015. Statistics relating to the ONS forecasts were collected from the ONS's flow forecast performance analysis report, available for access at <sintegre.ons.org.br>. The performance metrics of WANN's weekly streamflow forecasts (NSE and MAPE) were compared with those of the ONS forecast metrics.

In contrast to the approach of Freire et al. (2019), who developed streamflow forecasting models for the Xingó reservoir at different temporal scales using multilayer perceptron (MLP) neural networks and hybrid versions based on the wavelet transform, the present study introduces four relevant methodological advances. First, the spatial scope of the modeling is substantially expanded, as the previous study focused exclusively on a single system within the São Francisco River Basin. Second, the MLP architecture employed by the authors is replaced by an LSTM network, which is more suitable for representing long-term dependencies and nonlinear patterns characteristic of hydrological time series. Third, the set of evaluated wavelets is broadened by incorporating different families and orders to enhance the multiscale decomposition of the streamflow series. Finally, a monthly evaluation of predictive skill is adopted, in contrast to the aggregated validation commonly used in the literature, allowing the identification of seasonal variations in performance and providing a more robust diagnostic for operational applications.

## Results and Discussions

Figures 4, 5, and 6 present the adjustments of the forecasts to the observed data, the residuals of the forecasts, and the distribution of the residuals of the WANN forecasts for Três Marias, Sobradinho, and Retiro Baixo, respectively, across the 7-, 14-, 21-, 28-, 35-, and 42-day horizons.

When comparing the observed flows (blue) and those predicted by the neural network (dashed red) for Três Marias (Figure 4) and Sobradinho (Figure 5), it can be seen that the two series are well aligned, indicating that the model is able to capture the flow dynamics with good accuracy. For forecast horizons of 7–14 days, the residual errors are mostly close to zero, suggesting that the model has a good predictive performance for short-term horizons. The scatter plot of the residuals as a function of the predicted values shows a distribution of residuals around zero, with no obvious patterns of heteroscedasticity. The distribution of residuals shown in the histogram confirms that most of the errors are concentrated around zero, reinforcing the model's accuracy for shorter horizons.

For forecast horizons longer than 35 days, although significant peaks are well captured by the model, some discrepancies can be observed, especially in peak flow events, which may indicate that the model may have difficulties in predicting abrupt changes or extreme events. The pattern shown by the residuals confirms that there is a residual increase as the predicted values increase, indicating that the model has a bias toward underestimating high flow values.

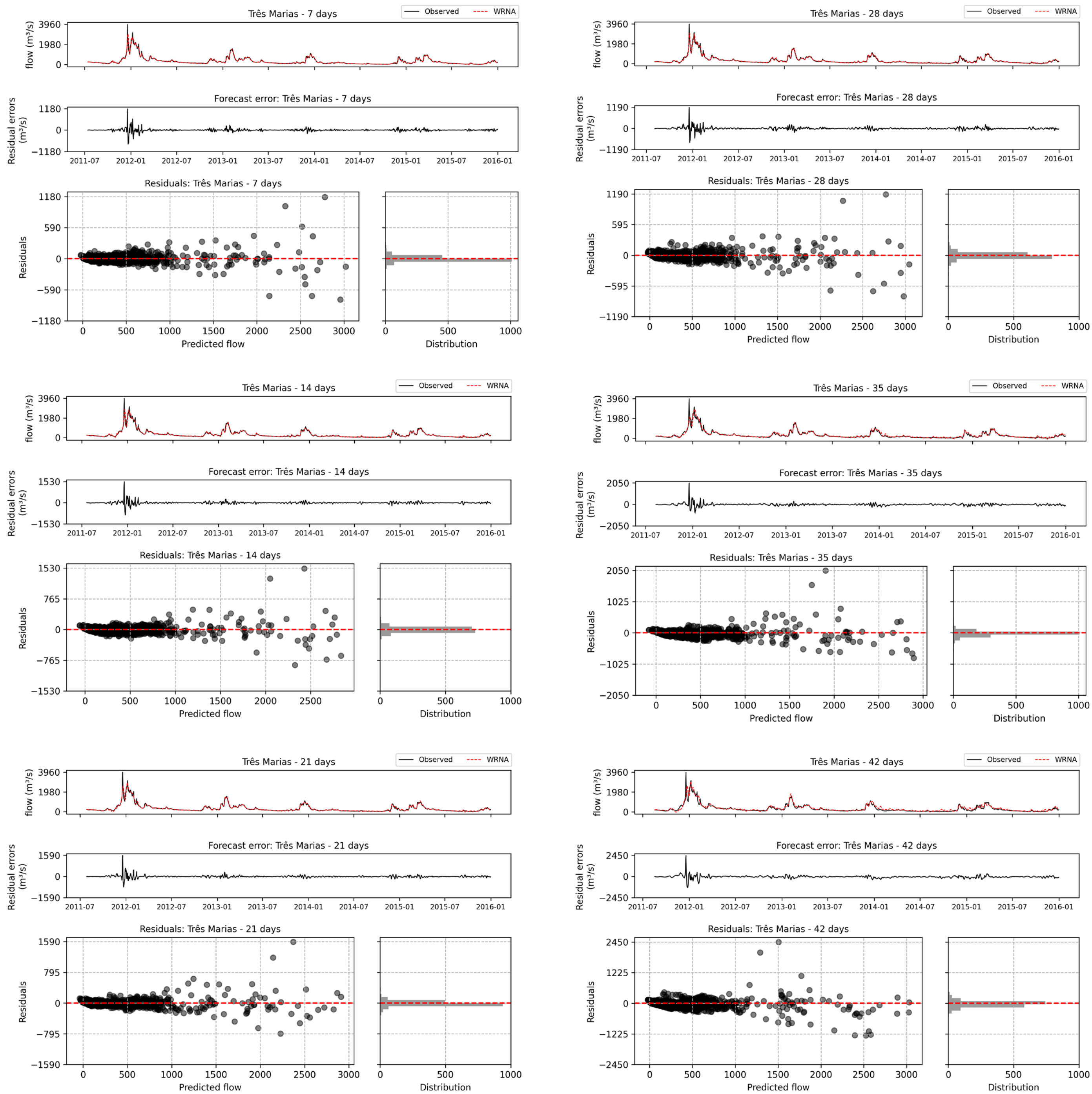


**Figure 4 – WANN forecasts for Três Marias across 7-, 14-, 21-, 28-, 35-, and 42-day horizons.**

For the forecasts made in the Retiro Baixo basin (Figure 6), the results show that the neural network was able to capture the trend and flow peaks with good accuracy, especially for shorter forecast horizons, 7 and 14 days, as indicated by the proximity between the observed and forecast flow lines. For forecast horizons of up to 21 days, the forecast error graph indicates that the residual errors are relatively small, remaining mostly within the range of ±65 m³/s, with some larger peaks. Analysis of the residuals reveals a random distribution around zero, with no obvious pattern, indicating that the model does not show significant systematic bias. The histogram of the residuals reinforces this observation, showing an approximately symmetrical distribution around zero.

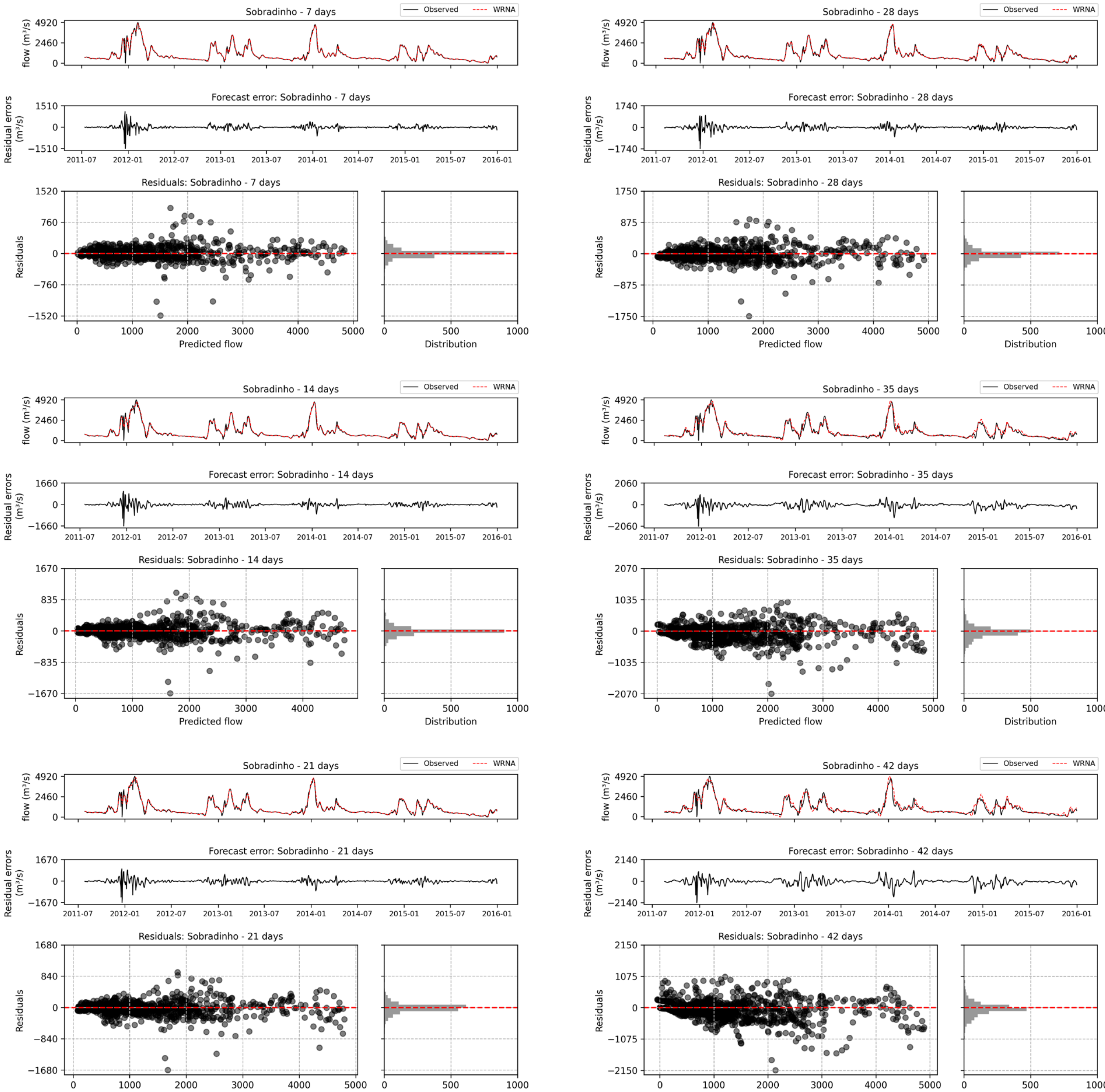


**Figure 5 – WANN forecasts for Sobradinho across 7-, 14-, 21-, 28-, 35-, and 42-day horizons.**

For forecasts with horizons longer than 28 days, the neural network begins to show greater difficulty in capturing the streamflow dynamics, especially in Retiro Baixo. Although the general trends are still well captured, the deviations between observed and forecast flows become more evident, especially during peak periods. The residual error graph shows higher variability and amplitude of the errors, with some points exceeding ±100 m³/s. Analysis of the residuals shows higher dispersion around zero and a more evident pattern of underestimation and overestimation of flows. The histogram of the residuals shows a wider and slightly skewed distribution, suggesting a decrease in the model's accuracy as the forecast horizon increases.

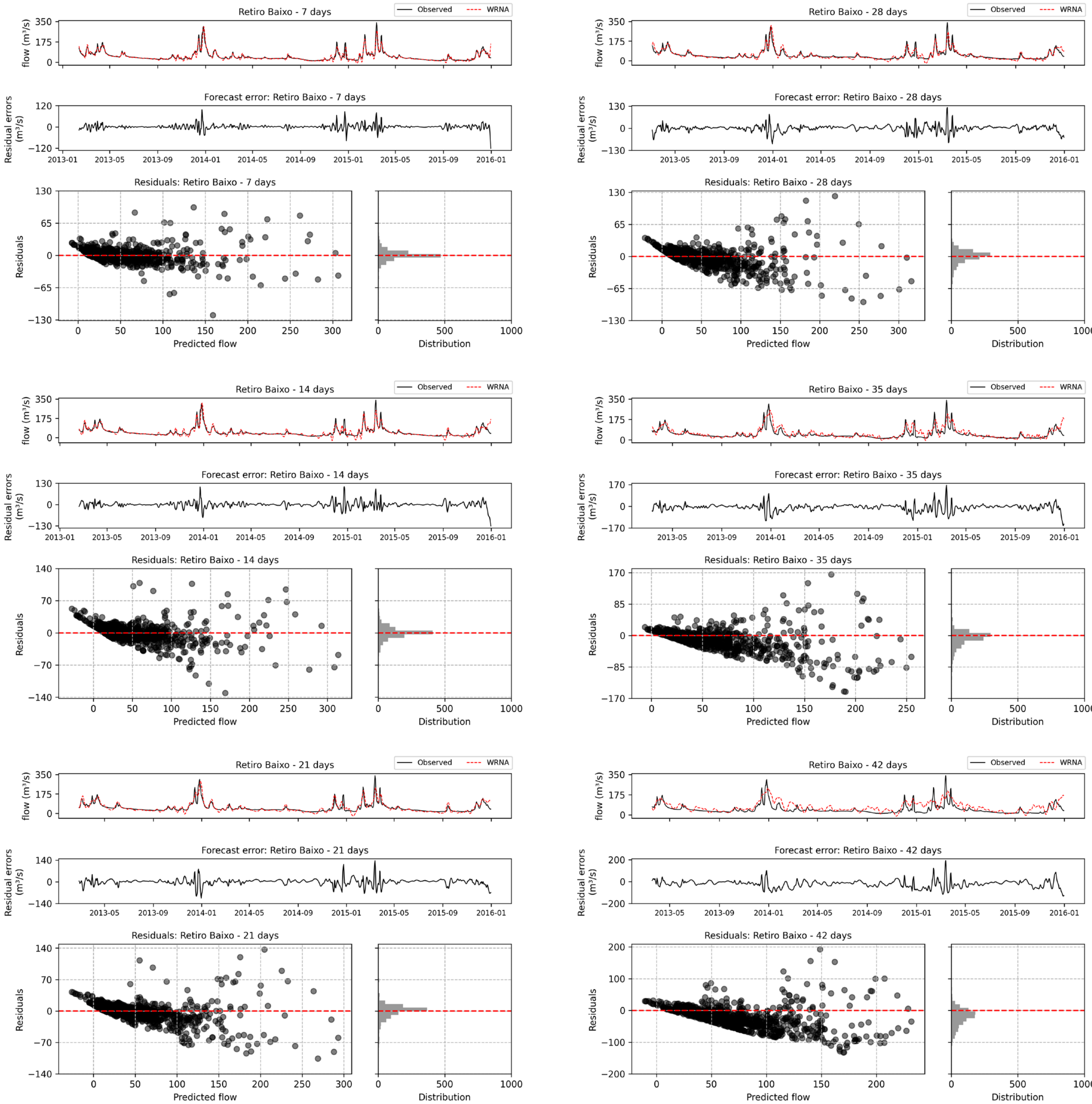


**Figure 6 – WANN forecasts for Retiro Baixo across 7-, 14-, 21-, 28-, 35-, and 42-day horizons.**

Figures 7 to 9 show the results, by month, of the r, NSE, and MAPE for the predictions made by WANN during the testing phase in the Três Marias, Sobradinho, and Retiro Baixo river basins, respectively.

Analysis of the statistical indexes for the Três Marias, Sobradinho, and Retiro Baixo basins reveals similar patterns, highlighting the accuracy of WANN's short-term forecasts. For forecast horizons of 7–28 days, the NSE and r indexes are high throughout the year, indicating good model performance. However, for longer forecasts (35 and 42 days), there is a sharp drop in performance, especially during certain critical months, which vary for each basin. For Retiro Baixo, except

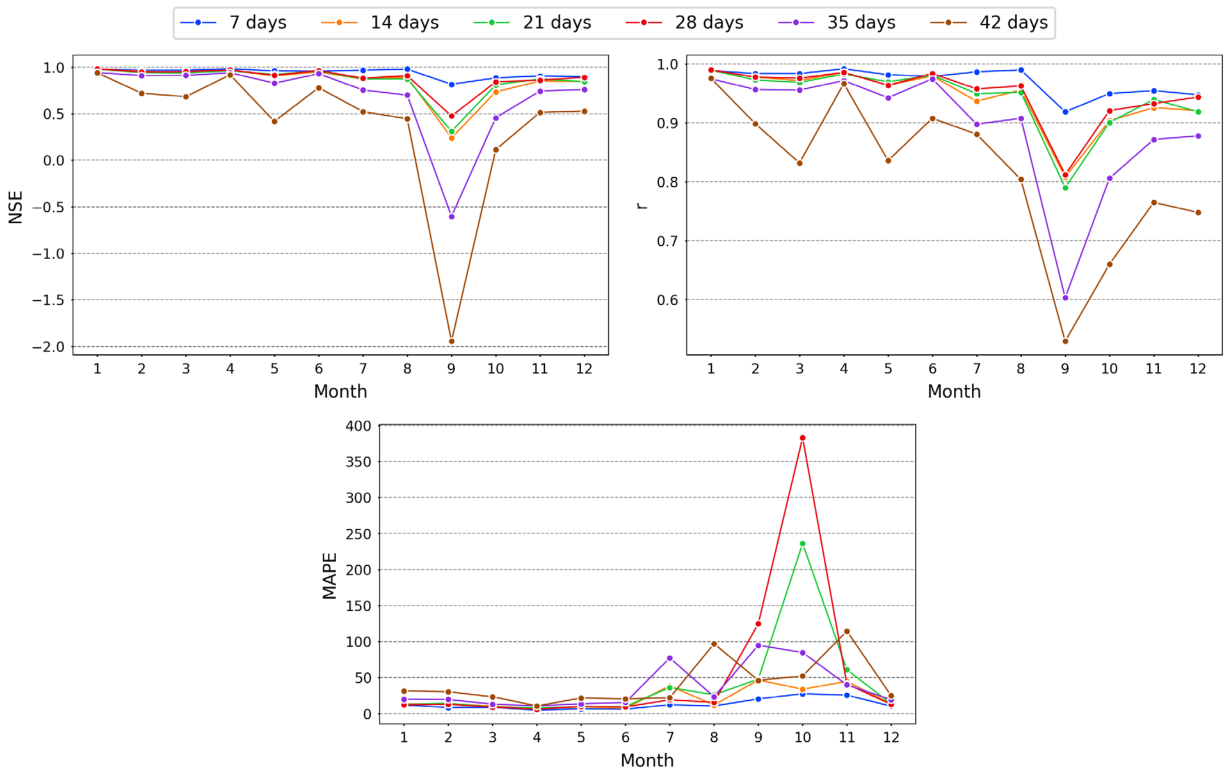


**Figure 7 – WANN forecast performance metrics for Três Marias.**

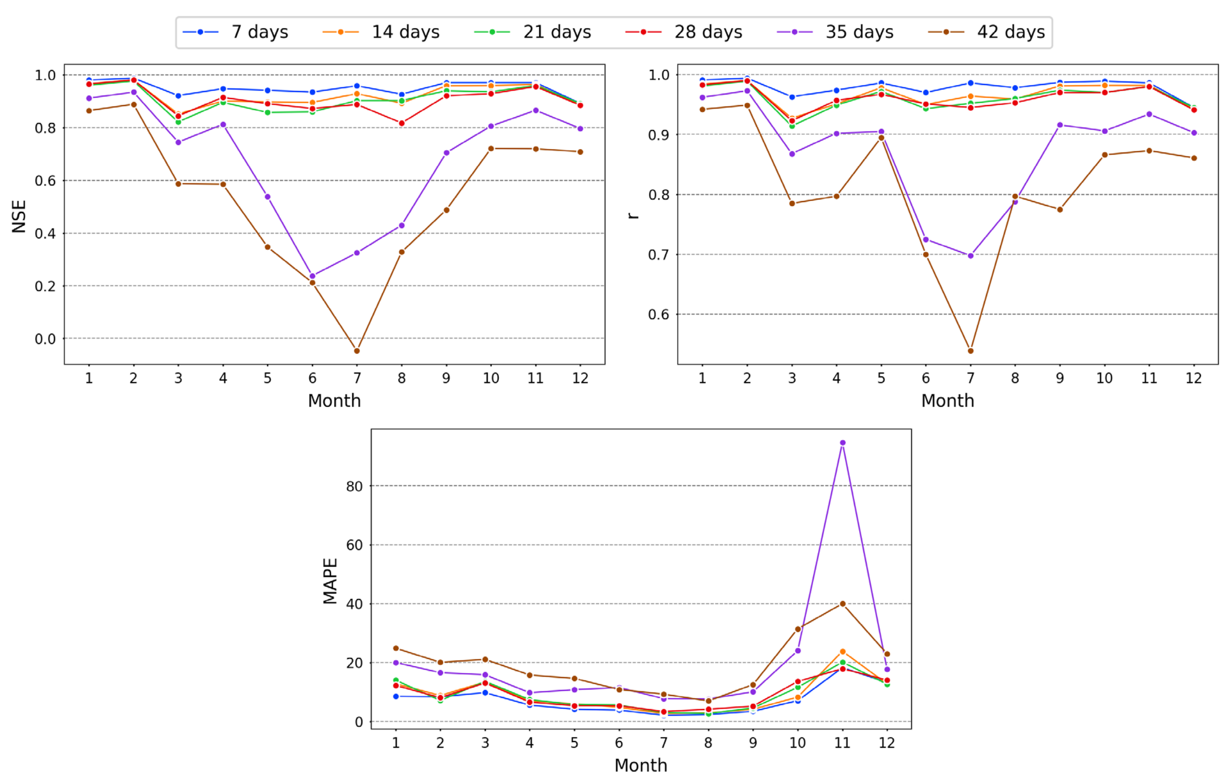


**Figure 8 – WANN forecast performance metrics for Sobradinho.**

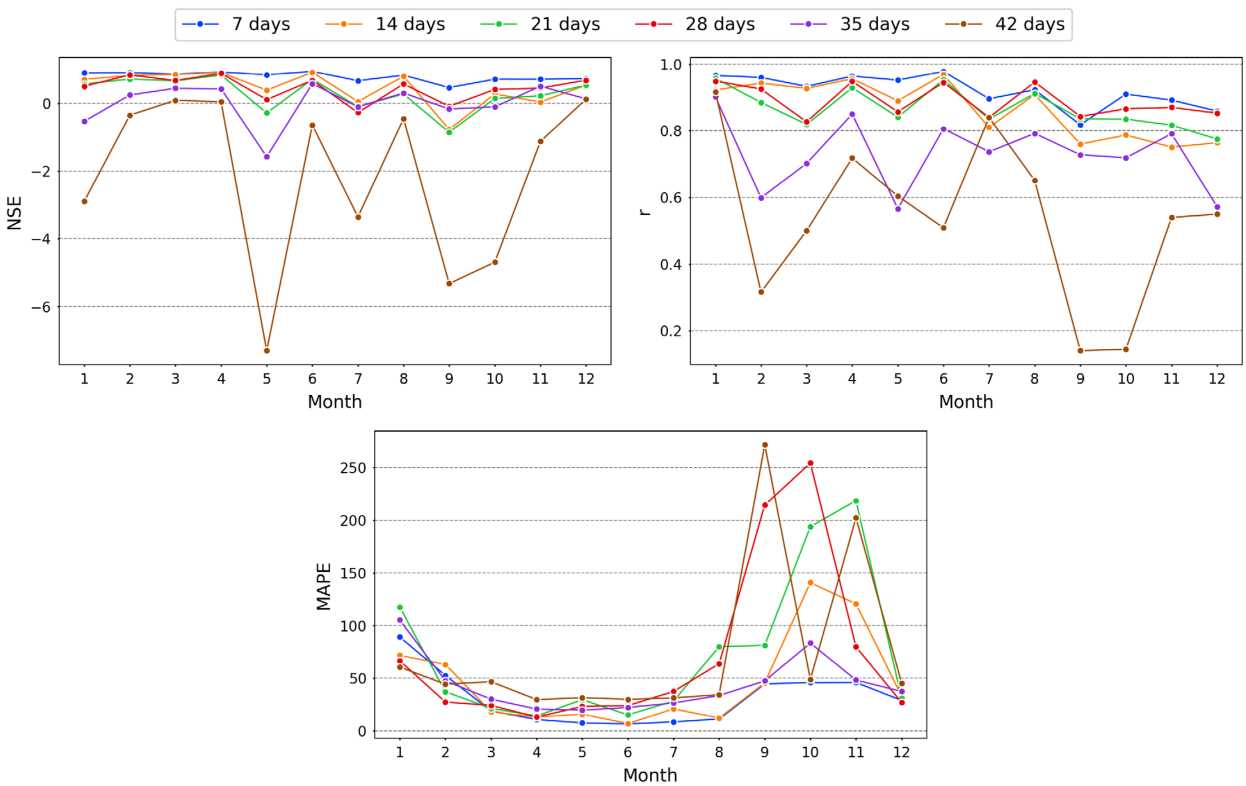


**Figure 9 – WANN forecast performance metrics for Retiro Baixo.**

at the 7-day forecast horizon, there is a sharp decline in NSE values during May, July, and September.

For the Três Marias basin (Figure 7), the NSE and r indexes are slightly lower in the months from April to October, particularly between June and July, due to lower water variability. During the months of November to March, which coincide with the rainy season in the region, the model's performance improves. However, the months of September and October, which precede the rainy season, present significant challenges, resulting in lower performance.

In the Sobradinho basin (Figure 8), the MAPE shows higher errors in November, coinciding with the start of the rainy season. During this period, flow variability increases due to more intense and frequent rainfall, resulting in flow peaks that are difficult to predict. The performance obtained for Sobradinho at the 7-day forecast horizon is similar to the results reported by Freire et al. (2019) and Saraiva et al. (2021) for the same reservoir and the same forecast horizon. Considering the entire simulation period analyzed by those authors, the NSE achieved was consistently above 0.90, comparable to that shown in Figure 8.

Retiro Baixo (Figure 9) shows a similar pattern to that of Sobradinho, with lower forecast errors during the months associated with the dry season (April to October) and a significant increase in MAPE in November and December, reflecting the difficulties associated with the higher variability of flows at the start of the rainy season.

Quedi et al. (2024) conducted hydrological forecasts with horizons of up to 6 weeks by coupling sub-seasonal climate prediction models with a hydrological–hydrodynamic model on a continental scale. The performance of the streamflow forecasts for the Sobradinho and Três Marias hydropower plants was inferior to that obtained with WANN across all forecast horizons considered. For Retiro Baixo, except at the 35- and 42-day horizons, the performance of WANN was also superior to that reported by Quedi et al. (2024).

Equally important, the WANN model demonstrated excellent performance for forecast horizons of up to 28 days in the Três Marias (Figure 7) and Sobradinho (Figure 9) reservoirs throughout the year, with the exception of September in the former, when performance reached a good level. At this horizon, the WANN model is able to rival even well-established monthly hydrological forecasting models in Brazil, such as the Soil Moisture Accounting Procedure (SMAP). As reported by Silva et al. (2020) and Silva et al. (2021), SMAP achieved metrics comparable to those obtained by WANN at the 28-day horizon for both reservoirs. However, unlike SMAP at the monthly scale, WANN provides forecasts at a daily temporal resolution, which offers additional benefits for the daily operation of the SIN.

Figure 10 shows the NSE and MAPE index performances for the forecasts made by WANN and ONS in the Três Marias, Retiro Baixo, and

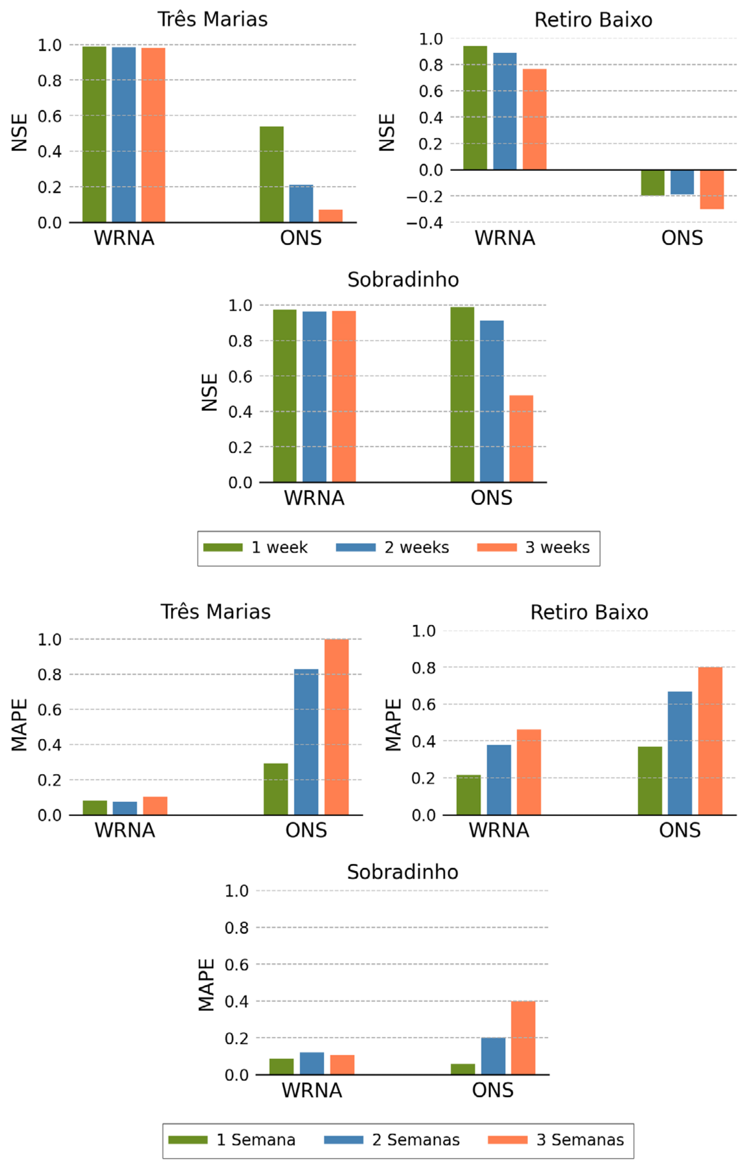


**Figure 10 – Performance metrics for ONS and WANN weekly streamflow forecasts for the Três Marias, Sobradinho, and Retiro Baixo basins in the first quarter of 2015.**

Sobradinho basins. These indexes represent the comparison between the observed weekly natural flow averages and the weekly natural flow averages forecast 1, 2, and 3 weeks in advance for the first quarter of 2015.

Regarding the forecasts made for Retiro Baixo and Três Marias, it can be seen that the WANN performed better than the forecasts made by the ONS at all forecast horizons. The NSE values for WANN were very close to one, indicating good agreement between the forecasts and the observed data. On the other hand, the models used by the ONS registered negative NSE values at Retiro Baixo, suggesting less accuracy in the forecasts. The same pattern was observed for the MAPE index, where WANN obtained lower values compared to the forecasts made by ONS.

In the Sobradinho basin, both WANN and ONS showed favorable results in terms of agreement with the observed data. Both models managed to adequately capture the flow patterns for the different forecast horizons. However, it was observed that ONS performed slightly better for forecasts made 1 week ahead, while WANN performed slightly better for forecasts 2 and 3 weeks ahead.

One notable aspect is that WANN's performance indexes remained consistent in all the analyzed basins, even for forecasts with more distant horizons (2 and 3 weeks). This stability in the results contrasts with the difficulties observed in previous analyses, where the WANN showed problems predicting flows for more distant horizons. The im-

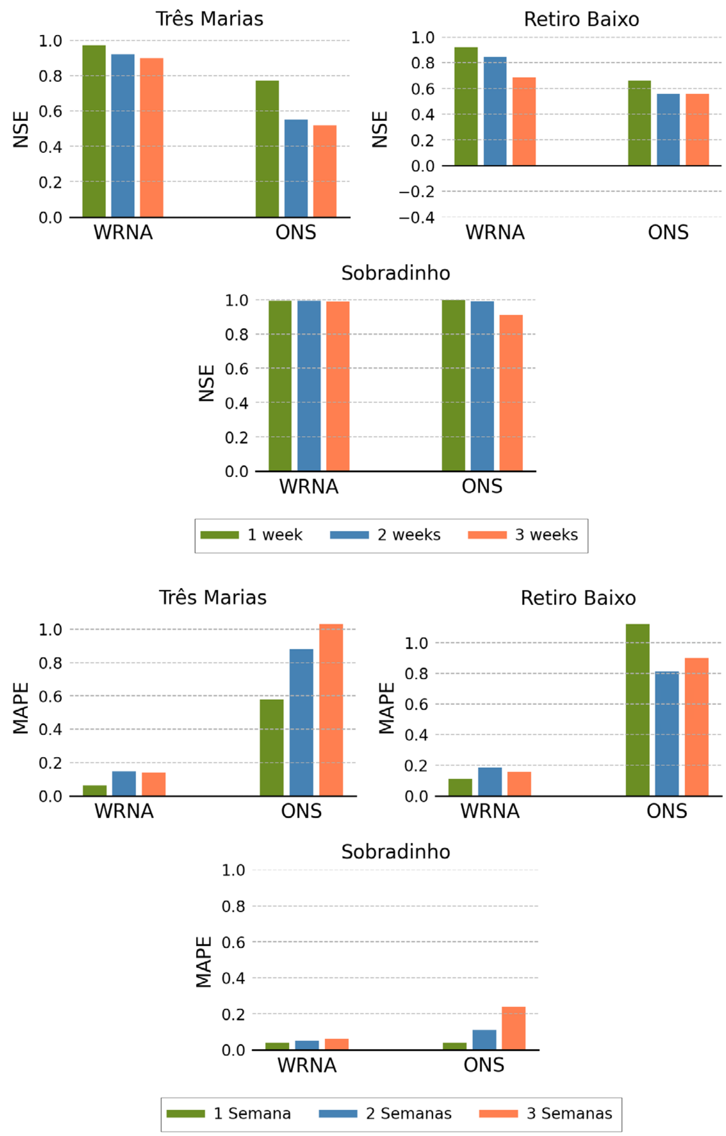


**Figure 11 – Performance metrics for ONS and WANN weekly streamflow forecasts for the Três Marias, Sobradinho, and Retiro Baixo basins in the first quarter of 2014.**

provement can be attributed to the fact that the current analysis considers the weekly average of flows, rather than the daily flow, which mitigates the errors.

Figure 11 shows the same evaluations as Figure 10, but for the first quarter of 2014. It can be seen that in the regions of the Três Marias and Retiro Baixo basins, WANN once again performed better than the projections made by the ONS. In the case of Sobradinho, the WANN and ONS estimates performed equally well.

The results obtained demonstrate significant gains in the predictive performance of the WRNA, consistent with studies in the literature that highlight the benefits of multiscale decomposition for hydrological forecasting. The application of the wavelet transform prior to neural modeling has been widely recognized for its ability to separate components of different frequencies, reducing high-variability noise, and enhancing relevant hydrological patterns. This strategy improves the representation of input signals and enhances network learning, as demonstrated by Khazaeiathar and Schmalz (2025), who reported a 66.43% reduction in root mean square error (RMSE) and a 45.49% reduction in MAPE for the forecasting model with wavelet-based decomposition compared to the LSTM. Similar results were observed by Agarwal et al. (2022), Ougahi and Rowan (2025), and Tuğrul and Hinis (2025), who found substantial improvements in NSE, r, and the ability to capture trends in basins with strong intra-annual variability. The convergence of these findings supports the robustness of hybrid methods in hydrological contexts dominated by multiscale phenomena.

The monthly performance assessment enabled the identification of seasonal variations in the model's predictive skill, avoiding the dilution of critical errors that occurs when metrics are calculated only in an aggregated manner. The hydrological verification literature highlights the importance of stratified approaches: Bellier et al. (2023) showed that validation using monthly windows enhances the operational interpretability of the results; Hurkmans et al. (2023) arguedd that temporal segmentation distinguishes periods governed by more deterministic hydrological processes from those influenced by greater stochastic variability; and Ferreira et al. (2023) emphasized that averaged analyses may conceal structural weaknesses in months associated with greater meteorological uncertainty.

The reduction in accuracy for the 35- and 42-day horizons follows what is widely documented in medium-range hydrological forecasting studies. The sub-seasonal horizon, which encompasses this period, is traditionally difficult to predict because the influence of initial conditions decreases and predictability becomes dependent on large-scale climate phenomena (Ferreira et al., 2023; Quedi et al., 2024).

## Conclusions

This study evaluated the performance of daily streamflow forecasts in three important Brazilian river basins of the SIN, located in the São Francisco River Basin (SFRB): Três Marias, Sobradinho, and Retiro Baixo. These predictions were made over time horizons of 7, 14, 21, 28, 35, and 42 days using an artificial neural network with inputs preprocessed by the wavelet transform (WANN).

The WANN was highly accurate in short-term forecasts (7–28 days), as indicated by the high NSE and r indexes. However, for long-term forecasts (35 and 42 days), there was a significant drop in performance, especially during the transition periods to the rainy season and in the dry months.

The performance metrics of the WANN forecasts were compared with the performance metrics obtained from ONS operational models. The WANN performed better overall than the ONS operational models. The difference in performance was especially significant in the Três Marias and Retiro Baixo basins, where WANN showed better ability to accurately represent the observed flows. In addition, WANN proved to be more efficient in long-term forecasts when considering weekly average flows, maintaining more stable and reliable performance indexes compared to those obtained by the ONS.

The model presented here proves to be a promising tool for dealing with the complex and dynamic challenges of hydrology, making it a valuable ally in water resources management.

## Authors' Contributions

**Castro**, E.S.: conceptualization; data curation; formal analysis; investigation; methodology; software; validation; visualization; writing — original draft, writing — review & editing. **Lima**, C.E.S.: formal analysis; methodology; software; validation; visualization; writing — review & editing. **Marcos Junior**, A.D.: formal analysis; methodology; software; validation; visualization; writing — review & editing. **Silveira**, C.S.: funding acquisition; project administration; supervision, validation; visualization; writing — review & editing. **Vasconcelos Júnior**, F.C.: validation; visualization; writing — review & editing.

## References

Agarwal, S.; Roy, P.; Choudhury, P.; Debbarma, N., 2022. Estudo comparativo sobre a previsão de vazão de rios usando GMNN e GMNN baseado em wavelet. Journal of Water and Climate Change, v. 13 (9), 3323-3337. https://doi.org/10.2166/wcc.2022.226.

Akujuobi, C.M., 2022. Wavelets and wavelet transform systems and their applications. Springer International Publishing. Berlin/Heidelberg, Germany.

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.

Arsenault, R.; Martel, J.L.; Brunet, F.; Brissette, F; Mai, J., 2023. Continuous streamflow prediction in ungauged basins: long short-term memory neural networks clearly outperform traditional hydrological models. Hydrology and Earth System Sciences, v. 27 (1), 139-157. https://doi.org/10.5194/hess-27-139-2023.

Bellier, J.; Whitin, B., Scheuerer, M.; Brown, J.; Hamill, T.M., 2023. A multi-temporal-scale modulation mechanism for the postprocessing of precipitation ensemble forecasts: Benefits for streamflow forecasting. Journal of Hydrometeorology, v. 24 (4), 659-673. https://doi.org/10.1175/JHM-D-22-0119.1.

Benedetto, J.J.; Michael, W.F., 2021. Wavelets: mathematics and applications. CRC Press, [S.l.].

Bettencourt, P.; Fulgêncio, C.; Grade, M.; Alcobia, S.; Monteiro, J.P.; Oliveira, R.; Leitão, J.C.; Leitão, P.C.; Fernandes, P.A.; Sousa, S.; Brites, S.; Fernandes, J.; Simões, J.; Scarton, M.; Santiago, E.; Aguiar, R.; Giffoni, M.; Melo, F.; Paes, A.C., 2016. Plano de recursos hídricos da bacia hidrográfica do rio São Francisco. Recursos Hidricos, v. 37 (1), 73-80. https://doi.org/10.5894/rh37n1-cti3.

Bevacqua, E.; Zappa, G.; Lehner, F.; Zscheischler, J., 2022. Precipitation trends determine future occurrences of compound hot–dry events. Nature Climate Change, v. 12 (4), 350-355. https://doi.org/10.1038/s41558-022-01309-5.

Ekwueme, B.N., 2024. Deep neural network modeling of river discharge in a tropical humid watershed. Earth Science Informatics, v. 17 (2), 1161-1177. https://doi.org/10.1007/s12145-023-01219-w.

Ferreira, N.C.; Chou, S.C.; Dereczynski, C., 2023. Evaluation of subseasonal precipitation simulations for the Sao Francisco River Basin, Brazil. Climate, v. 11 (11), 213. https://doi.org/10.3390/cli11110213.

Freire, P.K.M.M.; Santos, C.A.G.; Silva, G.B.L., 2019. Analysis of the use of discrete wavelet transforms coupled with ANN for short-term streamflow forecasting. Applied Soft Computing, Elsevier, v. 80, 494-505. https://doi.org/10.1016/j.asoc.2019.04.024.

Höge, M.; Scheidegger, A.; Baity-Jesi, M.; Albert, C.; Fenicia, F., 2022. Improving hydrologic models for predictions and process understanding using neural ODEs. Hydrology and Earth System Sciences, v. 26 (19), 5085-5102. https://doi.org/10.5194/hess-26-5085-2022.

Hunt, K.M.; Matthews, G.R.; Pappenberger, F.; Prudhomme, C., 2022. Using a long short-term memory (LSTM) neural network to boost river streamflow forecasts over the western United States. Hydrology and Earth System Sciences, v. 26 (21), 5449-5472. https://doi.org/10.5194/hess-26-5449-2022.

Hurkmans R.T,; Van Den Hurk, B.A; Schmeits, M.; Wetterhall, F. Pechlivanidis, I.G., 2023. Seasonal streamflow forecasting for fresh water reservoir management in the Netherlands: An assessment of multiple prediction systems. Journal of hydrometeorology, v. 24 (7), 1275-1290. https://doi.org/10.1175/JHM-D-22-0107.1

Khazaeiathar, M.; Schmalz, B., 2025. Advancing Flood Forecasting With Wavelet-LSTM: The Role of Nonlinearity in Discharge Prediction. Journal of Flood Risk Management, v. 18 (4), e 70148. https://doi.org/10.1111/jfr3.70148.

Li, W.; Pan, B.; Xia, J.; Duan, Q., 2022. Convolutional neural network-based statistical post-processing of ensemble precipitation forecasts. Journal of Hydrology, v. 605, 127301. https://doi.org/10.1016/j.jhydrol.2021.127301.

Li, X.; Xu, W.; Ren, M.; Jiang, Y.; Fu, G., 2022. Hybrid CNN-LSTM models for river flow prediction. Water Supply, v. 22 (5), 4902-4919. https://doi.org/10.2166/ws.2022.170.

Maceira, M.; Damázio, J.; Ghirardi, A.; Dantas, H., 1999. Periodic arma models applied to weekly streamflow forecasts. In: IEEE. Power Tech Budapest 99. Abstract Records. (Cat. No. 99EX376). [S.l.], p. 86.

Nearing, G.; Cohen, D.; Dube, V.; Gauch, M.; Gilon, O.; Harrigan, S.; Matias, Y., 2024. Global prediction of extreme floods in ungauged watersheds. Nature, v. 627 (8004), 559-563. https://doi.org/10.1038/s41586-024-07145-1.

Operador Nacional do Sistema Elétrico (ONS), 2015. ONS Relatório Anual de Avaliação das Previsões de Vazões – 2015. [S.l.], ONS.

Ougahi, J.H; Rowan, J.S., 2025. Enhanced streamflow forecasting using hybrid modelling integrating glacio-hydrological outputs, deep learning and wavelet transformation. Scientific Reports, v. 15 (1), 2762. https://doi.org/10.1038/s41598-025-87187-1.

Paiva, L.F.G.; Acioli, G.C.L., 2007. Previsão de vazões com o modelo cpins – cálculo e previsão de vazões naturais e incrementais a sobradinho. In: Anais do I Workshop de previsão de vazões. ONS, Rio de Janeiro.

Quedi, E.; Fan, F.; Siqueira, V.; Collischonn, W.; Paiva, R.; Petry, I.; Gama, C.; Silveira, R.; Paranhos, C.; Freitas, C., 2024. Sub-seasonal streamflow forecasts for hydropower dams in the Brazilian Eletrical Interconnected System. RBRH, v. 29, e7. https://doi.org/10.1590/2318-0331.292420230109.

Ribeiro, S.C., 2017. Caracterização geoambiental da sub-bacia do rio salgado na mesorregião sul cearense–parte i–clima e arcabouço geológico. Geoconexões, v. 1, 4-16. https://doi.org/10.15628/geoconexoes.2017.6290.

Santana, A.S.; Santos, G.R., 2020. Impactos da seca de 2012-2017 na região semiárida do Nordeste: notas sobre a abordagem de dados quantitativos e conclusões qualitativas. Boletim Regional, Urbano e Ambiental (BRUA), v. 22 (22), 119-129. https://doi.org/10.38116/brua22art9.

Santos, A.L.L., 2022. Previsão de vazão afluente da UHE-Tucuruí por redes neurais recorrentes LSTM. Orientador: Raphael Barros Teixeira. 2022. [9], 44 f. Undergraduate Thesis (Bachelor's Degree in Electrical Engineering) – Faculdade de Engenharia Elétrica, Campus Universitário de Tucuruí, Universidade Federal do Pará, Tucuruí. Retrieved 2025-10-20, from https://bdm.ufpa.br/handle/prefix/4578.

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 (107081), 107081. https://doi.org/10.1016/j.asoc.2021.107081.

Sharma, P.; Singh, S.; Sharma, S.D., 2022. Artificial neural network approach for hydrologic river flow time series forecasting. Agricultural Research, v. 11 (3), 465-476. https://doi.org/10.1007/s40003-021-00585-5.

Sherstinsky, A., 2020. Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network. Physica D: Nonlinear Phenomena, v. 404, 132306. https://doi.org/10.1016/j.physd.2019.132306.

Silva, M.V.M.; Silveira, C.S.; Costa, J.M.F.; Martins, E.S.P.R; Júnior, F.C.V, 2021. Projection of climate change and consumptive demands projections impacts on hydropower generation in the São Francisco River Basin, Brazil. Water, v. 13 (3), 332. https://doi.org/10.3390/w13030332.

Silva, M.V.M; Silveira, C.S; Costa, J.M.F; Martins, E.S.P.R; Júnior, F.C.V., 2020. Projections of climate change in streamflow and affluent natural energy in the Brazilian hydroelectric sector of CORDEX models. RBRH, v. 25. https://doi.org/10.1590/2318-0331.252020200020.

Souza, T.S.D.; Nascimento, P.D.S., 2022. Sequential climatic water balance applied in the hydrographic region of Paraguaçu, in Brazil. Sustainable Water Resources Management, v. 8 (1), 18. https://doi.org/10.1007/s40899-022-00608-1.

Tian, B.; Kong, X.; Wu, G., 2024. The application of the LSTM neural networks on the hydrology forecast. In: Proceedings of the 2024 6th International Conference on Pattern Recognition and Intelligent Systems. pp. 93-97. https://doi.org/10.1145/3689218.3689233.

Tuğrul, T.; Hinis M.A., 2025. Performance enhancement of models through discrete wavelet transform for streamflow forecasting in Çarşamba River, Türkiye. Journal of Water and Climate Change, v. 16 (2), 736-756. https://doi.org/10.2166/wcc.2025.709.

Wilks, D.S., 2011. Statistical methods in the atmospheric sciences. Academic Press, [S.l.]. v. 100.

Yilmaz, M., Tosunoğlu, F., Kaplan, N.H., Üneş, F., Hanay, Y.S., 2022. Predicting monthly streamflow using artificial neural networks and wavelet neural networks models. Modeling Earth Systems and Environment, v. 8 (4), 5547-5563. https://doi.org/10.1007/s40808-022-01403-9.

Zanial, W.N.C.W., Malek, M.B.A., Reba, M.N.M., Zaini, N., Ahmed, A.N., Sherif, M.; Elshafie, A., 2023. River flow prediction based on improved machine learning method: Cuckoo Search-Artificial Neural Network. Applied Water Science, v. 13 (1), 28. https://doi.org/10.1007/s13201-022-01830-0.

Zhao, X.; Wang, H.; Bai, M.; Xu, Y.; Dong, S.; Rao, H.; Ming, W., 2024. A comprehensive review of methods for hydrological forecasting based on deep learning. Water, v. 16 (10), 1407. https://doi.org/10.3390/w16101407.