HOLOS
ISSN 1807 - 1600

UTILIZING LONG SHORT-TERM MEMORY (LSTM) NETWORKS FOR RIVER FLOW PREDICTION IN THE BRAZILIAN PANTANAL BASIN

C. S. DESCOVI*, A. C. ZUFFO, S.M. MOHAMMADIZADEH, L. F. MURILLO-BERMÚDEZ, D. A. SIERRA
State University of Campinas (UNICAMP)
ORCID ID: https://orcid.org/0000-0002-2847-2064*
cassianodescovi@fec.unicamp.br *

Submitted June 5, 2023 – Accepted September 11, 2023

DOI: 10.15628/holos.2023.1631511, 2023

ABSTRACT

This article demonstrates the successful application of Long Short-Term Memory (LSTM) recurrent neural networks to simulate streamflow in the Aquidauana River basin, located in the Brazilian Pantanal. The LSTM network used daily precipitation data as input to predict future streamflow in the region. The results obtained from this research show a coefficient of determination (R^2) of 0.82, indicating a strong fit of the model to the observed data. Additionally, the Root Mean Squared Error (RMSE) was found to be 0.53, indicating the model's accuracy in predicting streamflow compared to actual data. These findings highlight the effectiveness of LSTM networks in hydrological modeling for the Pantanal region, which is crucial for water resource planning and sustainable management in this ecologically significant area. This study is expected to serve as a catalyst for further research and make a substantial contribution to the advancement of streamflow prediction techniques in complex watersheds such as the Aquidauana River basin.

KEYWORDS: LSTM, River flows, Pantanal.

REDES DE MEMÓRIA DE LONGO E CURTO PRAZO (LSTM) PARA PREDIÇÃO DE FLUXO DE RIO NA BACIA DO PANTANAL BRASILEIRO

RESUMO

Este artigo mostra uma aplicação bem-sucedida de rede neural recorrente - Long Short-Term Memory (LSTM), para simular a vazão na bacia do rio Aquidauana, dentro dos limites do Pantanal brasileiro. Os dados diários de precipitação serviram como variáveis de entrada para permitir que a rede LSTM previsse o fluxo futuro na região. Os resultados obtidos demonstram um coeficiente de determinação (R^2) de 0,82, indicando um ajuste favorável do modelo aos dados observados, juntamente com um erro quadrático médio (RMSE) de 0,53, demonstrando precisão na previsão do modelo em comparação com a vazão real. Tais métricas ressaltam a eficiência das redes LSTM para modelagem hidrológica na região do Pantanal, um aspecto crucial para o planejamento e gestão sustentável dos recursos hídricos na área. Espera-se que este estudo inspire novas pesquisas e contribua significativamente para o avanço das técnicas de previsão de vazões em bacias hidrográficas complexas e com deficiência de dados, como a bacia do Rio Aquidauana.

Palavras chave: LSTM, simulação de vazão, Pantanal.

1 INTRODUCTION

The Pantanal, located at the heart of South America, stands as the largest floodplain in the world. This region is renowned not only for its exceptional ecological diversity but also for its critical contribution to the regulation of regional hydrology (Assine et al., 2016; Couto & Oliveira, 2010; Macedo et al., 2014). Among the numerous rivers that weave through this vast wetland, the Aquidauana River plays a significant role in preserving the delicate equilibrium of the Pantanal ecosystem. It is an intricate web of rivers, marshes, and diverse habitats, providing a haven for countless species of wildlife. The comprehension and precise modeling of river flow within this region are of paramount significance for the management of water resources, flood control, and environmental conservation (Joia et al., 2018).

However, despite the ecological importance of the Pantanal, there exists a notable challenge - the scarcity of hydrological data. This shortage of data presents a formidable obstacle to the development of reliable river flow models. In particular, remote and less accessible areas of the Pantanal lack comprehensive hydrological observations. This scarcity impedes conventional modeling approaches, emphasizing the need for innovative techniques capable of capturing the intricate spatiotemporal dynamics of river flow. To address these limitations, this study embraces the capabilities of Long Short-Term Memory (LSTM) networks, a class of recurrent neural networks renowned for their proficiency in capturing sequential patterns and long-range dependencies (J. Fan et al., 2018; Kratzert et al., 2018; T. Liu et al., 2019).

The central aim of this investigation is to employ LSTM for modeling river flow within the Aquidauana River Basin, with precipitation data serving as a critical input variable. The LSTM model will be trained to learn the historical correlations between precipitation patterns and river flows, facilitating the generation of accurate flow predictions even during periods with limited available data. By utilizing LSTM, the study not only achieves precise river flow modeling but also offers a means to fill missing data gaps, a vital consideration given the scarcity of hydrological records in the Pantanal region.

This manuscript outlines the methodology employed for preprocessing and integrating the available hydrological and precipitation data. The process involves collecting and organizing historical data pertaining to river flow and precipitation within the Aquidauana River Basin. The data undergoes normalization, ensuring it is standardized on a common scale, which is essential for the LSTM model to function effectively. Subsequently, the LSTM architecture and training procedure are comprehensively detailed, providing insights into the neural network's inner workings and how it processes data sequences.

The research goes further by evaluating the LSTM model's proficiency in capturing the spatiotemporal fluctuations of river flow within the Aquidauana River Basin. This evaluation aims to demonstrate how well the LSTM network can adapt to the unique hydroclimatic conditions of the Pantanal region and provide accurate predictions. Additionally, a comparative analysis will be conducted, contrasting the LSTM's predictions with those generated by traditional hydrological models. This comparative assessment will determine the efficacy of LSTM in managing data limitations and capturing intricate hydrological dynamics.

In conclusion, this study represents a significant advancement in the realm of hydrological modeling within the Pantanal region. It showcases the capabilities of LSTM networks in effectively utilizing limited hydrological data for river flow prediction and data gap filling. The provision of precise river flow simulations through this research can significantly contribute to well-informed decision-making processes concerning sustainable water resource management, environmental preservation, and flood control initiatives in the Pantanal. These are critical aspects of maintaining the ecological balance and safeguarding the biodiversity of this unique region.

Moreover, the insights gained from the utilization of LSTM in addressing data scarcity challenges can offer valuable implications for analogous circumstances in other regions globally. The approach taken in this study can be a blueprint for addressing data limitations in hydrological modeling across various ecosystems, providing an innovative solution for regions where traditional data collection may be challenging. LSTM networks, with their ability to capture long-range dependencies, have the potential to revolutionize the field of hydrology, enabling more accurate predictions and better management of water resources in ecologically sensitive areas like the Pantanal. This research not only furthers our understanding of the Pantanal's hydrology but also offers a promising path towards sustainable management of water resources in other vulnerable regions worldwide. The contribution of LSTM networks to hydrological modeling is a step towards a more resilient and ecologically conscious approach to water resource management, with implications far beyond the Pantanal.

2 BIBLIOGRAFIC REVIEW

2.1 Hydrological Modeling in Wetland Ecosystems

The Pantanal, situated at the core of South America, is a remarkable example of a wetland ecosystem characterized by its dynamic hydrological processes. Wetlands, such as the Pantanal, are unique ecosystems that serve as vital ecological hubs and are often referred to as the Earth's "kidneys" due to their role in water purification and regulation (Couto & Oliveira, 2010). These regions are particularly susceptible to the impacts of climate change, as alterations in precipitation patterns and temperature can significantly influence the hydrology of wetlands (Cui et al., 2021; S. M. Mohammadizadeh, Filho, Descovi, Murillo-Bermúdez, & Sierra, 2023; Murillo Bermudez et al., 2023; Sierra et al., 2023). Consequently, accurate hydrological modeling is indispensable for understanding the response of wetland ecosystems to environmental changes, ensuring their conservation, and supporting the communities dependent on these regions for their livelihoods.

Wetlands are characterized by their fluctuating water levels, which are influenced by various factors, including precipitation, evapotranspiration, groundwater interactions, and the intricate network of rivers and channels. River flow within wetlands is not only critical for the ecological balance of these regions but also plays a crucial role in nutrient cycling and supporting diverse wildlife (Back et al., 2023; Xi et al., 2021). The Pantanal, with its vast floodplain and extensive network of rivers, exemplifies the intricate interplay between wetland hydrology and ecosystem health. Therefore, gaining insights into the hydrological dynamics of the Pantanal, and wetlands in general, is essential for sustainable environmental management, particularly in the face of changing climate conditions.

Furthermore, the challenges associated with hydrological modeling in wetland ecosystems are multifaceted. These regions often lack comprehensive hydrological data, especially in remote areas. Traditional modeling approaches, which rely heavily on historical observations and well-established statistical methods, may be insufficient to capture the complexity of wetland hydrology (Kratzert et al., 2018). The incorporation of innovative techniques, such as advanced machine learning algorithms like Long Short-Term Memory (LSTM) networks, becomes imperative to bridge the gaps in our understanding of wetland hydrology. LSTM networks, with their ability to capture sequential patterns and adapt to changing conditions, offer promising solutions for modeling the intricate and dynamic nature of river flow within wetland ecosystems.

2.2 Long Short-Term Memory (LSTM) Networks

In recent years, Long Short-Term Memory (LSTM) networks have emerged as a powerful tool in the realm of artificial intelligence and machine learning, particularly for modeling time-dependent data. LSTMs belong to the class of recurrent neural networks (RNNs) but possess a unique architecture that addresses one of the key challenges in sequential data analysis: the vanishing gradient problem (Hochreiter & Schmidhuber, 1997). This problem occurs when traditional RNNs struggle to capture long-term dependencies in sequential data, making them less suitable for tasks that require remembering information over extended time periods. LSTMs were designed to overcome this limitation, making them well-suited for applications in time series forecasting, natural language processing, and, notably, hydrology.

At the core of an LSTM network is the memory cell, which can store information over extended time steps and selectively retain or forget information as it processes sequential data. This memory cell is complemented by three gates: the input gate, the forget gate, and the output gate. These gates control the flow of information into and out of the memory cell, allowing the LSTM to capture and retain relevant patterns and context while discarding less important information (Mohammadizadeh et al., 2023; Yu et al., 2019). This architecture makes LSTMs highly adept at modeling sequences with varying time lags, making them particularly valuable for applications where past observations significantly influence future outcomes, such as river flow modeling.

In the context of hydrology, where time series data plays a central role, LSTMs offer several advantages over traditional statistical models (Fan et al., 2021; Gavidia, Mohammadizadeh, et al., 2023; Kratzert et al., 2019; Lees et al., 2021; Nikeghbali et al., 2014; Xiang et al., 2020). They can capture complex, non-linear relationships between input variables, such as precipitation patterns, and the target variable, such as river flow. Additionally, LSTMs can adapt to changing patterns and seasonality in the data, making them versatile for modeling hydrological processes influenced by climate, weather, and other factors. Their ability to handle irregular time intervals and missing data further enhances their utility in hydrological modeling, especially in regions with limited data availability, like the Pantanal (Assine et al., 2016).

As a result, the application of LSTM networks to hydrology has gained considerable attention in recent research (Descovi et al., 2023; Gavidia, Chinelatto, et al., 2023; Kratzert et al., 2018; Li et al., 2023; S. Mohammadizadeh et al., 2021; S. M. Mohammadizadeh, Filho, Descovi, Murillo-Bermúdez, & Sierra, 2023; Sahoo et al., 2019). Researchers have leveraged LSTM models to improve the accuracy of river flow forecasting, flood prediction, and the understanding of complex hydrological phenomena (Liu et al., 2020). The inherent ability of LSTMs to capture both short-term fluctuations and long-term dependencies in hydrological data aligns with the challenges posed by wetland ecosystems like the Pantanal, where river flow is influenced by multiple interacting factors and exhibits intricate spatiotemporal dynamics. Therefore, the integration of LSTM networks into the study of river flow within the Pantanal presents a promising avenue for enhancing our understanding of this vital ecosystem.

3 METHODOLOGY

3.1 Study Area and Data Collection

The study area of significant hydrological importance pertains to the Aquidauana River Basin, a region located within the vast Pantanal, an extensive tropical wetland nestled at the geographical heart of South America. The Pantanal primarily spans across the Brazilian states of Mato Grosso and Mato Grosso do Sul, extending into portions of Bolivia and Paraguay. Covering an estimated area ranging from approximately 150,000 to 195,000 square kilometers, the Pantanal stands as the largest freshwater wetland on a global scale. It assumes a pivotal role as an ecological hotspot and an invaluable natural resource of great significance.

The Aquidauana River, one of the numerous rivers coursing through the Pantanal, holds a key position in shaping the region's hydrological dynamics. Serving as a tributary of the Paraguay River, a major river in South America, the Aquidauana River substantially influences the overall flow and water regime of this intricate wetland. The river basin encompasses a diverse and intricate landscape, featuring wetlands, savannas, and tropical forests. The basin's topography, in conjunction with its tropical climate, gives rise to a complex hydrological system characterized by seasonal variations that include periods of flooding and drought.

The hydrographic basin of the Aquidauana River (as illustrated in Figure 1) constitutes a sub-basin of the Miranda River, itself one of the tributaries of the Paraguay River. Geographically, it is situated between the latitudes of 19° 19' 01'' and 21° 13' 49'' south, and the longitudes of 56° 49' 11'' and 54° 16' 44'' west. This basin is located in the north-central-western part of the state of Mato Grosso do Sul, stretching from the Maracaju Mountain Range, positioned within the municipality of São Gabriel do Oeste, to the expansive Pantanal plain, where it converges with the Miranda River within the municipal boundaries of Aquidauana.

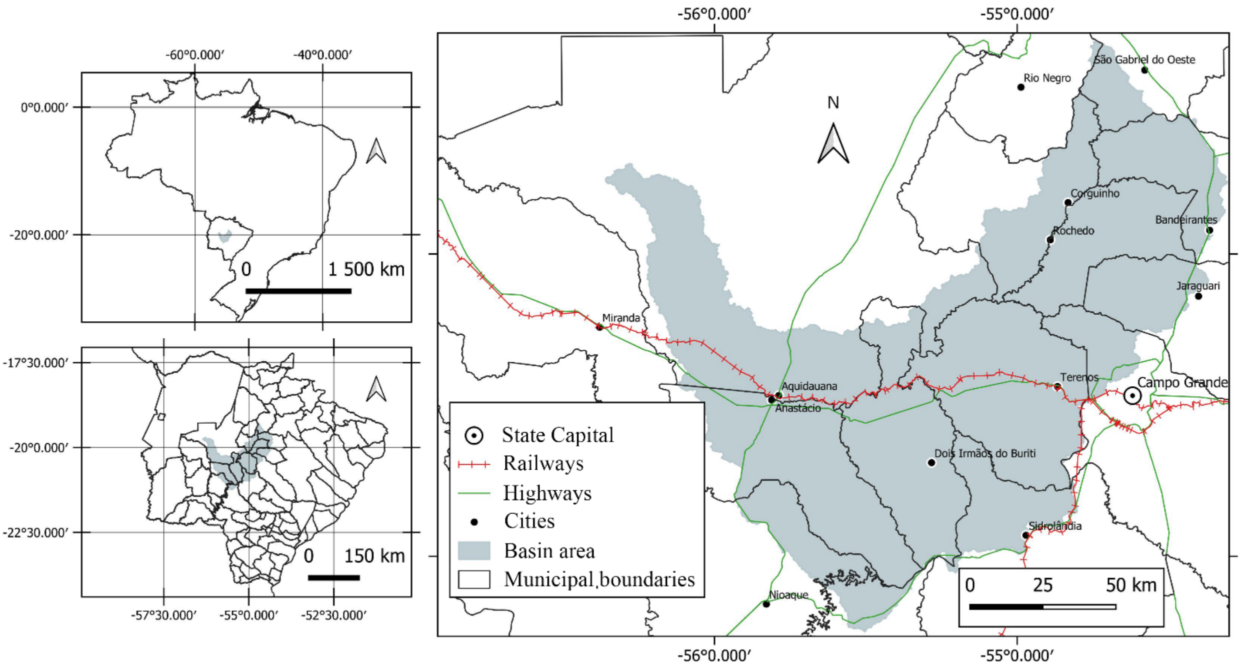


Figure 1: Location map of the Aquidauana River watershed.

The assessment of pluviometric stations encompassing the specified area was conducted using the database sourced from the NATIONAL WATER AGENCY (ANA, 2022), and the outcomes are outlined in Table 1. The survey identified a total of 6 pluviometric monitoring stations and 1 fluviometric monitoring station within the defined region.

Table 1 - Pluviometric Monitoring Stations within the Delimited Area.

	ID	Name	Location (MS)	Monitoring station
1	1956003	Entre Rios	Aquidauana	Precipitation
2	02055002	Palmeiras	Dois Irmãos do Buriti	Precipitation
3	02054009	Santa Elisa	Terenos	Precipitation
4	02054019	Jaraguari	Jaraguari	Precipitation
5	01954002	Rochedo	Rochedo	Precipitation
6	2155001	Nioaque	Nioaque	Precipitation
7	**66950000**	**Porto Ciriaco**	**Aquidauana**	**FLOW**

3.2 Long Short-Term Memory (LSTM) network

The LSTM network belongs to a distinct class of recurrent neural networks (RNNs), surpassing the constraints of conventional RNNs in effectively learning long-term dependencies. Initially proposed by Hochreiter and Schmidhuber (1997) and subsequently refined and popularized by Kawakami (2008), LSTM leverages its deep learning architecture to determine the timing for information retention and forgetting, accomplished through purposefully designed gates and memory cells (Figure 2).

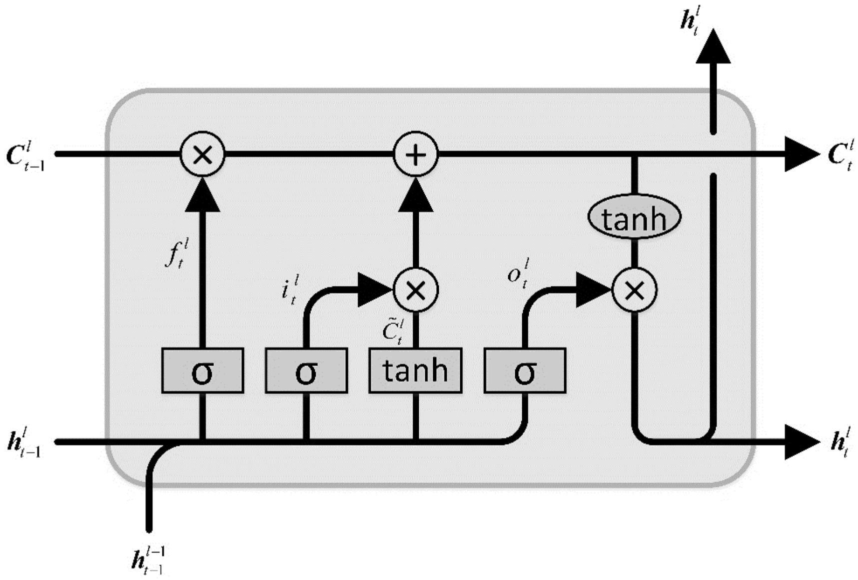


Figure 2 - The architecture of the LSTM cell. Source: FAN et al., (2018).

The key to LSTM is the cell state (C_t), which allows information to flow unchanged. The LSTM memory cell is regulated by three gates that optionally allow the passage of information. The first gate is called the forget gate, which controls which elements of the cell's previous state C_{t-1} will be forgotten.

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right) \quad (1)$$

Where f_t is an output vector from the sigmoid layer with values ranging from 0 to 1, indicating the degree of forgetting. W_f and b_f define the set of trainable parameters for the forget gate.

Next, the input gate decides which value will be updated:

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (2)$$

Where it is an output variable with a value ranging from 0 to 1. W_i and b_i are trainable parameters. Next, a candidate vector for the cell state is calculated using the current input (x_t) and the last hidden state (h_{t-1}):

$$\hat{C} = tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (3)$$

Where $\hat{C}$ is a vector with values ranging from 0 to 1, tanh is the hyperbolic tangent function, and W_c and b_c are trainable parameters. After that, you can update the old cell state C_{t-1} to the new cell state C_t by element-wise multiplication:

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \hat{C}_t \quad (4)$$

Finally, the output gate decides what will be the output through a sigmoid layer:

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (5)$$

In this way, o_t is a vector with values ranging from 0 to 1. W_o and b_o are trainable parameters defined for the output gate. The new hidden state h_t is then calculated by combining Equations 5 and 6:

$$h_t = o_t \cdot \tanh(C_t) \tag{6}$$

3.3 Adjustments to the size of the LSTM window

The quantity of preceding time steps is denoted as the window size, and this parameter exerts significant influence on prediction accuracy, necessitating careful selection to optimize the model's performance. In the present investigation, the window size is set to 5.

3.4 Adjustments to the hyperparameters

Neural networks conventionally encompass numerous hyperparameters, which are predefined prior to the commencement of the learning process. The optimization or tuning of these hyperparameters involves the search for a specific set of values that yield a model minimizing the loss function on the given data (GOODFELLOW; BENGIO; COURVILLE, 2016). In this study, the mean squared error (MSE) is employed as the loss function for hyperparameter optimization, following the work of Kratzert et al. (2018) and Fan et al. (2020).

Frequently encountered hyperparameters comprise the learning rate, the number of training epochs, the dimensionality of the output space, among others (GOODFELLOW; BENGIO; COURVILLE, 2016). The learning rate is a hyperparameter that signifies the step size in a gradient descent method (ZEILER, 2012). In this study, the Adam optimizer was employed as a stochastic optimization method (KINGMA; BA, 2017), with an initial learning rate set to 0.2. Additionally, a time-based decay rate was applied to update the learning rate during the training process.

Moreover, the number of epochs, typically denoting a complete pass through the entire dataset within the neural network, is employed to partition the training into discrete phases. Prolonged training can result in overfitting, wherein the model learns patterns exclusive to the training dataset (FAN et al., 2020). Conversely, insufficient training can lead to underfitting, signifying that the model fails to capture relevant patterns within the training data (GOODFELLOW; BENGIO; COURVILLE, 2016).

In this study, the number of training epochs will be set to 30, following the recommendation provided by Kratzert et al. (2018).

3.5 Validation of the LSTM model

The metrics used to evaluate the model's performance are the Coefficient of Determination (R^2) and the Root Mean Squared Error (RMSE). The calculation of R^2 is done using the equation:

$$R^2 = 1 - \frac{\sum(y_i - \hat{y})^2}{\sum(y_i - \acute{y})^2} \tag{7}$$

The calculation of RMSE (Root Mean Squared Error) is done using the following procedure:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y})^2}{n}} \quad (8)$$

Where:

- y_i is the observed flow at time *i*.
- $\hat{y}$ is the simulated flow at time *i*.
- $\acute{y}$ is the mean value of the observed flow data (average of y_i).
- n is the number of data points in the dataset.

4 RESULTS AND DISCUSSION

In this section, we present the outcomes of our study, which focused on harnessing the capabilities of the Long Short-Term Memory (LSTM) network to unravel the underlying patterns in historical time series data, particularly within the context of the Aquidauana River basin. Our methodology involved the utilization of daily precipitation data, denoted as y(t), as the input variable, and flow data, represented by x(t), as the target output variable. The temporal scope of our study encompassed the extensive period between January 1, 1999, and December 27, 2019, yielding a substantial dataset comprising 7,670 data samples.

4.1 Data Preprocessing

Normalization emerged as a pivotal preprocessing step aimed at standardizing the data on a common scale. The flow data underwent normalization, which involved rescaling it to a range between 0 and 1 or transforming it to have a mean of zero and a standard deviation of one. This normalization process played a pivotal role in stabilizing the training of our LSTM network, especially when dealing with variables that exhibit significant differences in magnitudes. Furthermore, we divided the dataset into training and validation subsets, adopting a 75/25 split ratio, where 75% of the normalized data was earmarked for training, and the remaining 25% for validation (Table 2 provides an overview of this partitioning).

Table 2 – Data division for training and testing.

Data	Samples	Percentage (%)
Training	5.752	75
Testing	1.918	25
Total	7.670	100

Following the data preprocessing steps, our LSTM network underwent the training process, which was assessed by monitoring the Mean Squared Error (MSE) for both the training and validation datasets (Figure 3). The plotted data indicated that the model converged to an approximately constant MSE value during training, with no significant increase as training epochs progressed. This consistent behavior for both training and validation data points toward a smooth training process without any notable anomalies. The stability and minimal MSE values strongly affirm the model's proficiency and its ability to effectively capture data patterns.

HOLOS
ISSN 1807 - 1600

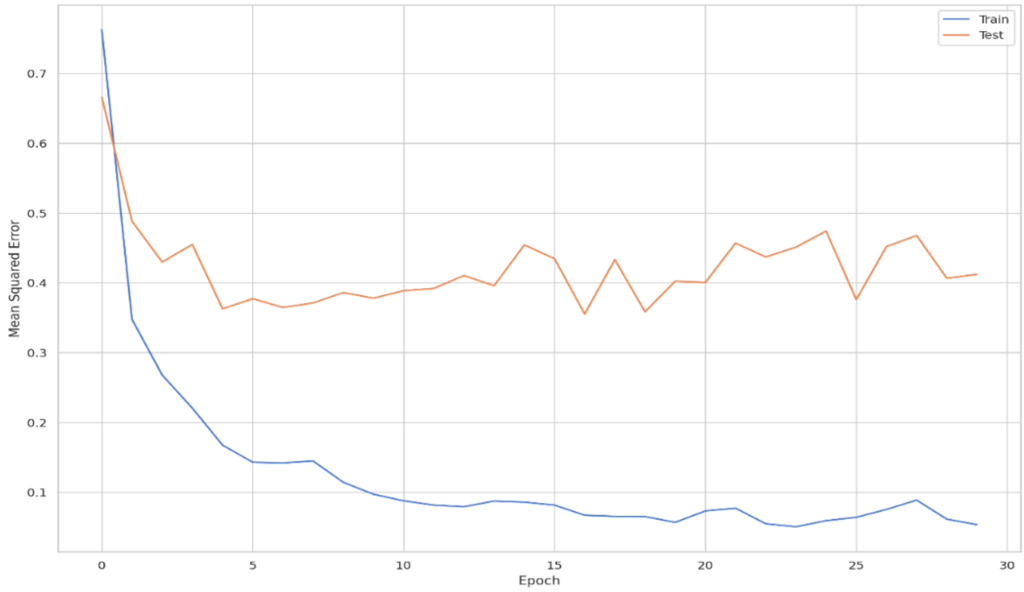
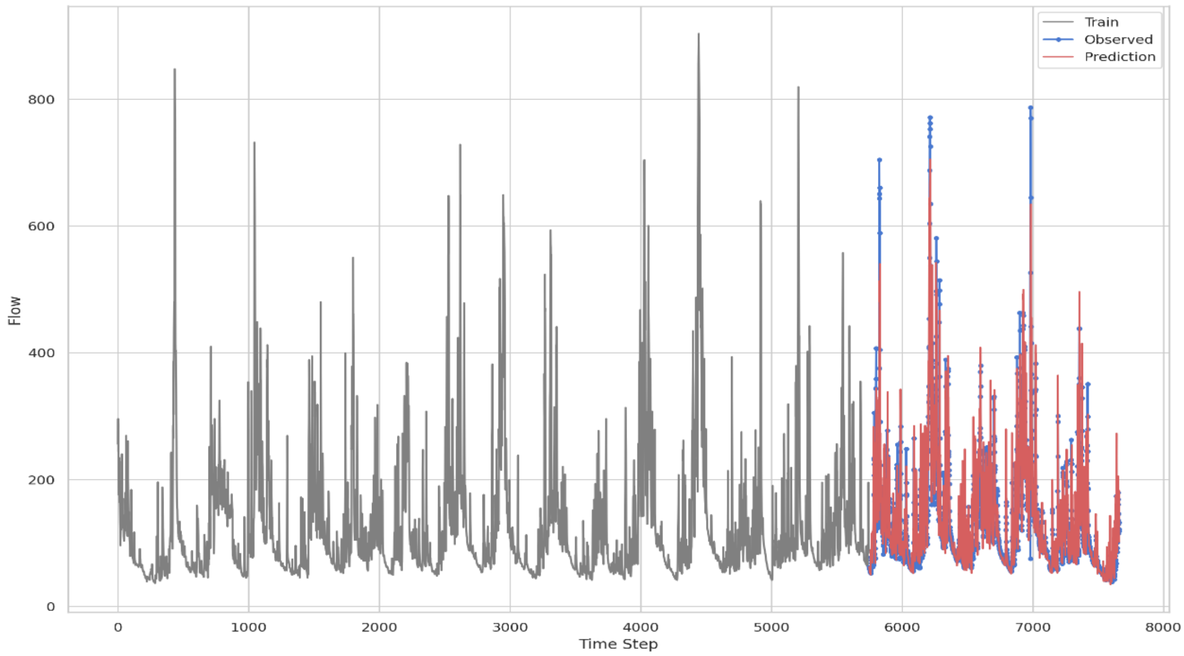


Figure 3 – Graph of the training epochs number.

4.2 Data Preprocessing

Upon completion of the training phase, we assessed the model's efficacy in flow prediction using the validation dataset. The validation process involved generating input-output sequences from the dataset, resulting in 1918 validation samples. The model's prediction mechanism relied on the sequences generated, producing output samples representing the target variable - in this case, the flow for the subsequent day. In this manner, the model was consistently supplied with the original data, with the primary objective of forecasting the target variable for the next day.

The simulation represented one-step-ahead predictions generated by the trained model. Input data sequences from pluviometric stations were fed into the model, enabling it to generate the subsequent flow sample, as depicted in Figure 4.



Figure

Figure 4 – Graph of the LSTM model results in the Aquidauana basin.

4.3 Model Performance

The performance of our LSTM model was rigorously evaluated using two key metrics: the coefficient of determination (R^2) and the Root Mean Squared Error (RMSE). R^2 is a commonly used metric for assessing the model's fit to observed data, while RMSE quantifies the accuracy of the model's predictions compared to actual data. These metrics capture the proportion of the total variance in the dependent variable (simulated flows) that can be attributed to the independent variables (precipitation data).

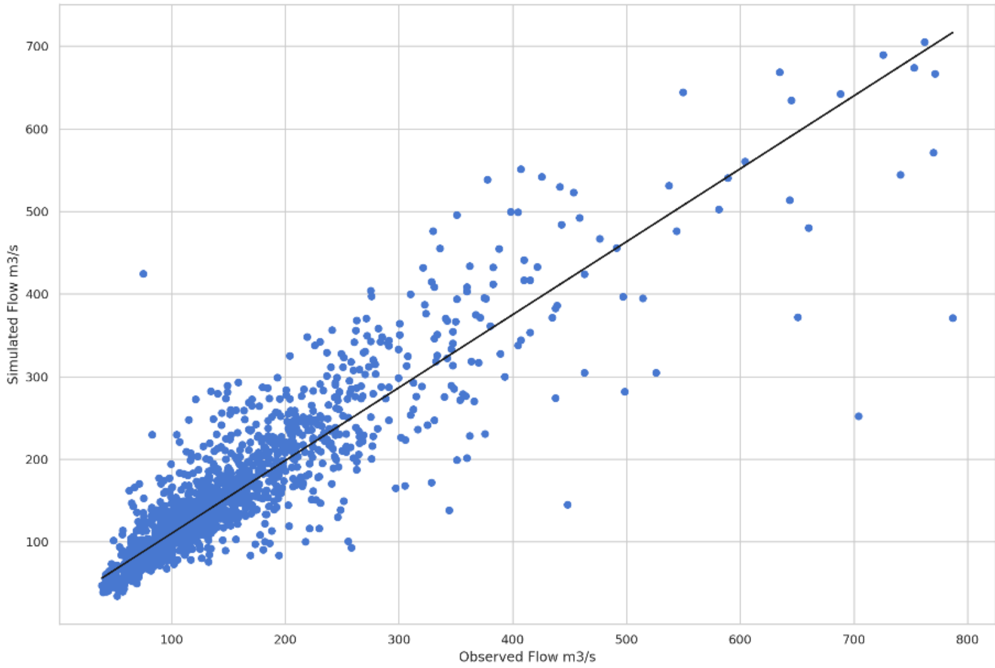


Figure 5 – R^2 value of 0,82.

The R^2 value of 0.82 signifies that approximately 82% of the variability observed in the simulated flows can be explained by the fluctuations in the precipitation data used as input for the LSTM model (Figure 5). A higher R^2 value indicates a more favorable fit of the model to the observed data, indicating a closer alignment between the model's predictions and the actual values. However, it is crucial to consider the specific application and context when interpreting the R^2 value accurately.

In conclusion, our results demonstrate the effectiveness of the LSTM network in modeling and predicting streamflow, offering a valuable tool for hydrological analysis in the Aquidauana River basin. This research not only contributes to the understanding of hydroclimatic relationships but also provides a robust framework for sustainable water resource management in complex watersheds. Further research could explore alternative neural network architectures and assess the model's performance under various climatic conditions to refine and extend these findings.

5 CONCLUSION

This study has successfully harnessed the power of the Long Short-Term Memory (LSTM) recurrent neural network to predict streamflow within the Aquidauana River basin, using daily precipitation data as its input variables. The impressive results we have obtained, exemplified by an R^2 value of 0.82 and an RMSE of 0.53, signify a remarkable alignment between our LSTM model and the observed streamflow data.

The key advantage of LSTM lies in its capability to capture long-term dependencies between precipitation and streamflow time series, enabling highly accurate predictions of future streamflows. The model's robust performance and its ability to discern and retain intricate temporal patterns highlight its suitability for a range of hydrological modeling tasks.

The practical significance of our approach is underscored by its critical role in providing precise streamflow predictions, a crucial component of sustainable water resources planning and management within complex watersheds. By integrating daily precipitation data with advanced machine learning techniques, our methodology emerges as a valuable asset in improving hydrological monitoring and decision-making, particularly in the face of intricate hydroclimatic conditions.

The practical significance of our approach is underscored by its critical role in providing precise streamflow predictions, a crucial component of sustainable water resources planning and management within complex watersheds. By integrating daily precipitation data with advanced machine learning techniques, our methodology emerges as a valuable asset in improving hydrological monitoring and decision-making, particularly in the face of intricate hydroclimatic conditions.In summary, this study constitutes a significant step forward in the application of LSTM networks for streamflow prediction using daily precipitation data. We hope that our findings will not only inspire further scholarly investigation but also encourage the practical adoption of this methodology to enhance water resource management across a wide spectrum of watersheds.

6 ACKNOWLEDGEMENTS

This work was supported by the Coordination for the Improvement of Higher Education Personnel (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES)) [grant numbers 88887.701529/2022-00, 88887.497812/2020-00, 88887. 701389/2022-00 and 88887.805051/2023-00].

7 REFERENCES

Assine, M. L., Merino, E. R., Pupim, F. N., Warren, L. V., Guerreiro, R. L., & McGlue, M. M. (2016). Geology and Geomorphology of the Pantanal Basin. Em I. Bergier & M. L. Assine (Orgs.), *Dynamics of the Pantanal Wetland in South America* (p. 23–50). Springer International Publishing. https://doi.org/10.1007/698_2015_349

Back, Á. J., Souza, G. da S., Galatto, S. L., Corseuil, C. W., & Poleto, C. (2023). EROSIVITY INDEX FOR BRASIL BASED ON CLIMATOLOGICAL NORMALS FROM 1991 TO 2020. *HOLOS*, *3*(39), Artigo 39. https://www2.ifrn.edu.br/ojs/index.php/HOLOS/article/view/16329

Couto, E., & Oliveira, V. (2010). The Soil Diversity of the Pantanal. Em *The Pantanal of Mato Grosso: Ecology, Biodiviersity and Sustainable Management of a Large Neotropical Seasonal Wetland* (p. 71–102).

Cui, Q., Ammar, M. E., Kariyeva, J., & Faramarzi, M. (2021). Regional wetland water storage changes: The influence of future climate on geographically isolated wetlands. *Ecological Indicators*, *120*, 106941. https://doi.org/10.1016/j.ecolind.2020.106941

Descovi, C. S., Zuffo, A. C., Mohammadizadeh, S. M., Murillo Bermudez, L. F., & Sierra, D. A. (2023, dezembro 29). *Modeling River Flows in the Brazilian Pantanal Basin Using Long Short-Term Memory (LSTM) Networks*. 7th SSS, Porto Alegre, Brazil.

Fan, D., Sun, H., Yao, J., Zhang, K., Yan, X., & Sun, Z. (2021). Well production forecasting based on ARIMA-LSTM model considering manual operations. *Energy*, *220*, 119708.

Gavidia, J. C. R., Chinelatto, G. F., Basso, M., Souza, J. P., Soltanmohammadi, R., Vidal, A. C., Goldstein, R. H., & Mohammadizadeh, S. M. (2023). Utilizing integrated artificial intelligence for characterizing mineralogy and facies in a pre-salt carbonate reservoir, Santos Basin, Brazil, using cores, wireline logs, and multi-mineral petrophysical evaluation. *Geoenergy Science and Engineering*, *231*, 212303. https://doi.org/10.1016/j.geoen.2023.212303

Gavidia, J. C. R., Mohammadizadeh, S. M., Chinelatto, G. F., Basso, M., Souza, J. P. da P., Vidal, A. C., & Goldstein, R. H. (2023). Bridging the Gap: Integrating Static and Dynamic Data for Improved Permeability Modeling and Super K Zone Detection in Vuggy Reservoirs. *Journal of Geoenergy Science and Engineering*. In Press.

Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation*, *9*(8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

Kratzert, F., Klotz, D., Brenner, C., Schulz, K., & Herrnegger, M. (2018). Rainfall–runoff modelling using Long Short-Term Memory (LSTM) networks. *Hydrology and Earth System Sciences*, *22*(11), 6005–6022. https://doi.org/10.5194/hess-22-6005-2018

Kratzert, F., Klotz, D., Shalev, G., Klambauer, G., Hochreiter, S., & Nearing, G. (2019). Benchmarking a catchment-aware long short-term memory network (LSTM) for large-scale hydrological modeling. *Hydrol. Earth Syst. Sci. Discuss*, *2019*, 1–32.

Lees, T., Reece, S., Kratzert, F., Klotz, D., Gauch, M., De Bruijn, J., Kumar Sahu, R., Greve, P., Slater, L., & Dadson, S. (2021). Hydrological concept formation inside long short-term memory (LSTM) networks. *Hydrology and Earth System Sciences Discussions*, *2021*, 1–37.

Li, B., Li, R., Sun, T., Gong, A., Tian, F., Khan, M. Y. A., & Ni, G. (2023). Improving LSTM hydrological modeling with spatiotemporal deep learning and multi-task learning: A case study of three mountainous areas on the Tibetan Plateau. *Journal of Hydrology*, *620*, 129401.

Liu, D., Jiang, W., Mu, L., & Wang, S. (2020). Streamflow Prediction Using Deep Learning Neural Network: Case Study of Yangtze River. *IEEE Access*, *8*, 90069–90086. https://doi.org/10.1109/ACCESS.2020.2993874

Mohammadizadeh, S. M., Filho, J. G. D., Descovi, C. S., Murillo-Bermúdez, L. F., Sierra, D. A., & Perez, G. A. C. P. (2023, dezembro 29). *Cavitation Jet Apparatus for the Evaluation of Cavitation Erosion of Solid Surfaces*. 7th SSS, Porto Alegre, Brazil,.

MOHAMMADIZADEH, S., Dalfré Filho, J. G., Sampaio Descovi, C., Murillo Bermúdez, L. F., Alfonso Sierra, D., & Corzo Perez, G. A. (2023). Assessment of cavitation erosion on solid surfaces using a cavitation jet apparatus. HOLOS, 5(39). https://doi.org/10.15628/holos.2023.16306

Mohammadizadeh, S., Moghaddam, M. A., & Talebbeydokhti, N. (2021). Analysis of flow in porous media using combined pressurized-free surface network. *Journal of Porous Media*, *24*(10). http://dx.doi.org/10.1615/JPorMedia.2021025407

Murillo Bermudez, L. F., Abreu, A. E. S. de, Pereira, S. Y., Mohammadizadeh, S. M., Descovi, C. S., Sierra, D. A., & Martim, A. L. S. S. (2023, dezembro 29). *CONTAMINAÇÃO DAS ÁGUAS SUBTERRÂNEAS POR ESGOTO SANITÁRIO: COMO AVALIAR A FONTE DA CONTAMINAÇÃO?* 7th SSS, Porto Alegre, Brazil.

Nikeghbali, P., Omidvar, P., Akbari, G., & Mohammadizadeh, S. M. (2014). THE STUDY ON SEDIMENT MOTION AND THE MODEL OF DAM BREAK ON THE MOVABLE BED BY SPH METHOD. *THE STUDY ON SEDIMENT MOTION AND THE MODEL OF DAM BREAK ON THE MOVABLE BED BY SPH METHOD*. INTERNATIONAL CONFERENCE ON COASTS, PORTS AND MARINE STRUCTURES (ICOPMAS). https://sid.ir/paper/948560/en

Sahoo, B. B., Jha, R., Singh, A., & Kumar, D. (2019). Long short-term memory (LSTM) recurrent neural network for low-flow hydrological time series forecasting. *Acta Geophysica*, *67*(5), 1471–1481.

Sierra, D. A., Luvizotto Jr, E., Mohammadizadeh, S. M., Descovi, C. S., Murillo Bermudez, L. F., & Filho, J. G. D. (2023, dezembro 29). *TRANSITÓRIOS HIDRÁULICOS EM LINHAS DE RECALQUE DE ESGOTO - REVISÃO DA LITERATURA*. 7th SSS, Porto Alegre, Brazil.

Xi, Y., Peng, S., Ciais, P., & Chen, Y. (2021). Future impacts of climate change on inland Ramsar wetlands. *Nature Climate Change*, *11*(1), 45–51.

Xiang, Z., Yan, J., & Demir, I. (2020). A rainfall-runoff model with LSTM-based sequence-to-sequence learning. *Water resources research*, *56*(1), e2019WR025326.

Yu, Y., Si, X., Hu, C., & Zhang, J. (2019). A review of recurrent neural networks: LSTM cells and network architectures. *Neural computation*, *31*(7), 1235–1270.

ABOUT THE AUTHORS

C. S. DESCOVI
Master in Civil Engineering Teaching, in the area of water, energy and environmental resources, from the State University of Campinas (Unicamp) and is currently pursuing a PhD at the same faculty. Research areas: hydrological modeling, hydrology, hydraulics.
E-mail: cassianodescovi@gmail.com
ORCID ID: https://orcid.org/0000-0002-2847-2064

A.C. ZUFFO
Master's degree in Civil Engineering - Concentration area in Hydraulic Engineering - University of São Paulo (EPUSP) and PhD in Hydraulic Engineering and Sanitation from EESC-USP. Professor at FEC-Unicamp. Post Doctorate from the University of Toronto - Ontario Canada.
E-mail: zuffoac@unicamp.br
ORCID ID: https://orcid.org/0000-0002-2186-9755

S. MOHAMMADIZADEH
Ph.D. candidate at School of Civil Engineering, Architecture and Urbanism (FECFAU), University of Campinas (UNICAMP), Brazil. He received a MS.c (2013) from University of Sistan and Baluchestan (USB), Iran, in the field of Civil Engineering – Hydraulic Structures. His main research interests are Cavitation, Porous Media, Computational Fluid Dynamics (CFD), Ultra High Performance Fibre Reinforced Concrete (UHPFRC), Petroleum Geology, as well as Hydrology.
E-mail: mohammadizadeh@fec.unicamp.br
ORCID ID: https://orcid.org/0000-0001-6995-066X

L. F. M. BERMÚDEZ
MSc and PhD student in Civil Engineering Education, area of water, energy and environmental resources, from the Campinas State University (Unicamp).
E-mail: luismurillo@fec.unicamp.br
ORCID ID: https://orcid.org/0000-0001-7564-487X

D. A. SIERRA
MSc in Civil Engineering Education, area of water, energy and environmental resources, from the Campinas State University (Unicamp) and PhD student at the same faculty. Research areas: hydraulic transients, bio-inspired algorithms, transient mitigation devices.
E-mail: d192229@dac.unicamp.br
ORCID ID:https://orcid.org/0000-0003-0699-4512

Editor: Francinaide de Lima Silva Nascimento
***Ad Hoc* Reviewer:** Fernando Fan e Fernando Neves Lima

Submitted June 5, 2023
Accepted September 11, 2023
Published December 18, 2023