

Article

Improving Streamflow Forecasting with Multisource Data and ANNs: A Case Study in the Miranda River Basin, Brazil

Christian Pascal Silva Bouix ^{1,*} , Vinícius Villa e Vila ¹ , Marcos Roberto Benso ¹ , Sergio Nascimento Duarte ¹ , Carlos Roberto Padovani ², Roseli Aparecida Francelin Romero ³  and Patricia Angélica Alves Marques ¹ 

¹ Biosystems Engineering Department, “Luiz de Queiroz” College of Agriculture, University of São Paulo, Padua Dias Avenue 11, Piracicaba 13418-900, SP, Brazil; viniciusvilla@usp.br (V.V.e.V.); marcosbenso@usp.br (M.R.B.); snduarte@usp.br (S.N.D.); paamarques@usp.br (P.A.A.M.)

² Embrapa Pantanal, 21 de Setembro, 1880, Corumbá 79320-900, MS, Brazil; carlos.padovani@embrapa.br

³ Institute of Mathematical and Computer Sciences, University of São Paulo, Trabalhador Saocarlene Avenue 400, São Carlos 13566-590, SP, Brazil; rafrance@icmc.usp.br

* Correspondence: christianbouix@usp.br

Abstract

The escalating frequency of extreme hydrological events under environmental uncertainty poses a severe socio-economic threat to floodplains such as the Brazilian Pantanal, the world’s largest tropical wetland. Mitigating dynamic flooding and drying cycles is highly challenging due to a critical scarcity of in situ monitoring, leaving flood risks poorly understood. To address these data gaps, this study presents an advanced deep learning forecasting framework that integrates multisource environmental data, fusing satellite-derived precipitation (CHIRPS) and global land data assimilation evapotranspiration (GLDAS) data with historical river gauge telemetry. Multi-layered neural network architectures were optimized and combined with progressive moving average filters (10– and 15–day windows) to capture the complex hydrometeorological patterns of the data-scarce Miranda River Watershed. The optimal deep learning configuration, utilizing a robust two-hidden-layer topology (15 and 60 neurons), consistently outperformed standard baselines. Although purely exogenous data blocks successfully minimized satellite noise and captured seasonal trends ($NSE \geq 0.92$), structural underestimation of peak flows was observed. When incorporating the previous day’s streamflow (lag $t-1$) as a physical anchor, this limitation was noticeably alleviated, increasing both the Nash–Sutcliffe Efficiency (NSE) and Coefficient of Determination (R^2) values above 0.99. While this performance surge is driven by the strong temporal persistence inherent to the autoregressive lag, it introduces an operational trade-off by restricting the forecast to a reactive 24 h window. In this regard, an evaluation of the operational forecast horizons revealed that the exogenous deep learning blocks maximize warning lead times, providing a vital tool for proactive civil defense and disaster risk reduction. Ultimately, this multisource framework establishes a methodological foundation for automated decision support systems, providing the high-accuracy streamflow forecasting capability required to support future flood mitigation frameworks.

Keywords: AI-driven forecasting; artificial intelligence; hydrological modeling; machine learning



Academic Editor: Rashid Farooq

Received: 23 June 2026

Revised: 24 July 2026

Accepted: 28 July 2026

Published: 2 August 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

The frequency and intensity of extreme weather events, such as prolonged droughts and heavy rainfall driven by global warming, have significantly increased across the Brazil-

ian territory [1,2]. These environmental instabilities pose a major challenge to operational hydrology, as the resulting non-linearity of hydro-climatological processes limits the effectiveness of traditional physical models. Although these models are widely applicable, they often present disadvantages due to the high number of variables involved and the extensive sequence of calibration steps required [3,4]. Furthermore, many of these models depend on specific Geographic Information System (GIS) software, limiting their operational independence. This complexity is further exacerbated by the high costs and low density of gauge stations in several regions of Brazil—a systemic failure that is particularly critical in complex landscapes such as the Brazilian Pantanal, the world’s largest tropical wetland. In the Pantanal, the scarcity of ground-based data hinders the ability to predict events that impact local ecosystems and economies [5,6].

To overcome these limitations, computational modeling and programming have become integral in the field of hydrology, supporting studies on regionalization [7], data gap filling [8], and streamflow forecasting [9–11]. Among various artificial intelligence (AI) methodologies, data-driven models [12], particularly Artificial Neural Networks (ANNs), have gained significant prominence. ANNs are information processing systems that simulate the human brain’s ability to learn and recognize patterns through trial and error [13]. Utilizing layered structures where algorithms automatically adjust and learn from experience, these systems can identify complex non-linear relationships within provided datasets, making them a promising tool for achieving high accuracy in hydrological predictions [14].

The extensive adoption of ANNs in hydrological forecasting stems from their continuous architectural evolution, beginning with the foundational concepts of artificial neurons proposed by McCulloch and Pitts [15] and the introduction of the Perceptron—a pattern recognition model—by Rosenblatt [16]. As a prominent class of ANNs, the Multi-layer Perceptron (MLP) expands upon these principles through interconnected layers where feedforward propagation maps inputs to outputs and backpropagation algorithms systematically adjust weights and biases to minimize prediction errors. This process is particularly efficient when equipped with mathematically efficient activation functions such as the Rectified Linear Unit (ReLU), which minimizes computational costs while retaining a high capacity to solve continuous non-linear regression problems [17]. Due to this structural adaptability, data-driven AI models have become widespread globally, proving highly effective across diverse applications such as sediment modeling, water quality monitoring, evapotranspiration estimation, and flood forecasting [18,19]. In recent years, the application of data-driven models in operational hydrology has shifted from basic structural concepts to sophisticated architectural frameworks that are capable of handling high-dimensional environmental datasets. MLP networks, in particular, have been widely validated as robust universal approximators for complex hydrological modeling due to their mathematical efficiency in mapping non-linear signatures without requiring physically explicit watershed parameters [19,20].

In this context, the Miranda River sub-basin—a tributary of the Upper Paraguay River Basin (UPRB)—exemplifies the key challenge addressed in this study. The UPRB encompasses the Brazilian Pantanal, and its hydrological health depends directly on the streamflow generated in plateau sub-basins like the Miranda. However, this region lacks an adequate density of rainfall and river stations [5,21,22], making it difficult to apply conventional models to monitor the water volumes that sustain the Pantanal wetlands. Consequently, satellite-derived information has emerged not just as an alternative, but as a necessary tool for climatological monitoring. These spaceborne platforms can overcome geographic and economic barriers by providing continuous spatial and temporal data streams over complex terrains where physical infrastructure is unfeasible, thereby capturing localized hydrometeorological dynamics across vast and poorly accessible regions [23–25].

By incorporating variables such as precipitation and evapotranspiration into these ANN models, it is possible to bypass the limitations of traditional modeling and map complex patterns, even in data-scarce environments such as ungauged basins.

Thus, the integration of multisource remote sensing products has emerged as a standard practice for hydro-climatological monitoring. Satellite-derived data, such as the Climate Hazards Center Infrared Precipitation with Stations (CHIRPS) dataset, have been extensively validated across the Brazilian territory and shown to accurately capture spatial rainfall variability even in complex transition zones [26]. Concurrently, the Global Land Data Assimilation System (GLDAS) provides consistent, continuous evapotranspiration streams [27], establishing a reliable input configuration for machine learning frameworks operating in ungauged or data-scarce basins.

As hydrological systems function as natural low-pass filters whose intrinsic memory, governed by basin storage, stabilizes streamflow into seasonal and multi-year trends, modeling based exclusively on raw, unfiltered data generates a scale inconsistency that limits the accuracy of system predictions [28,29]. Consequently, the direct use of raw satellite data in ANN models can introduce significant high-frequency noise, potentially compromising their forecasting accuracy. Therefore, ensuring a robust relationship between the model's inputs and outputs requires both careful selection of input variables and rigorous data quality assessment [30]. To mitigate these limitations, exploring data pre-processing techniques, such as moving averages and temporal scale adjustments, is essential to smooth these signals and better capture the basin's underlying hydrological memory [31–33]. Mathematically, a moving average functions as a smoothing filter that replaces each individual data point with the average value over a temporal window of size k , where k represents the window memory size or the smoothing temporal scale [34]. This filtering approach effectively attenuates day-to-day fluctuations and periodic anomalies, isolating the underlying long-term signal and providing a cleaner dataset for enhanced network generalization.

Thus, based on the premise that hydrological systems function as natural low-pass filters and considering the capability of ANNs to approximate complex non-linear relationships in data-scarce environments, this study hypothesizes that the integration of ANNs with temporal smoothing techniques and satellite-derived hydrometeorological data can effectively mitigate high-frequency remote sensing noise, capture the upstream catchment's hydrological memory, and provide accurate forecasts of the streamflow entering the Pantanal biome.

Therefore, the objective of this study is to model and evaluate different ANN architectures for continuous streamflow forecasting in the Miranda River Watershed (MRW). The research specifically investigates the impact of utilizing raw daily data alongside 10-day and 15-day moving average windows on model performance, aiming to identify the most robust configuration to support water resource management in this strategic tributary. Additionally, the effect of incorporating time-lagged variables with a one-day-delayed streamflow ($t - 1$) as model inputs is evaluated, aiming to capture the immediate hydrological response and the temporal persistence of the basin's flow right before entering the Pantanal.

2. Materials and Methods

For the execution of this work, a watershed within the plateau surrounding the Pantanal biome was selected, as watersheds in this region act as entry points for the Pantanal biome plain and enable quantification of the precipitation and flow entering it. According to Tucci and Clarke [35], the water volume of a watershed primarily depends on vertical processes such as precipitation and evapotranspiration, which ultimately influence the

basin's streamflow. Guided by this hydrological principle, data collection and processing were conducted for flow, precipitation, and real evapotranspiration, all on a daily scale, serving as input for the tested neural networks.

Subsequently, ANN architectures were constructed and trained. To ensure the credibility and reproducibility of the predictive models, a systematic optimization approach was employed. A manual grid search was conducted to evaluate six distinct network topologies, varying the structural hyperparameters. A summary of the hyperparameter grid and optimization methods is presented in Table 1.

Network performance was evaluated using the following metrics: Mean Absolute Error (MAE), Root Mean Square Error (RMSE), Nash–Sutcliffe coefficient (NSE), and coefficient of determination (R^2). Figure 1 presents the methodological framework developed in this study, illustrating the step-by-step workflow from satellite data acquisition to the final streamflow forecasting blocks.

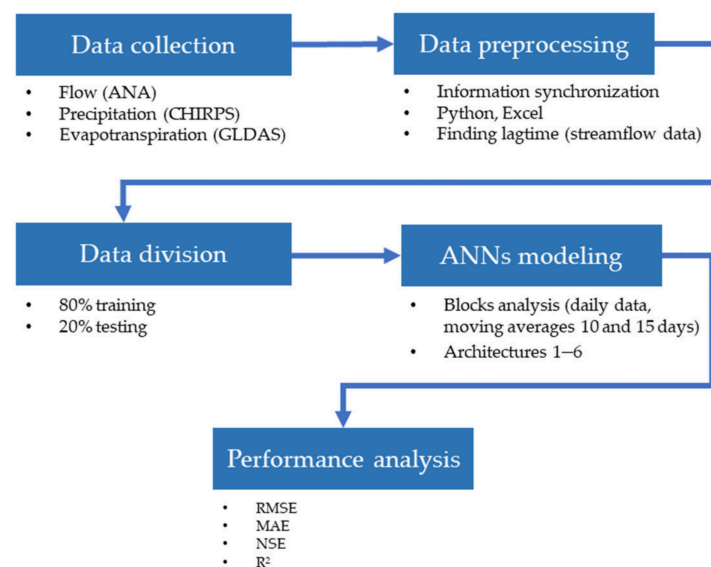


Figure 1. Workflow from satellite data acquisition to the final streamflow forecasting.

Table 1. Hyperparameter grid and optimization settings for the evaluated ANNs.

Parameter	Configuration/Value	Search Range
Optimization method	Manual grid search (empirical topology evaluation)	Fixed
Cross-validation	Time-series split (5 folds)	3 to 6 folds
Optimizer	Adam	Fixed
Learning rate	0.001	0.0001; 0.001; 0.01; 0.1
Adam β_1 (first moment decay)	0.9	Fixed
Adam β_2 (second moment decay)	0.999	Fixed
Adam ϵ	1×10^{-7}	Fixed
Loss function	Mean Squared Error (MSE)	Fixed
Activation function (hidden layers)	Rectified Linear Unit (ReLU)	ReLU; Sigmoid; Tanh (Hyperbolic tangent)
Activation function (output layer)	Linear	Fixed
Early stopping (patience)	25 Epochs (restoring best weights based on validation loss)	Fixed
Batch size	32	16; 32; 64
Variable hyperparameters	Hidden layers (1 to 3); neurons (10 to 80 per layer)	Fixed hidden layers (1 to 3); 2 to 150 neurons

Fixed denotes hyperparameters that were kept constant throughout the optimization process.

2.1. Characterization of the Study Area

The study area of this research is the MRW, located in the state of Mato Grosso do Sul, Brazil (Figure 2). The total area covered by the MRW is approximately 42,023 km², with the Miranda and Aquidauana rivers being the main rivers in the sub-basins. A portion of the MRW is situated within the Pantanal biome, and the Miranda and Aquidauana sub-basins serve as entry points for flows into the Pantanal. This study specifically considers only the Miranda River Basin (MRB).

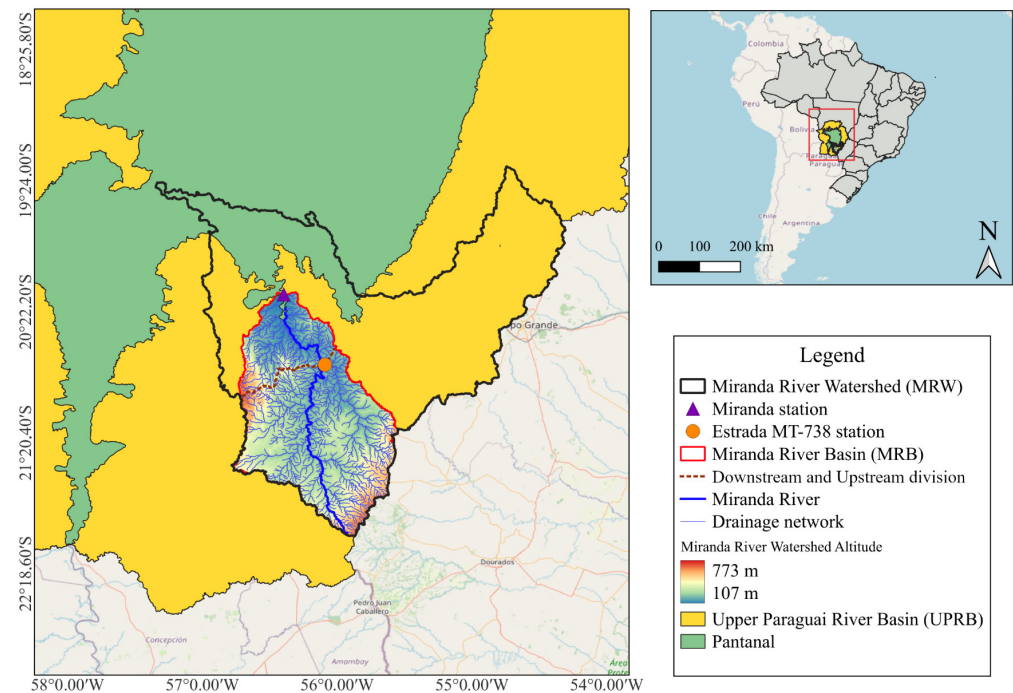


Figure 2. Location of the Miranda River Watershed and delineated sub-basins.

The MRB was divided into two portions, delineated using the GIS software QGIS (v. 3.28.0). The river gauge station in the city of Miranda (code 66910000) is considered the outlet point of the total portion of the sub-basin, encompassing its entire area up to the end of the Miranda River's extension. The upstream portion of the sub-basin (delimited by the brown dashed line in Figure 2) has its outlet at the Estrada MT–738 River gauge station (code 66900000) in the city of Bonito. This area extends from the middle section of the sub-basin to the end of the Miranda River.

The division of the sub-basin into two portions was performed to establish a cause-and-effect relationship between the upstream and downstream river gauge stations. The flows measured at the upstream portion's station serve as an indicator of the flows at the downstream station of the basin. In this way, during the ANN training phase, it is possible for the networks to learn the patterns in the data and establish the cause-and-effect relationship between the input and output variables. The characterization of the total and upstream portions of the sub-basin is detailed in Table 2.

The precipitation and flow information for these stations is available from the Hidroweb Portal of the National Water Resources Information System (SNIRH) of the National Water Agency (ANA), accessible at the following link: <https://www.snirh.gov.br/hidroweb/serieshistoricas> accessed on 15 July 2026. The World Meteorological Organization (WMO) [36] recommends certain minimum coverage densities for rain gauges for hydrological simulation in specific physiographic regions—for example, (a) small islands: 25 km² per station; (b) mountainous regions: 250 km² per station; (c) interior plains and mountainous/undulating areas: 575 km² per station; and (d) coastal areas: 900 km² per

station. These values were also used in the study of Cristaldo et al. [5], who found a low station density for the Aquidauana River basin, another plateau basin adjacent to the Pantanal. Considering the Miranda River sub-basin's total area of 15,190.99 km², the rain gauge density in this region falls below the recommended WMO values. Therefore, aiming for comprehensive coverage of precipitation within the study area of the sub-basin, data extracted from satellite imagery were used.

Table 2. Characterization of the portions of the Miranda River sub-basin.

Basin Portion	Area (km ²)	Perimeter (km)	Average Altitude (m)	Minimum Altitude (m)	Maximum Altitude (m)
Miranda RB (total portion)	15,190.99	686.70	307.78	107.00	773.00
Miranda RB (downstream portion)	11,552.14	589.92	330.50	135.00	754.00

2.2. Streamflow Rate Data

The streamflow rate information for the fluvial stations in the basins was collected from the historical series available from the Hidroweb Portal of SNIRH/ANA, accessible at the web address: <https://www.snirh.gov.br/hidroweb/serieshistoricas>, accessed on 15 July 2026. Flow rate data for the fluvial stations Estrada MT–738 (code 66900000) and Miranda (code 66910000) were extracted from 1 January 2000 to 31 December 2020, in order to obtain up- and downstream flows in the sub-basin portions. These fluviometric stations are located approximately 81 km apart and are managed by ANA (Table 3). The stations are concentrated only in the cities of Miranda/MS and Bonito/MS, and do not cover the entire basin area [5,21].

Table 3. Characteristics of fluviometric stations managed by ANA in the municipalities of Miranda and Bonito.

Station Name	Code	City	Coordinates	Drainage Area (km ²)
Miranda	66910000	Miranda	20°24' S, 56°39' W	15,000.00
Estrada MT–738	66900000	Bonito	20°76' S, 56°09' W	11,600.00

2.3. Precipitation Data

Daily precipitation data were extracted from satellite images of the CHIRPS v. 2.0 Final collection, provided by the Climate Hazards Center, US, Santa Barbara, through the Google Earth Engine platform. The collection has a spatial resolution of 0.05° and is available for use at the web address: https://developers.google.com/earth-engine/datasets/catalog/UCSB-CHG_CHIRPS_DAILY accessed on 15 July 2026. Precipitation data were extracted from shapefile files outlining the MRB for both its total and upstream portions, resulting in cumulative precipitation for each section of the sub-basin. The temporal interval was the same as that for the flow rate data (Table 4).

Table 4. Summary of the daily variables used in the study.

Variable	Source	Resolution	Period	Unit
Precipitation	CHIRPS collection v. 2.0	0.05°	1 January 2000–31 December 2020	mm d ^{−1}
Evapotranspiration	GLDAS collection v. 2.0	0.25°	1 January 2000–31 December 2020	mm d ^{−1}

Table 4. *Cont.*

Variable	Source	Resolution	Period	Unit
Streamflow	Hidroweb (SNIRH)	na *	1 January 2000–31 December 2020	m ³ s ^{−1}

* Not applicable.

2.4. Evapotranspiration Data

The daily average evapotranspiration was extracted and calculated from satellite images of the NASA Global Land Data Assimilation System (GLDAS) v. 2.0 collection, with a spatial resolution of 0.25°, using the methodology outlined in Section 2.3 for rainfall data (Table 4). The collection is available at the following address: <https://ldas.gsfc.nasa.gov/gldas>, accessed on 15 July 2026. The data were retrieved using Google Earth Engine, similar to the precipitation data described in Section 2.3.

2.5. Pre-Analysis of the Collected Data

Initially, an exploratory data analysis was conducted to assess the intrinsic characteristics of the dataset. A correlation analysis was performed using Pearson's linear coefficient (r). The choice of Pearson's coefficient was intentional; first, it serves to demonstrate the structural lack of pure linearity in the system's hydrological routing and rainfall–runoff processes, providing empirical justification for the adoption of advanced non-linear algorithms such as ANNs. To ensure consistency with the hydro-climatological monitoring standards established for the Pantanal basin, the temporal alignment and synchronization of the hydrological time-series followed the regional protocol applied by Santos Antoniassi [37] for Embrapa Pantanal.

This methodology relies on the Pearson correlation coefficient as the baseline metric to evaluate the linear association between the downstream and upstream station data, which is subsequently validated through calculation of the maximum R^2 value between the time-series. Adhering to this institutional standard is critical, as it ensures that our preprocessing pipeline remains fully consistent and directly comparable with official hydrological assessments and historical databases maintained by Embrapa Pantanal for the Miranda River Basin, establishing a standardized foundation prior to the application of the non-linear machine learning architectures.

Subsequently, each variable was visualized using line graphs and box plots to comprehensively analyze their temporal dynamics and statistical distributions. Line graphs were employed to map the continuous chronological evolution of the hydro-climatological series, allowing for clear identification of seasonal trends, cyclic patterns, and the basin's long-term baseline behavior. Concurrently, box plots were utilized to summarize the data's statistical distribution. This graphical tool is essential for visualizing the dispersion (interquartile ranges) and central tendencies (medians) of data and, most importantly, for the detection of outliers [38]. In the context of this study, identifying outliers is crucial, as they represent extreme atypical events such as severe flood peaks or prolonged droughts that govern the non-linear hydrological responses of the watershed.

2.6. Data Preprocessing

For analysis of the collected daily data, a check for missing data in the satellite product readings or in the ANA database for the period of interest was first performed. The data collected was visualized in Excel® spreadsheet format for all variables. Data from the Hidroweb fluvial stations (SNIRH/ANA) showed gaps in flow readings for the two studied fluvial stations, while no periods of missing data were identified in the satellite products. Bertoni and Tucci [39] have analyzed linear regression and regional weighting

methodologies for filling flow gaps, and concluded that these methodologies are not suitable for this purpose at daily and hourly scales. Therefore, for more comprehensive analysis of variables and testing of the neural network architectures, it was decided to exclude daily rows for all variables instead of filling data gaps, aiming for synchronization of all information. Once the dataset was cleaned and synchronized, the data were submitted to a systematic preprocessing pipeline to handle the inherent non-linearities and temporal dependencies of the hydrological variables. To address input scale discrepancies, all features were normalized to the [0, 1] range using the Min–Max scaling technique.

All data preprocessing was conducted in Python v. 3.9, through open-source software, using the Jupyter Notebook platform v. 7.4.2. To definitively prevent data leakage and preserve the temporal integrity of the hydrological series, the processed data was chronologically split into two main subsets: 80% for training and 20% for final out-of-sample testing. While Hanoon et al. [40] utilized a distinct configuration of 50% for training, 25% for validation, and 25% for testing, the proportional allocation in this study follows a similar methodological principle of rigorously reserving a substantial, strictly isolated subset for final model evaluation. To dynamically optimize the models without exposing them to future data, the 80% training subset was used for 5-fold time-series cross-validation. This technique creates internal training and validation folds that strictly respect the chronological sequence of the observations, allowing the Early Stopping mechanism to monitor the learning process while keeping the final 20% of the dataset completely unseen for predictive generalization assessment.

The raw dataset comprised 7671 daily observations. Following the synchronization and filtering process, 7654 records were retained, corresponding to 99.78% of the total raw data, while only 0.22% were excluded due to missing values. The dataset was chronologically partitioned as follows for model training and evaluation: the training subset spanned from 8 January 2000 to 26 October 2016, accounting for 80% of the observations, while the validation and testing subset covered the period from 27 October 2016 to 31 December 2020, representing the remaining 20%. This temporal split was strictly maintained to ensure that the models were evaluated against unseen future data, enabling rigorous assessment of their predictive and generalization capabilities.

2.7. Artificial Neural Networks Construction

As parallel distributed information processing systems [41], ANNs (particularly MLPs) are structured in interconnected layers. In the MLP architecture, information from the input layer is propagated to the subsequent hidden layers. Each neuron j in a hidden layer receives the input signals x_i , which are multiplied by their respective connection weights w_{ij} . The neuron then aggregates these weighted inputs to produce a net input signal. This net sum is passed through an activation function f , which scales the input to determine the neuron's activation level and compute its final output y_j . This process is mathematically described by Equation (1) [17],

$$y_i = f\left(\sum_{i=1}^n x_i w_{ij}\right) \quad (1)$$

where y_j is the computed output of the hidden neuron j , f represents the selected activation function, x_i denotes the input variables from node i , w_{ij} indicates the weight associated with the connection between input node i and hidden node j , and n is the total number of input nodes.

The activation function determines the numerical range of the output and the neuron's activation level. This study utilizes the ReLU function, which is widely adopted in ANN architectures due to its low computational cost compared to other functions. The ReLU

function returns zero for negative values and the value itself for non-negative values, as schematically defined in Equation (2) [17]:

$$f(a) = \begin{cases} 0, & \text{for } a < 0 \\ a, & \text{for } a \geq 0 \end{cases} \quad (2)$$

Once the signal reaches the output layer, the network computes the error between the model's prediction and the observed data. To minimize this error, MLPs employ the backpropagation algorithm, which iteratively updates the initial weights and biases. According to Ali and Shahbaz [17], these parameter adjustments are governed by Equations (3) and (4):

$$E_k = (d_k - y_k) f' \left(\sum_{j=1}^m y_j w_{jk} \right) \quad (3)$$

$$\Delta W_{jk} = \alpha E_k y_k \quad (4)$$

where d_k is the observed output value, y_k is the predicted output value obtained with the model, f' is the derivative of the activation function, W_{jk} is the weight between the hidden and output layer neurons, and α is the learning rate of the model.

The number of neurons in the input and output layers of the network depends on the desired model, according to the studied problem. The neurons in the hidden layer are parameters that must be determined to optimize the performance of the proposed model. MLP-ANNs were used in the study, and different architectures were tested until the optimal coefficients were found.

To achieve this, the input and output layer data were defined, following which the number of intermediate layers and their neurons were optimized through trial and error. The output data was the same for all tested architectures. The k-fold time-series cross-validation technique was used, following the methodology of Jung et al. [42] and Rohani et al. [43], with a fixed 5-fold data split for all models tested to increase the training speed.

In the training process of the neural networks, the Adam (Adaptive Moment Estimation) optimizer was used. The goal of this algorithm is to obtain the optimal parameters for the neural network (in this case, the MLP) [44]. Synaptic weights were randomly initialized to prevent bias and ensure an objective learning process during the model training stage. The overall architecture of the tested neural networks can be seen in Figure 3.

2.8. Performance Analysis of the Models

To evaluate the predictive performance of the six proposed ANN-MLP architectures, the statistical metrics MAE, RMSE, R^2 , and NSE were utilized. The mathematical relationships for calculating these evaluation metrics are expressed in Equations (5)–(8), as established in the literature by Apaydin et al. [45], Oliveira et al. [46], Zakaria et al. [47], and Ali and Shahbaz [17].

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (X_i - Y_i)^2} \quad (5)$$

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |Y_i - X_i| \quad (6)$$

$$R^2 = \left[\frac{\sum_{i=1}^N (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^N (X_i - \bar{X})^2} \sqrt{\sum_{i=1}^N (Y_i - \bar{Y})^2}} \right]^2 \quad (7)$$

$$\text{NSE} = 1 - \frac{\sum_{i=1}^N (X_i - Y_i)^2}{\sum_{i=1}^N (X_i - \bar{X})^2} \quad (8)$$

where X_i is the observed variable (with its mean denoted as $\bar{X}$), Y_i is the predicted variable (with its mean denoted as $\bar{Y}$), and N is the number of instances.

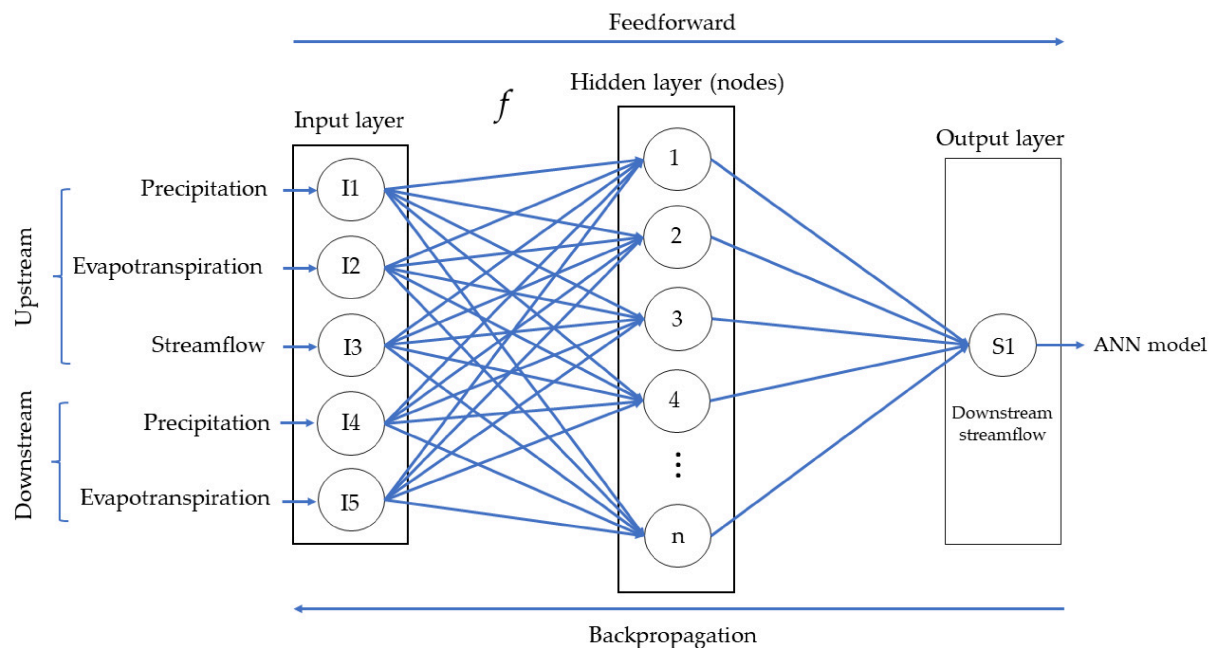


Figure 3. Overall architecture of the tested Artificial Neural Networks.

3. Results and Discussion

3.1. Exploratory Data Analysis

The generated correlation matrix (Figure 4) illustrates the degree of linear association among precipitation and evapotranspiration variables for both the upstream and downstream portions of the sub-basin, alongside the streamflow records from the Estrada MT–738 (code 66900000) and Miranda (code 66910000) stations. Upstream and downstream evapotranspiration data exhibited a perfect correlation ($r = 1.0$), while precipitation data demonstrated a near-perfect correlation ($r = 0.99$).

This high degree of linear association is physically consistent with the regional climatology. At the sub-basin scale, atmospheric drivers such as evapotranspiration exhibit high spatial homogeneity [48]. Similarly, the near-perfect correlation in precipitation indicates strong temporal coherence across the 81 km distance separating the stations. This result confirms that the variability in rainfall is spatially consistent throughout the study area, with both regions responding uniformly to large-scale meteorological events. These findings align with those of Oliveira-Júnior et al. [48], who evaluated CHIRPS estimates across the Brazilian Midwest and identified high correlation coefficients as a robust indicator of the product's capability to capture regional precipitation dynamics. The authors emphasized that such high temporal coherence highlights the reliability of the CHIRPS dataset in capturing large-scale rainfall patterns, even if the absolute volumetric agreement may vary slightly due to the inherent effects of interpolation.

Conversely, the streamflow records yielded an instantaneous correlation of $r = 0.60$, corresponding to a relatively low coefficient of determination ($R^2 = 0.36$). This weak concurrent correlation indicates that same-day upstream data is insufficient for accurate downstream prediction when used alone, as it fails to account for the natural travel time of the flood wave along the river channel. Recent research on hydrological modeling frameworks has demonstrated that estimating the optimal flow travel time through cross-correlation analysis is essential to resolve such a temporal gap [49].

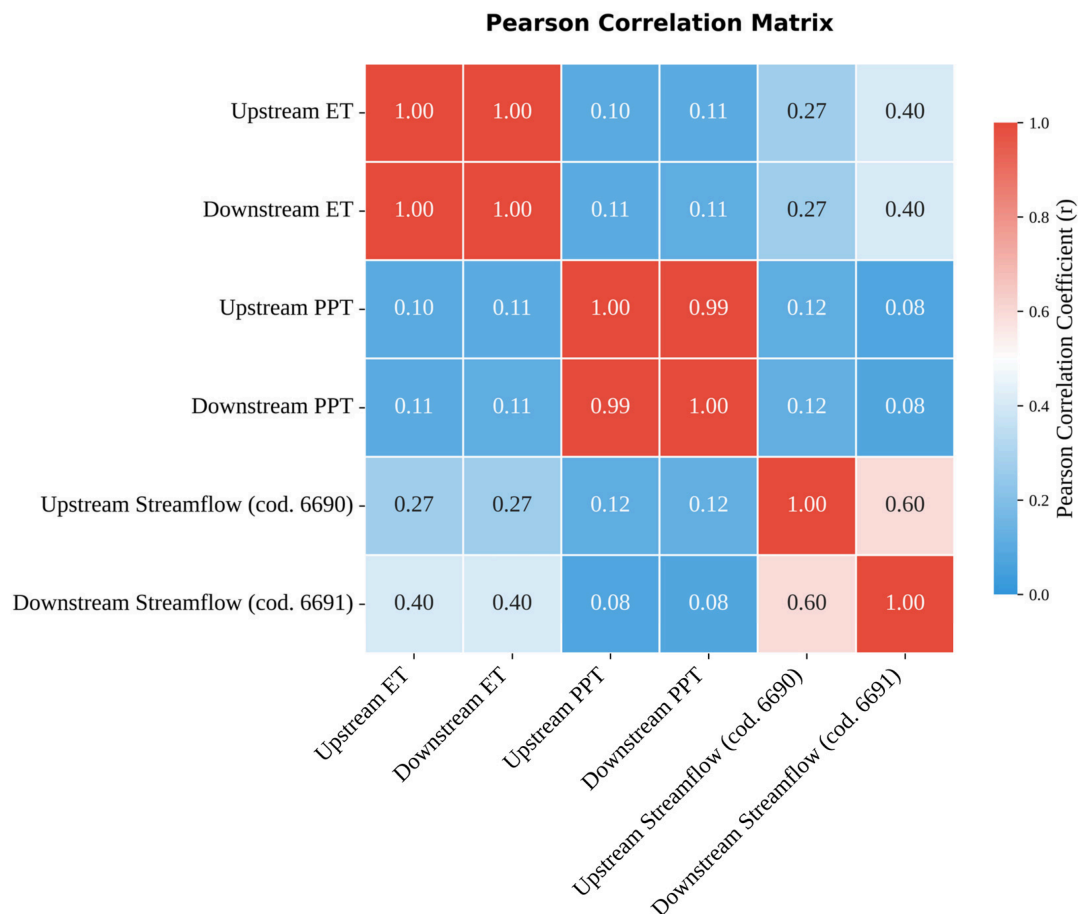


Figure 4. Correlation matrix for precipitation (PPT), evapotranspiration (ET), and flow (Q) data from the Estrada MT–738 (code 66900000) and Miranda (code 66910000) stations for the upstream and downstream portions of the Miranda River sub-basin.

Following the protocol established by Santos Antoniassi [37] for the Pantanal basin, we performed series synchronization to identify the maximum correlation between the datasets. This analysis revealed a 6-day lag, representing the travel time of the flood wave from the upstream to the downstream station (Figure 5). Incorporating this lag significantly improved the correlation, increasing the R^2 from 0.36 to 0.78. This substantial improvement underscores the existence of a robust hydrodynamic cause-and-effect relationship, establishing this 6-day lag as the baseline temporal window for the upstream streamflow inputs in the ANN architectures.

3.1.1. Precipitation Data Behavior

The daily precipitation time-series (Figure 6A) clearly illustrates the region's strong seasonal variability and the high degree of spatial meteorological coherence between the upstream and downstream portions of the Miranda River Basin. Visual analysis of the daily data revealed intense, localized rainfall events capable of reaching approximately 73 mm day^{-1} (notably during the 2003–2004 period), with extreme isolated spikes approaching 86 mm day^{-1} between 2013 and 2014. Despite these acute daily extremes, both monitoring points exhibited a synchronized macro-behavior, consistently transitioning into recurrent dry periods where daily precipitation drops to near zero, reflecting the basin's distinct cyclic nature.

Lagtime (Days)	<i>r</i>	<i>R</i> ²
0	0.5955	0.3546
1	0.6456	0.4167
2	0.6929	0.4802
3	0.7330	0.5373
4	0.7639	0.5836
5	0.7852	0.6165
6	0.7962	0.6340
7	0.7948	0.6317
8	0.7799	0.6082
9	0.7528	0.5667
10	0.7170	0.5141

r = Pearson correlation, *R*² = Coefficient of determination. Bold values show the maximum correlation.

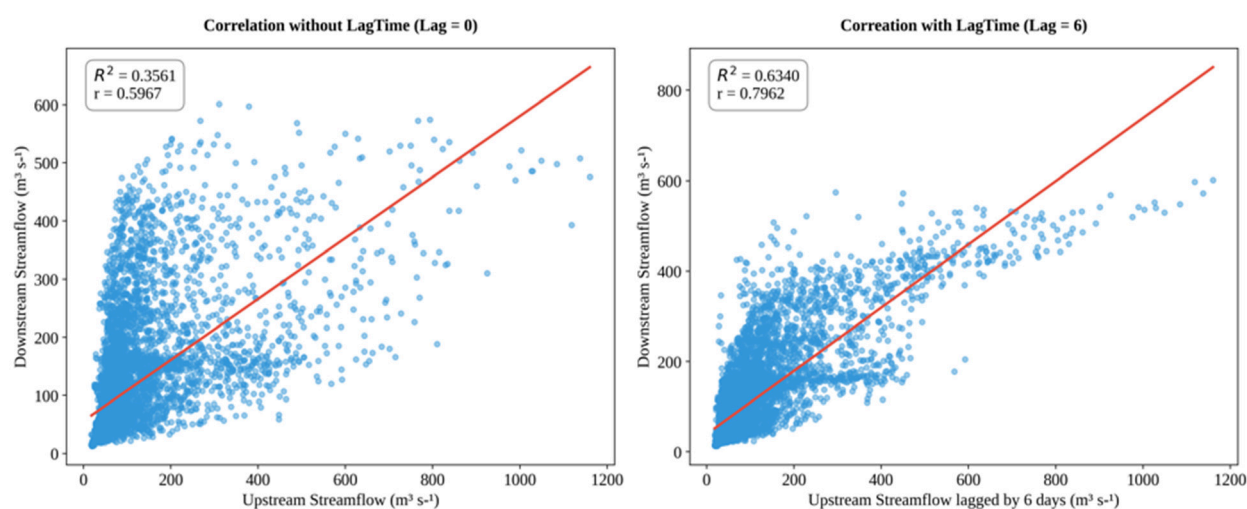


Figure 5. Calculated correlations between upstream and downstream streamflow stations for lagtime identification.

The monthly mean precipitation distribution (Figure 6B) further demonstrates this pronounced seasonality. A well-defined dry season occurs from April to September, characterized by a drastic reduction in rainfall volumes (generally falling below 5 mm day^{−1}). During this period, the winter months of June, July, and August represent critical meteorological valleys. Within these most severe dry spells, the absolute minimum of the monthly means dropped to negligible rates of 0.22 mm day^{−1} upstream and 0.19 mm day^{−1} downstream. This strict and prolonged precipitation deficit strongly correlates with the baseflow recessions observed in the streamflow data, confirming the basin's precise response to regional winter drought conditions.

Conversely, the wet season (spanning from October to March) concentrates the vast majority of the basin's annual rainfall. The overall median of the monthly means across the entire historical series stabilizes at 3.83 mm day^{−1} for both the upstream and downstream portions, reinforcing the homogeneous macro-distribution of rainfall across the sub-basin throughout the hydrological year. During the wet season, the monthly averages stay consistently above this baseline, with upper quartile values frequently approaching 10 mm day^{−1}. Furthermore, the box plots reveal a high frequency of upper-bound outliers during this wet period, particularly in February, March, and October. These atypical outliers represent exceptionally rainy months where the average daily precipitation deviated significantly from the standard interquartile range, reaching absolute maximum monthly

means of $10.66 \text{ mm day}^{-1}$ upstream in December and $10.28 \text{ mm day}^{-1}$ downstream in January. These sustained, high-volume atmospheric inputs act as the primary meteorological drivers of the massive volumetric routing and flood events identified in the basin's hydrographs.

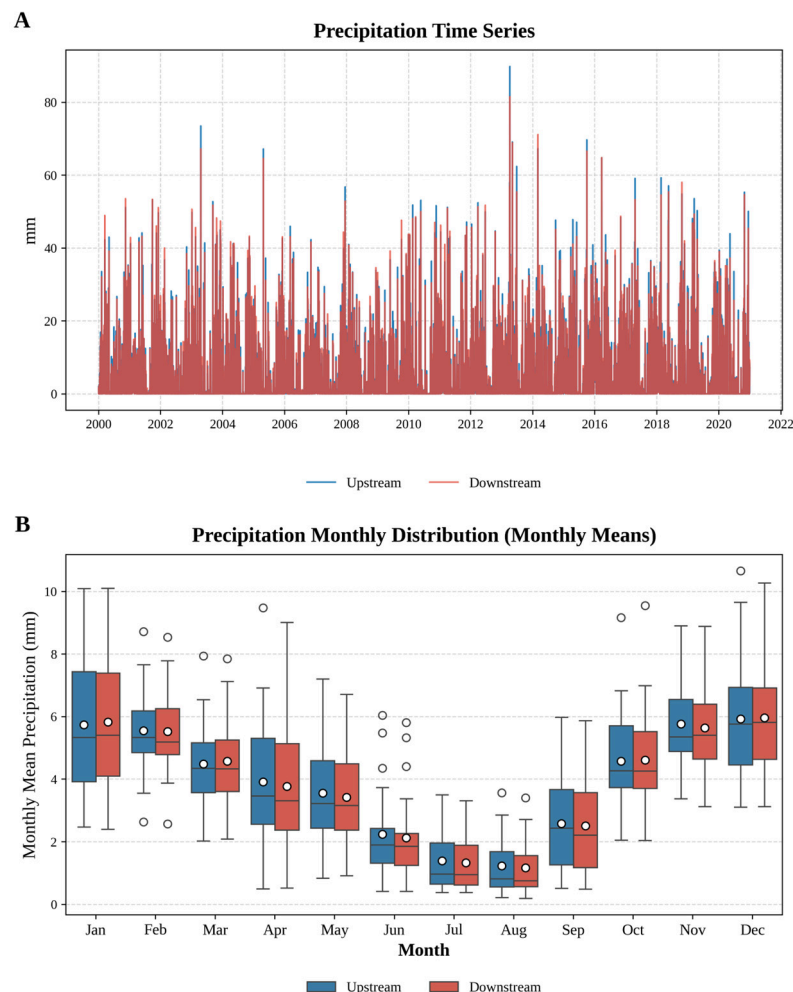


Figure 6. (A) Daily precipitation from 2000 to 2020 for the upstream and downstream portions of the Miranda River sub-basin. (B) Box plot of the daily average precipitation values for the months of January to December from 2000 to 2020 for the upstream and downstream portions of the Miranda River sub-basin.

3.1.2. Evapotranspiration Data Behavior

Figure 7 presents a comprehensive overview of the evapotranspiration (ET) dynamics over the study period. The daily evapotranspiration time-series (Figure 7A) reveals well-defined, highly synchronized interannual cycles where daily ET rates frequently peak near 7.0 mm day^{-1} during the wet summer months, contrasting with sharp declines to values near 1.0 mm day^{-1} during the dry winter season. This strong temporal synchronization and the nearly perfect graphical overlap between the upstream and downstream time-series visually corroborate the perfect positive spatial Pearson correlation coefficient ($r = 1.00$) calculated for the dataset (Figure 4). This identical behavior confirms a homogeneous atmospheric demand and vegetative response across the sub-basin segments throughout the analyzed decades.

The monthly mean distribution of ET (Figure 7B) further reveals this seasonal behavior, operating as a direct response to solar radiation and temperature patterns. The highest monthly mean evapotranspiration rates are concentrated during the wet season

(i.e., from October to March), with values stabilizing between approximately 4.0 and 5.0 mm day⁻¹ and reaching absolute monthly mean maxima of 5.50 mm day⁻¹ upstream and 5.37 mm day⁻¹ downstream.

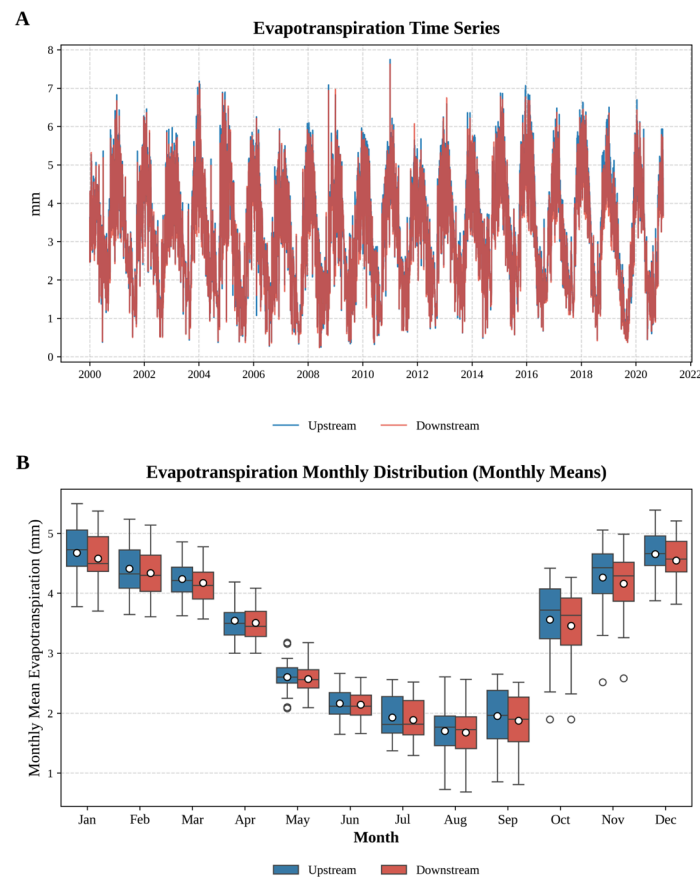


Figure 7. (A) Daily evapotranspiration from 2000 to 2020 for the upstream and downstream portions of the Miranda River sub-basin. (B) Box plot of the daily average real evapotranspiration values for the months of January to December from 2000 to 2020 for the upstream and downstream portions of the Miranda River sub-basin.

Conversely, the onset of the dry season triggers a progressive reduction in ET rates due to declining winter temperatures and severe soil moisture limitations. The lowest average rates are systematically recorded between June and September, with monthly averages falling in the range of 1.8 to 2.0 mm day⁻¹. Within this recession period, the winter months of July and August represent the core seasonal valleys of the ET cycle, where monthly means drop to absolute minima of 0.73 mm day⁻¹ upstream and 0.69 mm day⁻¹ downstream under extreme drought conditions. The compression of the interquartile ranges during these months highlights a highly stable energy-limited and moisture-constrained biophysical state across the entire MRB.

3.1.3. Streamflow Data Behavior

Regarding the streamflow time-series (Figure 8A), the visual behavior between the upstream and downstream hydrographs reflects the unadjusted data prior to the synchronization methodology discussed in Section 3.1. Visually, the upstream station (Estrada MT-738, code 66900000) presents much more pronounced and sharper daily peak flows, generating intense, short-lived spikes that reach an absolute maximum of 1160.18 m³ s⁻¹ in 2018. In contrast, the downstream station (Miranda, code 66910000) presents attenuated daily maxima reaching 601.35 m³ s⁻¹ (also in 2018). Furthermore, a trend of increasing extreme flows over the last six assessed years (2015 to 2020) is noticeable for both stations.

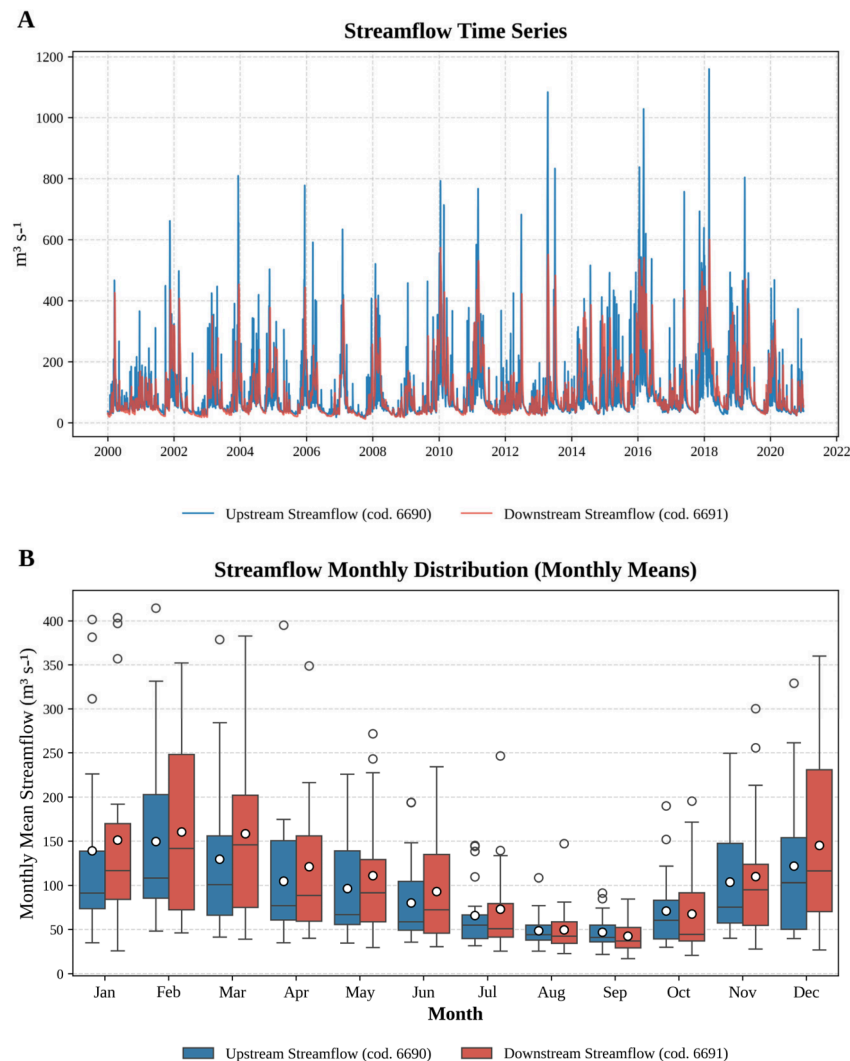


Figure 8. (A) Daily flow hydrograph from 2000 to 2020 for the fluvial stations Estrada MT–738 (code 66900000) and Miranda (code 66910000) in the Miranda River sub-basin. (B) Box plot of the daily average flow values for the months of January to December from 2000 to 2020 for the fluvial stations Estrada MT–738 (code 66900000) and Miranda (code 66910000) in the Miranda River sub-basin.

A distinct hydrological behavior can be observed when comparing these instantaneous peaks with the monthly averages (Figure 8B), which illustrate the basin's volumetric routing. Despite the higher daily peaks upstream, both the medians and the upper quartiles of the monthly mean streamflow are generally higher for the Miranda station. This dynamic indicates that, while the upstream section responds rapidly to localized precipitation—clearly reflected by the numerous extreme outliers above the whiskers—the downstream station captures the cumulative runoff of a much larger contributing drainage area. Consequently, the downstream portion maintains a more sustained baseflow and routes a larger overall volume of water, resulting in consistently wider interquartile ranges and higher monthly minima rather than transient extreme peaks. The absolute maximum monthly means remained closely matched, reaching $414.53 \text{ m}^3 \text{s}^{-1}$ upstream in February and $403.62 \text{ m}^3 \text{s}^{-1}$ downstream in January.

The data distribution further reveals pronounced seasonal cycles. A well-defined dry season occurs from April to September, with the winter months of June through August marking the critical hydrological valleys. During this recession period, observed monthly flows decrease significantly to medians near $50 \text{ m}^3 \text{s}^{-1}$, sustained primarily by groundwater contributions, and drop to absolute minimum monthly means of $22 \text{ m}^3 \text{s}^{-1}$ upstream and

$17.21 \text{ m}^3 \text{ s}^{-1}$ downstream under extreme drought conditions. Conversely, the wet season (October to March) drives significant volumetric expansion, with medians approaching or exceeding $100 \text{ m}^3 \text{ s}^{-1}$. This period is characterized by prominent upper-bound outliers, particularly in the upstream data, with exceptionally wet years pushing monthly averages above $350 \text{ m}^3 \text{ s}^{-1}$. This sustained downstream accumulation during the summer reinforces the spatial consistency of rainfall over the region and the lower basin's capacity to aggregate and progressively route massive seasonal runoff.

3.2. Experimental Setup and Input Configurations

To systematically evaluate the predictive performance of the Artificial Neural Networks (ANNs), the experimental setup was organized into six distinct operational input blocks (Table 5). In all scenarios, the output (target variable) of the models was the downstream streamflow at the current time ($Q_{dw}(t - 0)$). Under two predictive input configurations, three levels of temporal data conditioning were cross-examined: (i) raw daily data; (ii) 10-day moving average smoothing; and (iii) 15-day moving average smoothing.

Table 5. Summary of the experimental input blocks, data conditioning, feature configurations, and target output for ANN modeling.

Block	Data Conditioning (Smoothing)	Input Features Vector
1	None (Raw Daily Data)	Baseline Variables = ($Q_{up}(t - 0)$, $Q_{up}(t - 5)$, $Q_{up}(t - 6)$, $Q_{up}(t - 7)$, $P_{up}(t - 0)$, $ET_{up}(t - 0)$, $P_{dw}(t - 0)$, $P_{dw}(t - 1)$, $ET_{dw}(t - 0)$, $ET_{dw}(t - 1)$)
2	None (Raw Daily Data)	Baseline Variables + $Q_{dw}(t - 1)$
3	10-Day Moving Average	Baseline Variables = ($Q_{up}(t - 0)$, $Q_{up}(t - 5)$, $Q_{up}(t - 6)$, $Q_{up}(t - 7)$, $P_{up}(t - 0)$, $ET_{up}(t - 0)$, $P_{dw}(t - 0)$, $P_{dw}(t - 1)$, $ET_{dw}(t - 0)$, $ET_{dw}(t - 1)$)
4	10-Day Moving Average	Baseline Variables + $Q_{dw}(t - 1)$
5	15-Day Moving Average	Baseline Variables = ($Q_{up}(t - 0)$, $Q_{up}(t - 5)$, $Q_{up}(t - 6)$, $Q_{up}(t - 7)$, $P_{up}(t - 0)$, $ET_{up}(t - 0)$, $P_{dw}(t - 0)$, $P_{dw}(t - 1)$, $ET_{dw}(t - 0)$, $ET_{dw}(t - 1)$)
6	15-Day Moving Average	Baseline Variables + $Q_{dw}(t - 1)$

Q = streamflow; P = precipitation; ET = evapotranspiration. The subscripts 'up' and 'dw' denote upstream and downstream data, respectively. The terms in parentheses ($t - n$) indicate the temporal lag in days relative to the current forecast day ($t - 0$). "Baseline Variables" refers to the standard set of input features defined in the immediately preceding odd-numbered block (e.g., Block 1 for Block 2), to which the antecedent downstream streamflow $Q_{dw}(t - 1)$ is added as an autoregressive term.

The selection of 10- and 15-day moving average windows serves as an essential data pre-processing low-pass filter. Consequently, data smoothing techniques are strongly recommended to isolate deterministic physical routing signals before network training, preventing the algorithms from overfitting to local sensor oscillations and aleatoric uncertainties [26]. It is important to emphasize that these moving average windows were calculated exclusively using historical observations, ensuring that no future information was incorporated during the data conditioning process.

Regarding the spatial inputs, the temporal lag window for the upstream streamflow (Q_{up}) was established through statistical cross-correlation analysis, as described in Section 3.1. This procedure identified a peak correlation at a lag of 6 days ($t - 6$), reflecting the average travel time of the flood wave. To accommodate physical variations in flow celerity under different hydrological regimes, a 3-day window centered around this peak ($t - 5$, $t - 6$, and $t - 7$) was selected. This statistical approach is fundamentally aligned with the methodology proposed by Sudheer et al. [50], who demonstrated that the identification of appropriate input vectors should rely on quantitative cross-correlation properties to define significant antecedent inputs, rather than arbitrary trial-and-error procedures. It is

important to clarify the forecasting horizon of the proposed framework. In particular, the model operates on a hybrid temporal basis.

- (a) Hydrodynamic Forecasting (6-day horizon): By utilizing lagged upstream streamflow inputs (Q_{up} at $t - 5$, $t - 6$, $t - 7$), the model effectively acts as a forecasting engine, leveraging the physical travel time of the flood wave to project downstream responses before the peak arrives. This provides a clear predictive lead time of up to 6 days.
- (b) Diagnostic Nowcasting (0-day horizon): The inclusion of contemporaneous meteorological variables (precipitation and evapotranspiration) functions as a diagnostic component, refining the model's accuracy by incorporating real-time atmospheric boundary conditions.

Therefore, while the model is constrained by meteorological data availability (nowcasting) for its contemporary inputs, it intrinsically provides a predictive lead-time based on upstream hydrodynamic propagation. In addition to these lagged variables, the instantaneous upstream streamflow, $Q_{up}(t - 0)$, is explicitly included in the input vector. While the lagged window ($t - 5$ to $t - 7$) implicitly maps the historical routed volume, $Q_{up}(t - 0)$ provides the network with a real-time spatial boundary condition. This anchors the model to the instantaneous hydrological baseline state of the upper basin, establishing the current spatial gradient before applying the lagged flood wave. This inclusion mirrors the foundational framework, where a sequence of temporal states provides the necessary physical context for the network. Therefore, the input structure defined in this study stands in strict alignment with established physical and statistical benchmarks in data-driven streamflow modeling.

The baseline feature vector (applied to Blocks 1, 3, and 5) consists of 10 input variables: upstream streamflow (Q_{up}) at lags $t - 0$, $t - 5$, $t - 6$, and $t - 7$; upstream precipitation (P_{up}) and evapotranspiration (ET_{up}) at lag $t - 0$; and downstream precipitation (P_{dw}) and evapotranspiration (ET_{dw}) at lags $t - 0$ and $t - 1$. Alternatively, Blocks 2, 4, and 6 introduce an autoregressive component by incorporating the downstream streamflow from the previous day (Q_{dw} at lag $t - 1$) as an additional input, totaling 11 predictive variables. For each of these six blocks, the same six ANN architectures were systematically trained and tested to ensure an unbiased and comprehensive comparative analysis of model performance under different smoothing and memory scenarios.

3.3. ANN Performance Analysis for Daily-Scale Raw Data

Aiming to clarify the modeling approach and identify the most robust topological configuration, six distinct ANN architectures varying in depth and neuron density (as detailed in Table 6) were systematically trained and evaluated using raw daily data. To ensure rigorous model evaluation, the dataset was randomly partitioned into an 80% training set and an independent 20% hold-out test set, as described in Section 2.6. All proposed architectures were modeled to forecast the continuous daily streamflow at the downstream monitoring point of the MRB (station code 66910000). Furthermore, to comprehensively assess the models' predictive performance, the input vectors were structured into two distinct setups: a purely exogenous configuration (Block 1) and an autoregressive configuration (Block 2).

The results (consolidated in Table 6) reveal that, for Block 1, Architectures 3 and 6 achieved the most consistent performance, consistently outperforming other configurations. Architecture 3 (two hidden layers: 15 and 60 neurons) yielded the lowest RMSE ($42.33 \text{ m}^3 \text{ s}^{-1}$) and one of the highest NSE values (0.851), while Architecture 6 (three hidden layers: 80, 50, and 25 neurons) achieved a slightly higher R^2 (0.853). Given the marginal differences in performance, both configurations were considered optimal, showing that a structural depth of two to three hidden layers is sufficient to capture the non-linear characteristics of the system. Further increases in complexity beyond these configura-

tions did not yield substantial improvements, suggesting that the model performance reached a plateau where additional neurons or layers provide diminishing returns relative to computational cost.

Table 6. Results of the analysis of the models tested with daily data (Blocks 1 and 2).

Analysis	Architecture	Hidden Layers (Neurons)	Cross-Validation Performance (Mean of 5 Folds)				Final Out-of-Sample Test (Unseen 20%)			
			RMSE	MAE	R ²	NSE	RMSE	MAE	R ²	NSE
Block 1: Baseline Variables	1	10	43.739	28.176	0.790	0.779	42.733	27.559	0.855	0.848
	2	40	43.356	27.596	0.793	0.783	43.204	27.703	0.848	0.845
	3	15, 60	42.745	26.949	0.795	0.789	42.329	28.089	0.852	0.851
	4	50, 30	43.328	27.394	0.792	0.784	43.114	26.957	0.852	0.846
	5	20, 40, 80	43.598	27.460	0.787	0.782	43.973	28.793	0.842	0.839
	6	80, 50, 25	43.200	27.118	0.790	0.786	42.533	26.674	0.853	0.849
Block 2: Baseline Variables + $Q_{dw}(t - 1)$	1	10	11.573	6.822	0.985	0.985	11.433	7.340	0.989	0.989
	2	40	10.833	6.247	0.986	0.986	10.329	6.243	0.991	0.991
	3	15, 60	10.694	6.232	0.987	0.987	10.092	6.145	0.992	0.992
	4	50, 30	10.756	6.419	0.986	0.986	10.374	6.257	0.991	0.991
	5	20, 40, 80	11.109	6.406	0.985	0.985	10.661	6.278	0.990	0.990
	6	80, 50, 25	10.552	6.354	0.987	0.987	9.867	6.039	0.992	0.992

Baseline Variables = standard set of input features defined in the immediately preceding odd-numbered block; $Q_{dw}(t - 1)$ = downstream stream flow from the previous day. The cross-validation metrics represent the aggregated mean (RMSE, MAE, R², and NSE) across the 5 temporal folds during the training phase, while the final test metrics represent the model's generalization on the completely isolated 20% dataset.

The superiority of this multi-layered approach over simpler models corroborates the findings of Kassem et al. [51], who demonstrated that the structural depth of ANNs allows for superior mapping of complex non-linear rainfall–runoff relationships compared to traditional statistical methods such as ARIMA, which often fail to capture extreme daily variability. The validation hydrograph and scatter plot for Block 1 (Figure 9A) show that the ANN successfully captured recession limbs and baseflow periods, which is vital for water resource management in the Pantanal biome. This precision during dry periods (MAE of 28.09 m³ s^{−1}) aligns with the conclusions of Apayadin et al. [45], who highlighted the ability of ANNs to maintain high accuracy for baseflow discharge in basins with low in situ monitoring density.

However, it should be noted that a systematic underestimation of extreme peaks above 500 m³ s^{−1} was observed. This is evidenced by the scatter plot in Figure 9A, where observed values above this threshold can be seen to consistently fall below the 1:1 parity line. As discussed by Zhang et al. [52] in their recent investigation into flood forecasting, static MLP architectures often struggle with rapid response events because they lack an explicit temporal memory structure, treating satellite inputs as isolated snapshots rather than a continuous physical sequence. Furthermore, as observed by Apayadin et al. [45], this underestimation bias during flood events is typically exacerbated by the lower mathematical frequency of catastrophic extremes in historical training series.

This limitation was fully overcome in the second experimental stage. The inclusion of the downstream streamflow from the previous day $Q_{dw}(t - 1)$, as a predictive variable in Block 2 resulted in a massive reduction in model uncertainty (Table 6). In Architecture 3, the RMSE dropped by 76%, while the NSE and R² surged to 0.992. This substantial leap in performance is governed by the concept of hydrological memory [28,29]; that is, the physical principle that streamflow on any given day is highly dependent on the flow from the preceding day. By feeding the model with this autoregressive $t - 1$ variable, the network receives a physical “anchor” that implicitly accounts for unmeasured factors, such as accumulated soil moisture, groundwater contribution, and the basin's immediate storage state. Consequently, as demonstrated in the unified scatter plot (Figure 9B), the inclusion of this term eliminated the structural lag and the underestimation bias observed in Block 1, with data points tightly clustered along the 1:1 identity line even for high-flow events.

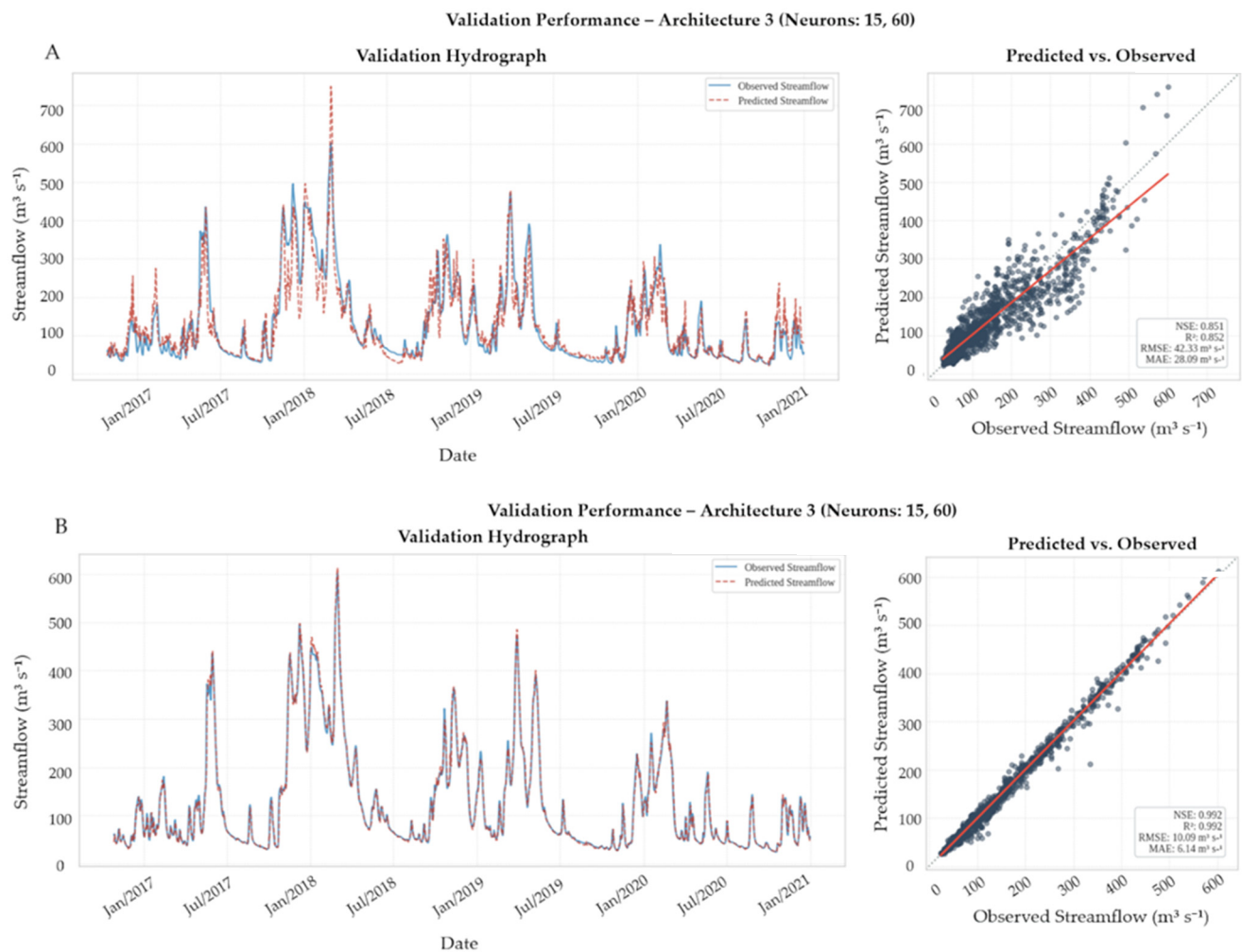


Figure 9. Performance evaluation of ANN Architecture 3 during the validation period (2017–2020). (A) Block 1 with baseline variables; (B) Block 2 with baseline variables + $Q_{dw}(t - 1)$.

3.4. ANN Performance Analysis for 10–Day Moving Average Smoothing

As summarized in Table 7 and illustrated in Figure 10A, this smoothing technique effectively attenuated the high-frequency noise inherent in meteorological inputs. The application of a 10–day moving average filter (Block 3) produced a substantial improvement in the predictive performance of the evaluated ANNs compared to the raw daily data scenarios. Consequently, the architectures were able to better assimilate the underlying hydrological memory of the basin. The evaluated models demonstrated strong generalization capabilities, with NSE and R^2 values reaching over 0.90 in the independent Out–of-Sample Test set for most architectures (Table 7) while cross-validation metrics remained robust, indicating a significant improvement over the raw daily data scenarios.

Table 7. Results of the analysis of the models tested with 10–day moving average-smoothed daily data (Blocks 3 and 4).

Analysis	Architecture	Hidden Layers (Neurons)	Cross-Validation Performance (Mean of 5 Folds)				Final Out–of-Sample Test (Unseen 20%)			
			RMSE	MAE	R^2	NSE	RMSE	MAE	R^2	NSE
Block 3: Baseline Variables	1	10	37.171	23.424	0.855	0.840	34.013	22.685	0.906	0.904
	2	40	36.504	22.688	0.856	0.847	33.120	21.979	0.909	0.909
	3	15, 60	36.272	22.394	0.857	0.848	32.736	20.969	0.916	0.911
	4	50, 30	37.433	22.960	0.852	0.838	34.242	22.035	0.911	0.903
	5	20, 40, 80	36.171	22.742	0.855	0.850	33.637	21.913	0.907	0.906
	6	80, 50, 25	36.378	22.738	0.859	0.848	33.540	21.785	0.914	0.907

Table 7. Cont.

Analysis	Architecture	Hidden Layers (Neurons)	Cross-Validation Performance (Mean of 5 Folds)				Final Out-of-Sample Test (Unseen 20%)			
			RMSE	MAE	R ²	NSE	RMSE	MAE	R ²	NSE
Block 4: Baseline Variables + $Q_{dw}(t-1)$	1	10	10.991	6.556	0.986	0.986	11.398	7.194	0.989	0.989
	2	40	10.693	6.228	0.987	0.987	10.804	6.338	0.990	0.990
	3	15, 60	9.977	5.951	0.989	0.988	10.201	6.176	0.992	0.991
	4	50, 30	10.419	6.150	0.987	0.987	10.501	6.220	0.991	0.991
	5	20, 40, 80	10.287	6.206	0.988	0.987	10.571	6.421	0.991	0.991
	6	80, 50, 25	10.419	6.285	0.987	0.987	11.080	6.579	0.990	0.990

Baseline Variables = standard set of input features defined in the immediately preceding odd-numbered block; $Q_{dw}(t-1)$ = downstream stream flow from the previous day. The cross-validation metrics represent the aggregated mean (RMSE, MAE, R², and NSE) across the 5 temporal folds during the training phase, while the final test metrics represent the model's generalization on the completely isolated 20% dataset.

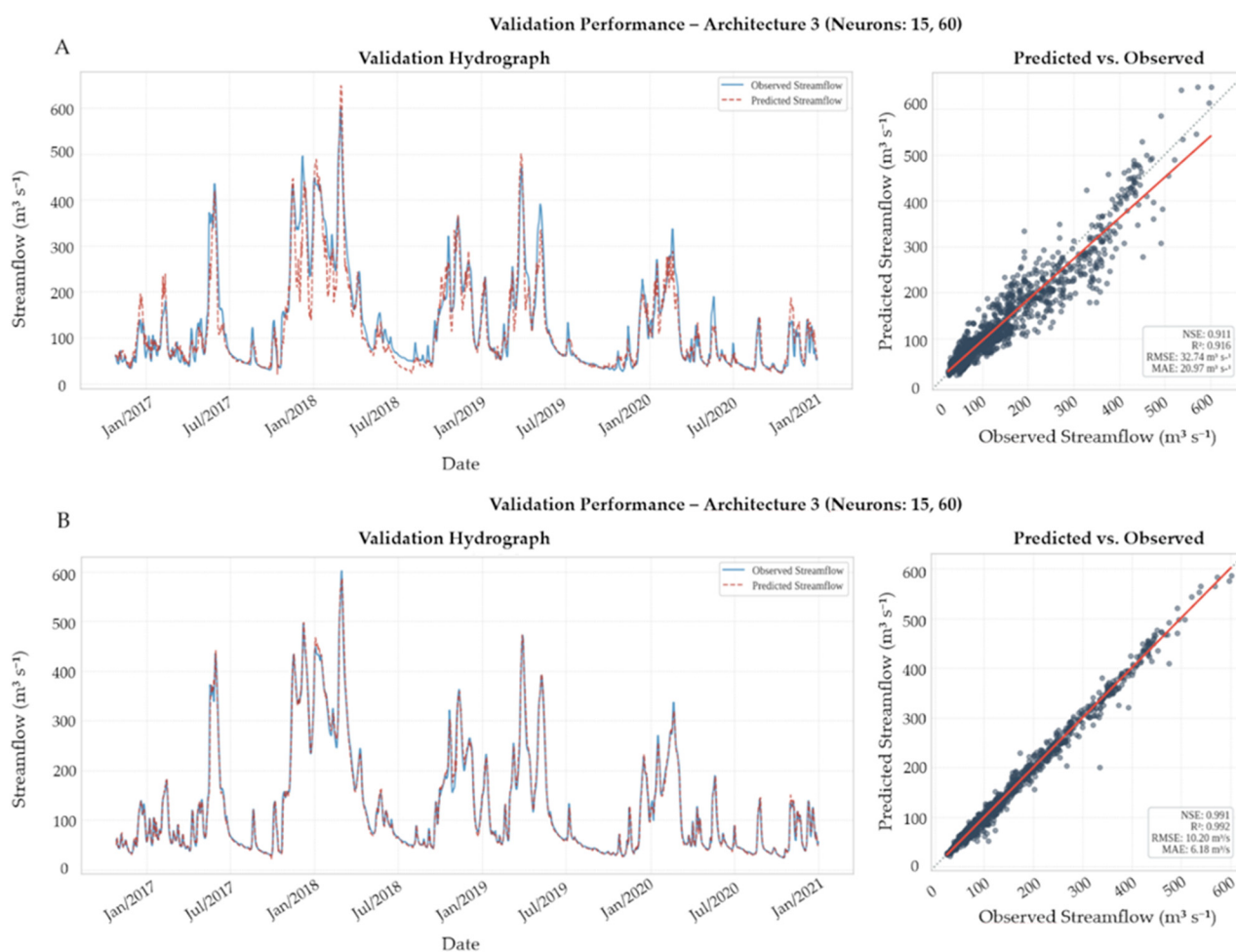


Figure 10. Performance evaluation of ANN Architecture 3 during the validation period (2017–2020). (A) Block 3 with baseline variables; (B) Block 4 with baseline variables + $Q_{dw}(t-1)$.

The improvement of ANN models fed with smoothed data has been widely documented in the literature. For example, Wu and Chau [32] applied a moving average smoothing methodology to daily and monthly precipitation data in the pre-processing stage prior to training MLP neural networks for time-series forecasting. The authors concluded that the smoothing method used for series pre-processing resulted in considerable improvements in model performance for rainfall forecasting.

Despite the overall metric improvement in Block 3, a natural consequence of wider moving average windows is the inherent flattening of short-term extremes. Consequently, the isolated 10-day moving average-smoothed models still presented a slight attenuation

of the sharpest flood peaks. Mehr and Kahya [33] utilized moving averages with 3-, 5-, and 7-day windows to smooth daily streamflow data, finding general improvements in forecasting performance but highlighting that larger smoothing windows prevent original peak flows from being adequately captured. Therefore, when developing comprehensive streamflow forecasting models that aim to accurately capture both wet and dry seasonal variations, this trade-off underscores the importance of evaluating multiple input configurations. In this study, exploring distinct block scenarios allowed for a detailed assessment of how data smoothing reduces satellite-derived noise, while the subsequent integration of autoregressive terms serves precisely to counteract peak flattening and ensure high-flow predictive accuracy.

The most expressive predictive leap materialized in Block 4, with the incorporation of the previous day's downstream streamflow $Q_{dw}(t - 1)$ into the 10-day smoothed data. This addition provided the ANN with a highly correlated, autoregressive physical proxy reflecting the immediate antecedent state of the system. As demonstrated in the results (Table 7 and Figure 10B), the error metrics plummeted, with the NSE value surpassing 0.98. In this scenario, the previous limitation regarding peak flow underestimation was mitigated, with the regression trend line better aligning with the 1:1 identity line. These results indicate that the synergistic combination of temporal smoothing to filter meteorological noise and a direct autoregressive proxy to maintain physical memory allows for the development of a relatively simple MLP architecture that achieves accurate daily streamflow predictions.

3.5. ANN Performance Analysis for 15-Day Moving Average Smoothing

The predictive performance of the neural network architectures trained with a 15-day moving average filter is summarized in Table 8. A systematic analysis of the statistical metrics reveals that Block 5 (Baseline Variables input with 15-day smoothing) achieved high stability and robust generalization across all evaluated topologies, with NSE and R^2 values consistently exceeding 0.92. Again, Architecture 3 was established as the optimal configuration for this block, exhibiting the lowest RMSE ($30.63 \text{ m}^3 \text{ s}^{-1}$) and MAE ($19.46 \text{ m}^3 \text{ s}^{-1}$). These numerical indicators confirm that expanding the moving average window to 15 days allowed high-frequency fluctuations and satellite-derived meteorological noise to be successfully filtered out, allowing the networks to map the long-term seasonal trends of the basin with a high degree of statistical accuracy.

Table 8. Result of the analysis of the models tested with 15-day moving average-smoothed daily data (Blocks 5 and 6).

Analysis	Architecture	Hidden Layers (Neurons)	Cross-Validation Performance (Mean of 5 Folds)				Final Out-of-Sample Test (Unseen 20%)			
			RMSE	MAE	R^2	NSE	RMSE	MAE	R^2	NSE
Block 5: Baseline Variables	1	10	32.533	20.419	0.888	0.877	30.337	21.510	0.926	0.924
	2	40	33.744	20.950	0.880	0.867	31.588	20.816	0.919	0.917
	3	15, 60	32.977	20.367	0.882	0.875	30.901	20.320	0.923	0.920
	4	50, 30	32.214	20.069	0.887	0.880	31.181	20.526	0.923	0.919
	5	20, 40, 80	32.593	20.552	0.882	0.878	32.157	21.539	0.916	0.914
	6	80, 50, 25	32.313	20.335	0.888	0.880	32.213	20.756	0.919	0.914
Block 6: Baseline Variables + $Q_{dw}(t - 1)$	1	10	11.065	6.503	0.985	0.985	9.899	5.718	0.992	0.992
	2	40	10.569	6.233	0.987	0.987	10.161	6.121	0.992	0.991
	3	15, 60	10.717	6.226	0.986	0.986	10.184	6.148	0.992	0.991
	4	50, 30	10.702	6.291	0.987	0.986	10.238	6.204	0.992	0.991
	5	20, 40, 80	10.426	6.118	0.987	0.987	10.223	6.187	0.991	0.991
	6	80, 50, 25	10.568	6.162	0.987	0.987	11.214	6.587	0.990	0.989

Baseline Variables = standard set of input features defined in the immediately preceding odd-numbered block; $Q_{dw}(t - 1)$ = downstream stream flow from the previous day. The cross-validation metrics represent the aggregated mean (RMSE, MAE, R^2 , and NSE) across the 5 temporal folds during the training phase, while the final test metrics represent the model's generalization on the completely isolated 20% dataset.

Notably, the numerical excellence of Block 5 was fundamentally transformed in Block 6 with the integration of the previous day's downstream streamflow $Q_{dw}(t-1)$. As shown in Table 8, the inclusion of this autoregressive proxy triggered a massive performance leap, causing the RMSE of Architecture 3 to plummet by approximately 67% (dropping to $10.15 \text{ m}^3 \text{ s}^{-1}$) while the NSE and R^2 values both surged to an exceptional 0.991. These results demonstrate that, even with a broader 15-day smoothing window applied to the climate inputs, the model still requires the immediate physical state of the system ($t-1$) lag to minimize residual uncertainties and accurately establish the baseline daily streamflow.

The physical implications of these two scenarios are clearly reflected in the validation hydrographs and scatter plots presented in Figure 11. For Block 5 (Figure 11A), the visual alignment confirms that while the 15-day moving average filter ensured accurate simulation of the recession limbs and dry-season baseflow, it induces a severe structural trade-off resulting in pronounced flattening and underestimation of extreme flood peaks, particularly those around $400 \text{ m}^3 \text{ s}^{-1}$ or higher. This visual lag and peak attenuation occur because the use of a wider temporal window smooths out the rapid, heavy precipitation events that trigger sharper responses in the watershed.

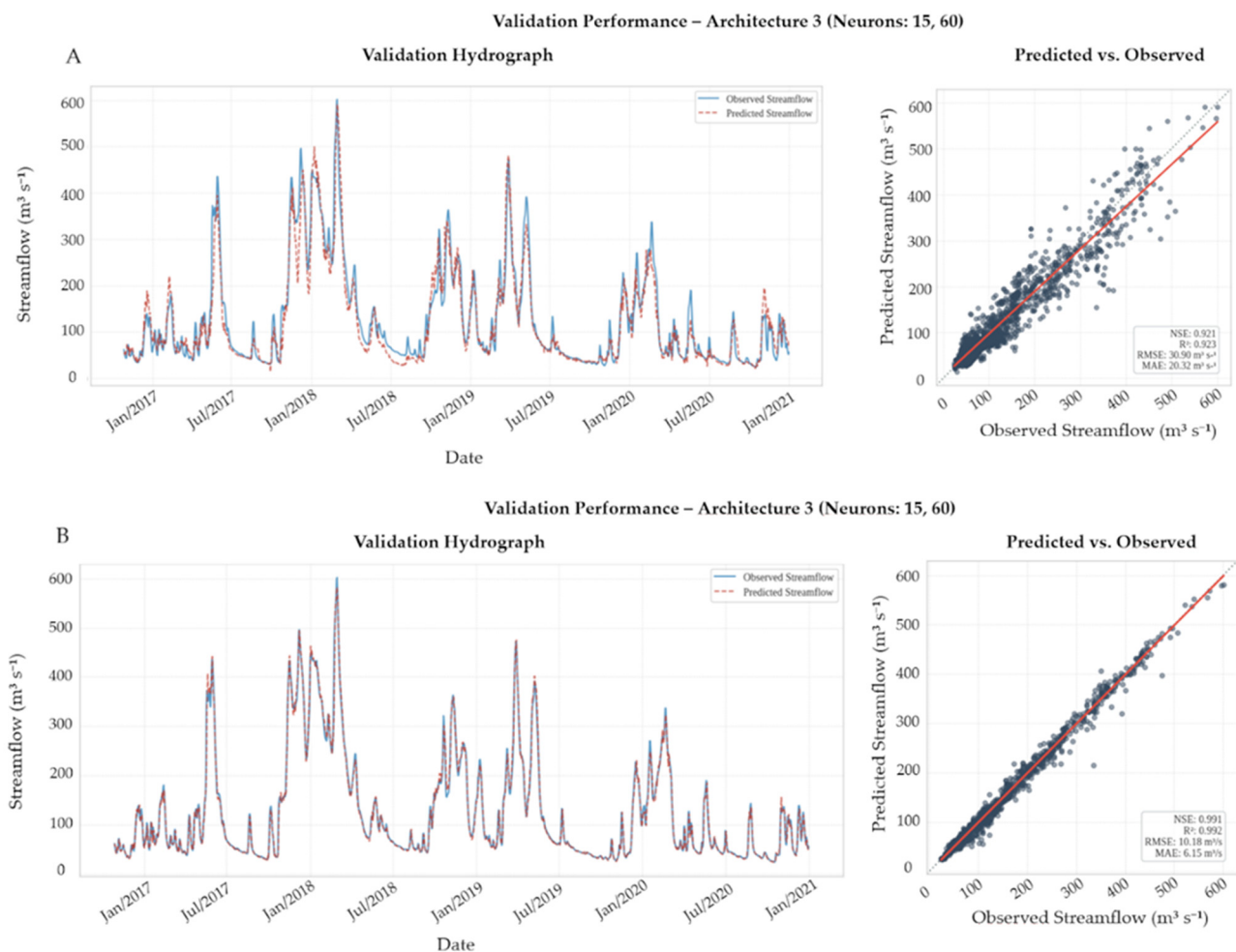


Figure 11. Performance evaluation of ANN Architecture 3 during the validation period (2017–2020). (A) Block 5 with baseline variables; (B) Block 6 with baseline variables + $Q_{dw}(t-1)$.

This critical limitation was effectively mitigated in Block 6 (Figure 11B). The hydrograph demonstrates that the inclusion of the autoregressive variable effectively acted as a physical corrector, bringing the simulated peaks back into perfect alignment (in terms of both phase and magnitude) with the observed streamflow data. Furthermore, the uni-

fied scatter plot shows a striking contraction of the data points, which became tightly clustered along the 1:1 identity line. This visual and statistical harmony underscores that the synergistic combination of a broad 15-day moving average filter (to eliminate noise) and the immediate antecedent streamflow indicator $Q_{dw}(t - 1)$ provides an optimized predictive framework.

In fact, in this advanced modeling scenario, even the simplest evaluated topology (Architecture 1, with a single 10-neuron layer) achieved high metrics. This finding suggests that appropriate temporal data conditioning simplifies mapping of the river's non-linear characteristics, demonstrating that shallow neural networks are sufficient to achieve high predictive accuracy. This enables maximization of computational efficiency without compromising model performance, avoiding the need for more complex structural configurations.

Despite the remarkable statistical performance achieved by the autoregressive models, it is fundamental to recognize the operational limitations of utilizing a $t - 1$ temporal lag for real-time disaster forecasting. From a civil defense and risk mitigation perspective, predicting an extreme hydrological event with only 24 h of advance notice—the practical limit when relying on the streamflow registered on the immediately preceding day—frequently fails to provide sufficient lead time to mobilize emergency resources, issue evacuation warnings, or implement effective preventive policies in the municipality of Miranda. Within the framework of Flood Early Warning Systems, the literature has emphasized that maximizing the forecast horizon is the most critical and determining factor for the practical effectiveness and success of a public alert [53].

Consequently, a clear trade-off is established between immediate statistical precision and the actual operational utility of the model [54]. Therefore, for the specific purpose of disaster response, architectures that utilize purely exogenous, medium-term predictive horizons (e.g., 5- to 15-day moving average configurations without the $t - 1$ anchor) have the potential to offer superior operational value. Although these exogenous configurations face greater uncertainty and a slight attenuation in the statistical modeling of peak flows, they may effectively overcome the reactive limitation of the 24 h lag, potentially supplying the vital temporal window required for governmental action [54].

The ultimate contribution of this study's findings lies in establishing a foundational mathematical framework that could support a future Decision Support System (DSS) for the Miranda River sub-basin. While the development of complete operational software is beyond the scope of the current study, the optimized neural network architectures developed herein are suitable for future integration as the analytical engine in such a system [55]. Conceptually, a future DSS could autonomously ingest real-time remote sensing and telemetry data to feed these pre-trained models. The continuous processing of these variables would enable the theoretical establishment of a dynamic dashboard that is capable of projecting streamflows and triggering automated alerts for local authorities. Transitioning these AI-based networks into operational software represents a highly recommended proactive step for future works, with the aim of enhancing the resilience of vulnerable Pantanal communities against extreme climate events [55].

3.6. Ablation Analysis

To definitively quantify the individual contributions of the proposed input modules, an ablation study using Architecture 3 as the reference model was conducted. Table 9 isolates the impacts of the (10- and 15-day) moving average smoothing filters and the lagged autoregressive term $Q_{dw}(t - 1)$ on the overall predictive performance within the test set.

Table 9. Ablation analysis quantifying the contributions of moving average filters and autoregressive inputs to streamflow prediction in the Rio Miranda basin.

Scenario	Smoothing Window	Lagged Variable $Q_{dw}(t - 1)$	NSE	RMSE ($\text{m}^3 \text{s}^{-1}$)
Baseline (Block 1)	None	No	0.851	42.33
Autoregressive Only (Block 2)	None	Yes	0.992	10.09
10-day Smoothing (Block 3)	10 days	No	0.911	32.74
10-day Smoothing + $Q_{dw}(t - 1)$ (Block 4)	10 days	Yes	0.991	10.20
15-day Smoothing (Block 5)	15 days	No	0.920	30.90
15-day Smoothing + $Q_{dw}(t - 1)$ (Block 6)	15 days	Yes	0.991	10.18

The baseline model (Block 1), relying solely on raw exogenous variables without historical streamflow data, establishes the lower bound of performance. The isolated addition of moving average smoothing (Blocks 3 and 5) yields noticeable improvements in the performance metrics, highlighting these filters' ability to mitigate daily noise and better represent the slower response times of the watershed, aligning with recent findings that preprocessing with temporal smoothing enhances model stability in data-driven hydrological forecasting [56].

However, the ablation analysis revealed that the lagged downstream streamflow from the previous day, $Q_{dw}(t - 1)$, acts as the primary driver of model accuracy. As anticipated, configurations that incorporated this immediate downstream antecedent yielded the highest predictive accuracies. This behavior is expected due to the strong temporal autocorrelation characteristic of large river basins, where the prior day's streamflow provides the network with a highly accurate baseline state of the system.

When the autoregressive term was introduced, even without prior data smoothing (Block 2), the model exhibited a substantial leap in precision. Ultimately, the synergistic application of both modules (Blocks 4 and 6) yielded the most robust configurations for predicting streamflow in the MRB. The moving average filters ensure stability across recession curves and baseflow periods, while the autoregressive component realistically anchors the magnitude of the predictions, thus optimizing the predictive capacity of the considered MLP models; this outcome is consistent with previous evidence that hybrid architectures combining smoothing filters and autoregressive inputs tend to outperform single-module approaches in operational streamflow prediction contexts [57].

3.7. Limitations and Future Directions

Despite the high predictive performance of the optimized MLP framework, several limitations are acknowledged, which delineate the scope of this study and suggest pathways for future research. Regarding model complexity, while our MLP architecture provides high inference speed and interpretability, offering clear advantages over more opaque recurrent architectures, there is a recognized trade-off regarding flood peak bias. Feedforward networks occasionally struggle to capture the absolute magnitudes of extreme events—a common limitation when models are optimized against aggregate loss functions [58]. Furthermore, while the current study utilized global performance metrics (NSE, RMSE, MAE, and R^2) which align with recent benchmarks for evaluating neural network architectures in streamflow forecasting, we acknowledge that these global indicators might be inadequate in the context of specific hydrological dynamics analysis.

Future research should incorporate more granular diagnostics, such as peak-flow timing errors, model performance during low-flow recession periods, and seasonal be-

haviors across extreme climate events [59]. Additionally, benchmarking the optimized MLP frameworks presented in this study against sequential deep learning models, such as long short-term memory and temporal convolutional networks, remains a valuable future direction. Such a comparison may allow researchers to determine whether the increased computational overhead of these sequential architectures favorably trade off against statistically significant improvements in handling the peak flow biases observed in our models.

In future work, it is recommended to explore the integration of the optimized MLP architectures presented in this study into an automated DSS, as well as coupling these networks with real-time telemetry and broader meteorological forecasts to extend the predictive horizon beyond 24 h. Ultimately, future research should focus on extending these architectures to multi-step-ahead forecasting (1–7 days) and evaluating their viability as components of operational early warning systems for flood risk mitigation in Pantanal communities.

Finally, the cross-basin generalizability of this framework remains untested. The structural parameters and moving average windows were specifically tailored to the hydrological response time of the MRB; thus, applying this framework to basins with different geomorphological or climatic characteristics would necessitate systematic retraining. To mitigate these computational and data requirements, future work should explore transfer learning techniques, allowing the framework to adapt pre-trained weights to adjacent ungauged basins. Importantly, we aim to transition from the current reactive 24 h forecasting window toward long-term multi-step forecasting horizons, which is essential to enhance the operational utility of the model for disaster risk reduction and proactive civil defense planning.

4. Conclusions

This study evaluated Multi-layer Perceptron Artificial Neural Networks (MLP-ANNs) for daily streamflow forecasting in the data-scarce Miranda River sub-basin, utilizing satellite-derived precipitation (CHIRPS) and evapotranspiration (GLDAS) products. The results indicated that data-driven models provide a viable methodological alternative for simulating streamflow dynamics in regions with limited in situ monitoring, contributing to the hydrological framework of the Pantanal biome.

Based on a performance analysis across different input combinations and topological variations, the following main conclusions were drawn:

- (a) Regarding the topological configuration, Architectures 3 (two hidden layers: 15, 60 neurons) and 6 (three hidden layers: 80, 50, 25 neurons) achieved the most consistent performance. This demonstrates that shallow-to-moderately deep networks (up to three hidden layers) are sufficient to systematically capture the non-linear characteristics of this basin, with further structural complexity yielding diminishing returns.
- (b) The application of temporal data conditioning—specifically, 10- and 15-day moving average filters—effectively attenuated the high-frequency variations in the satellite-derived datasets. This pre-processing step allowed the models to better capture baseflow dynamics and recession curves, which are relevant parameters for water resource management during dry periods.
- (c) The inclusion of an autoregressive variable—namely, downstream streamflow from the previous day, $Q_{dw}(t - 1)$ —significantly improved the predictive metrics (e.g., NSE reaching up to 0.99 on the test set) and mitigated the underestimation bias observed for extreme flood peaks ($>500 \text{ m}^3 \text{ s}^{-1}$). This confirms the importance of antecedent system memory in the mathematical mapping of rapid hydrological responses.

- (d) An operational trade-off was identified between statistical accuracy and forecast horizon. While the inclusion of the $Q_{dw}(t - 1)$ variable maximized model precision, it inherently limits the forecast to a 24 h lead time. Conversely, purely exogenous configurations (i.e., those using moving averages without the autoregressive term) present lower peak flow accuracy but offer extended predictive horizons.

Author Contributions: Conceptualization, C.P.S.B., S.N.D., C.R.P. and P.A.A.M.; methodology, C.P.S.B., S.N.D., C.R.P., R.A.F.R. and P.A.A.M.; software, C.P.S.B., C.R.P., R.A.F.R., M.R.B. and P.A.A.M.; validation, C.P.S.B., C.R.P. and P.A.A.M.; formal analysis, C.P.S.B., S.N.D., C.R.P., R.A.F.R., V.V.e.V., M.R.B. and P.A.A.M.; investigation, C.P.S.B., S.N.D., C.R.P. and P.A.A.M.; resources, P.A.A.M.; data curation, C.P.S.B., C.R.P., V.V.e.V. and P.A.A.M.; writing—original draft preparation, C.P.S.B. and P.A.A.M.; writing—review and editing, C.P.S.B., S.N.D., C.R.P., R.A.F.R., V.V.e.V., M.R.B. and P.A.A.M.; visualization, C.P.S.B. and P.A.A.M.; supervision, P.A.A.M.; project administration, P.A.A.M. All authors have read and agreed to the published version of the manuscript.

Funding: The authors acknowledge CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior [grant number 88887.817104/2023-00] and PROEX—Finance Code 001); FEALQ (Fundação de Estudos Agrários Luiz de Queiroz); INCT Combate à Fome CNPq 406774/2022-6; the Center for Artificial Intelligence and Machine Learning (CIAAM); and the Center of Artificial Intelligence (C4AI/IBM process 2019/07665-4) for supporting this research.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data will be made available on request.

Conflicts of Interest: Author Carlos Roberto Padovani were employed by the company Embrapa Pantanal, Brazil. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AI	Artificial Intelligence
ANA	National Water Agency
ANN	Artificial Neural Network
ET	Evapotranspiration
GIS	Geographic Information System
GLDAS	Global Land Data Assimilation System
MAE	Mean Absolute Error
ML	Machine Learning
MLP	Multi-layer Perceptron
MLP-ANN	Multi-layer Perceptron Artificial Neural Network
MRB	Miranda River Basin
MRW	Miranda River Watershed
NSE	Nash–Sutcliffe coefficient
PPT	Precipitation
Q	Streamflow
QGIS	Quantum Geographic Information System
R^2	Coefficient of Determination
ReLU	Rectified Linear Unit
RMSE	Root Mean Square Error
SNIRH	National Water Resources Information System
WMO	World Meteorological Organization

References

1. Silva, V.O.; De Mello, C.R.; Chou, S.C.; Guo, L.; Viola, M.R. Characteristics of extreme meteorological droughts over the Brazilian Pantanal throughout the 21st century. *Front. Water* **2024**, *6*, 1385077. [\[CrossRef\]](#)
2. Neto, A.R.; da Paz, A.R.; Marengo, J.A.; Chou, S.C. Hydrological processes and climate change in hydrographic regions of Brazil. *J. Water Resour. Prot.* **2016**, *8*, 1103–1127. [\[CrossRef\]](#)
3. Pechlivanidis, I.G.; Jackson, B.M.; McIntyre, N.R.; Wheeler, H.S. Catchment scale hydrological modelling: A review of model types, calibration approaches and uncertainty analysis methods in the context of recent developments in technology and applications. *Glob. NEST J.* **2011**, *13*, 193–214.
4. Devia, G.K.; Ganasri, B.P.; Dwarakish, G.S. A review on hydrological models. *Aquat. Procedia* **2015**, *4*, 1001–1007. [\[CrossRef\]](#)
5. Cristaldo, M.F.; Souza, C.C.; Jesus, L.; Padovani, C.R.; Oliveira, P.T.S.; Viganó, H.H.G. Analysis and distribution of the rainfall monitoring network in a Brazilian Pantanal region. *Rev. Bras. Meteorol.* **2017**, *32*, 199–205. [\[CrossRef\]](#)
6. Thielen, D.; Ramoni-Perazzi, P.; Puche, M.L.; Márquez, M.; Quintero, J.I.; Rojas, W.; Soto-Werschitz, A.; Thielen, K.; Nunes, A.; Libonati, R. The Pantanal under siege—On the origin, dynamics and forecast of the megadrought severely affecting the largest wetland in the world. *Water* **2021**, *13*, 3034. [\[CrossRef\]](#)
7. Diniz, L.S. Regionalização de Parâmetros de Modelo Chuva-Vazão Usando Redes Neurais. Ph.D. Thesis, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil, 2008.
8. Ruezzene, C.B.; Miranda, R.B.; Tech, A.R.B.; Mauad, F.F. Preenchimento de falhas em dados de precipitação através de métodos tradicionais e por inteligência artificial. *Rev. Bras. Climatol.* **2021**, *29*, 177–204.
9. Uliana, E.M.; Silva, D.D.; Moreira, M.C.; Pereira, D.R.; Almeida, F.T. Modelo hidrológico híbrido para previsão de vazões na bacia do Rio Piracicaba-MG. *Rev. Bras. Meteorol.* **2020**, *34*, 471–480.
10. Santos Neto, L.A.; Maniesi, V.; Querino, C.A.S.; Silva, M.J.G.; Brown, V.R. Modelagem hidroclimatológica utilizando redes neurais multilayer perceptron em bacia hidrográfica no sudoeste da Amazônia. *Rev. Bras. Climatol.* **2020**, *26*, 696–711.
11. Mendonça, L.M.; Gomide, I.S.; Souza, J.V.; Blanco, C.J.C. Modelagem chuva-vazão via redes neurais artificiais para simulação de vazões de uma bacia hidrográfica da Amazônia. *Rev. Gest. Água Am. Lat.* **2021**, *18*, e2. [\[CrossRef\]](#)
12. Santos, J.R.S.; Celeste, A.B. Avaliação de estratégias de modelagem guiada por dados para previsão de vazão em rio sergipano. *Rev. Ambient. Água* **2014**, *9*, 555–566. [\[CrossRef\]](#)
13. Patil, A.P.; Deka, P.C. An extreme learning machine approach for modeling evapotranspiration using extrinsic inputs. *Comput. Electron. Agric.* **2016**, *121*, 385–392. [\[CrossRef\]](#)
14. Mageto, G.A.S.; Silva, A.G.; Bulgarelli, R.F.; Bublit, C.F.; Valente, A.C.; Costa, D.A.G. Artificial intelligence applications in the agriculture 4.0. *Rev. Ciênc. Agron.* **2020**, *51*, e20207701. [\[CrossRef\]](#)
15. McCulloch, W.S.; Pitts, W. A logical calculus of the ideas immanent in nervous activity. *Bull. Math. Biophys.* **1943**, *5*, 115–133. [\[CrossRef\]](#)
16. Rosenblatt, F. The perceptron: A probabilistic model for information storage and organization in the brain. *Psychol. Rev.* **1958**, *65*, 386–408. [\[CrossRef\]](#) [\[PubMed\]](#)
17. Ali, S.; Shahbaz, M. Streamflow forecasting by modeling the rainfall–streamflow relationship using artificial neural networks. *Model. Earth Syst. Environ.* **2020**, *6*, 1645–1656. [\[CrossRef\]](#)
18. Yaseen, Z.M.; El-shafie, A.; Jaafar, O.; Afan, H.A.; Sayl, K.N. Artificial intelligence based models for stream-flow forecasting: 2000–2015. *J. Hydrol.* **2015**, *530*, 829–844. [\[CrossRef\]](#)
19. Ibrahim, K.S.M.H.; Huang, Y.F.; Ahmed, A.N.; Koo, C.H.; El-Shafie, A. A review of the hybrid artificial intelligence and optimization modelling of hydrological streamflow forecasting. *Alex. Eng. J.* **2022**, *61*, 279–303. [\[CrossRef\]](#)
20. Zounemat-Kermani, M.; Matta, E.; Cominola, A.; Xia, X.; Zhang, Q.; Liang, Q.; Hinkelmann, R. Neurocomputing in surface water hydrology and hydraulics: A review of two decades retrospective, current status and future prospects. *J. Hydrol.* **2020**, *588*, 125085. [\[CrossRef\]](#)
21. Collischonn, B.; Allasia, D.; Collischonn, W.; Tucci, C.E.M. Desempenho do satélite TRMM na estimativa de precipitação sobre a bacia do Paraguai superior. *Rev. Bras. Cartogr.* **2007**, *59*, 93–99.
22. Padovani, C.R. Dinâmica Espaço-Temporal das Inundações do Pantanal. Ph.D. Thesis, Escola Superior de Agricultura Luiz de Queiroz, Piracicaba, Brazil, 2010.
23. Guimarães, D.P.; Landau, E.C.; Santos, M.C.B.; Mendes, S.H.G.S. Caracterização das chuvas no Pantanal Mato-grossense. In *Anais do VII Simpósio de Geotecnologias no Pantanal*; Embrapa Informática Agropecuária/INPE: Jardim, Brazil, 2018; pp. 555–562.
24. Pereira, G.; Ramos, R.C.; Rocha, L.C.; Brunsell, N.A.; Merino, E.R.; Mataveli, G.A.V.; Cardozo, F.S. Rainfall patterns and geomorphological controls driving inundation frequency in tropical wetlands: How does the Pantanal flood? *Prog. Phys. Geogr.* **2021**, *45*, 669–686. [\[CrossRef\]](#)
25. Marengo, J.A.; Jimenez, J.C.; Espinoza, J.C.; Cunha, A.P.; Aragão, L.E.O. Increased climate pressure on the agricultural frontier in the Eastern Amazonia–Cerrado transition zone. *Sci. Rep.* **2022**, *12*, 457. [\[CrossRef\]](#) [\[PubMed\]](#)

26. De Barros de Sousa, L.; Montenegro, A.A.d.A.; da Silva, M.V.; Almeida, T.A.B.; de Carvalho, A.A.; da Silva, T.G.F.; de Lima, J.L.M.P. Spatiotemporal analysis of rainfall and droughts in a semiarid basin of Brazil: Land use and land cover dynamics. *Remote Sens.* **2023**, *15*, 2550. [\[CrossRef\]](#)
27. Ruhoff, A.; de Andrade, B.C.; Laipelt, L.; Fleischmann, A.S.; Siqueira, V.A.; Moreira, A.A.; Barbedo, R.; Cyganski, G.L.; Fernandez, G.M.R.; Brêda, J.P.L.F.; et al. Global evapotranspiration datasets assessment using water balance in South America. *Remote Sens.* **2022**, *14*, 2526. [\[CrossRef\]](#)
28. Mendoza, L.A.F.; Lima, A.G.G.; de Mello, H.D., Jr.; Pacheco, M.A.C. Data-Driven Detection of Climate–Streamflow Dependencies and Multi-Year Hydrological Memory in Brazilian Reservoir Systems. *Water* **2026**, submitted. [\[CrossRef\]](#)
29. Berghuijs, W.R.; Woods, R.A.; Anderson, B.J.; de Sánchez, A.L.H.; Hrachowitz, M. Annual memory in the terrestrial water cycle. *Hydrol. Earth Syst. Sci.* **2025**, *29*, 1319–1333. [\[CrossRef\]](#)
30. Linares-Rodriguez, A.; Lara-Fanego, V.; Pozo-Vazquez, D.; Tovar-Pescador, J. One-day-ahead streamflow forecasting using artificial neural networks and a meteorological mesoscale model. *J. Hydrol. Eng.* **2015**, *20*, 05015001. [\[CrossRef\]](#)
31. Yilmaz, M.U.; Onoz, B. A comparative study of statistical methods for daily streamflow estimation at ungauged basins in Turkey. *Water* **2020**, *12*, 459. [\[CrossRef\]](#)
32. Wu, C.L.; Chau, K.W. Prediction of rainfall time series using modular soft computing methods. *Eng. Appl. Artif. Intell.* **2013**, *26*, 997–1007. [\[CrossRef\]](#)
33. Mehr, A.D.; Kahya, E. A Pareto-optimal moving average multigene genetic programming model for daily streamflow prediction. *J. Hydrol.* **2017**, *549*, 603–615. [\[CrossRef\]](#)
34. Wu, C.L.; Chau, K.W.; Li, Y.S. Methods to improve neural network performance in daily flows prediction. *J. Hydrol.* **2009**, *372*, 80–93. [\[CrossRef\]](#)
35. Tucci, C.E.M.; Clarke, R.T. Impacto das mudanças da cobertura vegetal no escoamento: Revisão. *Rev. Bras. Recur. Hídricos* **1997**, *2*, 135–152. [\[CrossRef\]](#)
36. WMO. *Guide to Hydrological Practices: Volume I Hydrology—From Measurement to Hydrological Information*; WMO: Geneva, Switzerland, 2008.
37. Santos Antoniassi, R.A. Predição de Nível de Rios da Região Hidrográfica do Rio Paraguai Utilizando Algoritmos de Aprendizado de Máquina. Master's Dissertation, Universidade Federal de Mato Grosso do Sul, Campo Grande, Brazil, 2022.
38. Ng, W.W.; Panu, U.S.; Lennox, W.C. Chaos based analytical techniques for daily extreme hydrological observations. *J. Hydrol.* **2007**, *342*, 17–41. [\[CrossRef\]](#)
39. Bertoni, J.C.; Tucci, C.E.M. Precipitação. In *Hidrologia: Ciência e Aplicação*, 2nd ed.; Tucci, C.E.M., Ed.; UFRGS: Porto Alegre, Brazil, 2007; pp. 177–241.
40. Hanoon, M.S.; Abdullatif, A.A.; Ahmed, A.N.; Razzaq, A.; Birima, A.H.; El-Shafie, A. A comparison of various machine learning approaches performance for prediction suspended sediment load of river systems: A case study in Malaysia. *Earth Sci. Inform.* **2022**, *15*, 91–104. [\[CrossRef\]](#)
41. Haykin, S. *Neural Networks: A Comprehensive Foundation*; Prentice Hall: Englewood Cliffs, NJ, USA, 1994.
42. Jung, K.; Bae, D.H.; Um, M.J.; Kim, S.; Jeon, S.; Park, D. Evaluation of nitrate load estimations using neural networks and canonical correlation analysis with k-fold cross-validation. *Sustainability* **2020**, *12*, 400. [\[CrossRef\]](#)
43. Rohani, A.; Taki, M.; Abdollahpour, M. A novel soft computing model (Gaussian process regression with K-fold cross validation) for daily and monthly solar radiation forecasting (Part: I). *Renew. Energy* **2018**, *115*, 411–422. [\[CrossRef\]](#)
44. Salem, H.; Kabeel, A.E.; El-Said, E.M.S.; Elzeki, O.M. Predictive modelling for solar power-driven hybrid desalination system using artificial neural network regression with Adam optimization. *Desalination* **2022**, *522*, 115411. [\[CrossRef\]](#)
45. Apaydin, H.; Feizi, H.; Sattari, M.T.; Colak, M.S.; Shamshirband, S.; Chau, K.W. Comparative analysis of recurrent neural network architectures for reservoir inflow forecasting. *Water* **2020**, *12*, 1500. [\[CrossRef\]](#)
46. Oliveira, A.R.; Ramos, T.B.; Neves, R. Streamflow estimation in a mediterranean watershed using neural network models: A detailed description of the implementation and optimization. *Water* **2023**, *15*, 947. [\[CrossRef\]](#)
47. Zakaria, M.N.A.; Ahmed, A.N.; Malek, M.A.; Birima, A.H.; Khan, M.M.H.; Sherif, M.; Elshafie, A. Exploring machine learning algorithms for accurate water level forecasting in Muda river, Malaysia. *Heliyon* **2023**, *9*, e17993. [\[CrossRef\]](#) [\[PubMed\]](#)
48. de Oliveira-Júnior, J.F.; da Silva Junior, C.A.; Teodoro, P.E.; Rossi, F.S.; Blanco, C.J.C.; Lima, M.; de Gois, G.; Correia Filho, W.L.F.; de Barros Santiago, D.; dos Santos Vanderley, M.H.G. Confronting CHIRPS dataset and in situ stations in the detection of wet and drought conditions in the Brazilian Midwest. *Int. J. Climatol.* **2021**, *41*, 4478–4493. [\[CrossRef\]](#)
49. Zhang, Q.; Li, Z.; Dong, Y.; Wang, H.; Wang, Y.; Li, Z.; Feng, Q.; Huang, H. Spatiotemporal Characteristics and Identification of Typical Hydrological Patterns of Interval Inflow in the Three Gorges Reservoir Basin, China. *Hydrology* **2026**, *13*, 75. [\[CrossRef\]](#)
50. Sudheer, K.P.; Gosain, A.K.; Ramasastri, K.S. A data-driven algorithm for constructing artificial neural network rainfall-runoff models. *Hydrol. Process.* **2002**, *16*, 1325–1330. [\[CrossRef\]](#)
51. Kassem, A.A.; Raheem, A.M.; Khidir, K.M. Daily streamflow prediction for Khazir river basin using ARIMA and ANN models. *Zanco J. Pure Appl. Sci.* **2020**, *32*, 30–39. [\[CrossRef\]](#)

52. Zhang, Y.; Zhou, Z.; Deng, Y.; Pan, D.; Van Griensven Thé, J.; Yang, S.X.; Gharabaghi, B. Daily Streamflow Forecasting Using Networks of Real-Time Monitoring Stations and Hybrid Machine Learning Methods. *Water* **2024**, *16*, 1284. [[CrossRef](#)]
53. Mehmood, H.; Rasmy, M. Challenges and technical advances in flood early warning systems (FEWSs). In *Flood Impact Mitigation and Resilience Enhancement*; IntechOpen: London, UK, 2020; p. 19.
54. Mosavi, A.; Ozturk, P.; Chau, K.W. Flood prediction using machine learning models: Literature review. *Water* **2018**, *10*, 1536. [[CrossRef](#)]
55. Sit, M.; Demiray, B.Z.; Xiang, Z.; Ewing, G.J.; Sermet, Y.; Demir, I. A comprehensive review of deep learning applications in hydrology and water resources. *Water Sci. Technol.* **2020**, *82*, 2635–2670. [[CrossRef](#)] [[PubMed](#)]
56. Hammond, J.; Goodling, P.; Diaz, J.; Corson-Dosch, H.; Heldmyer, A.; Hamshaw, S.; McShane, R.; Ross, J.; Sando, R.; Simeone, C.; et al. Machine learning generated streamflow drought forecasts for the conterminous United States. *Front. Water* **2026**, *7*, 1709138. [[CrossRef](#)]
57. Khan, M.M.; Sarwar, M.K.; Zafar, M.A.; Rashid, M.; Tariq, M.A.U.R.; Haider, S.; Okasha, A.M.; Dewidar, A.Z.; Mattar, M.A.; Salem, A. Comparative analysis of inflow forecasting using machine learning and statistical techniques: Case study of Mangla reservoir and Marala Headworks. *Front. Environ. Sci.* **2025**, *13*, 1590346. [[CrossRef](#)]
58. Jesus, G.; Mardani, Z.; Alves, E.; Oliveira, A. Deep Learning-Based River Flow Forecasting with MLPs. *Sensors* **2025**, *25*, 2154. [[CrossRef](#)] [[PubMed](#)]
59. Yang, Y.; Pan, M.; Feng, D.; Xiao, M.; Dixon, T.; Hartman, R.; Shen, C.; Song, Y.; Sengupta, A.; Delle Monache, L.; et al. Improving streamflow simulation through machine learning: DI-LSTM and operational limitations. *Hydrol. Earth Syst. Sci.* **2025**, *29*, 5453–5470.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.