Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Predicting Flood Water Levels in Data-scarce Meso-Scale River Basins for Support Early Warning Systems

Vinicius Santanna Castiglio

viniciuscastiglio@Hotmail.com

Federal University of Rio Grande do Sul: Universidade Federal do Rio Grande do Sul
https://orcid.org/0000-0003-0460-5966

Gean Paulo Michel

Federal University of Rio Grande do Sul: Universidade Federal do Rio Grande do Sul

Research Article

Keywords: Artificial Neural Network, Ungauged Catchments, Flood forecasting, Water level prediction, Machine learning

Posted Date: April 8th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9045388/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Vinícius Santanna Castiglio[1] (Corresponding author)

Email: viniciuscastiglio@hotmail.com

Gean Paulo Michel[1]

Email: geanpmichel@gmail.com

PREDICTING FLOOD WATER LEVELS IN DATA-SCARCE MESO-SCALE RIVER BASINS FOR SUPPORT EARLY WARNING SYSTEMS

Abstract

This study proposes a methodology for predicting maximum river water levels during rainfall events in meso-scale river basins with limited data availability. The approach reformulates flood peak prediction as a multi-horizon conditional mapping problem, thereby reducing the dependence on long historical time series and recurrent architectures commonly employed in the literature. The proposed methodology estimates flood levels based on prior observations of river stage, precipitation data, and the elapsed time since the onset of the event, aiming to support early flood warning systems. For the proposed approach, water level and accumulated precipitation data were used for forecast horizons of 1, 2, 6, 12, and 14 hours prior to the occurrence of the predicted level. The methodology is based on an artificial neural network model, with the dataset divided into training (70.0%) and validation (30.0%) subsets. The model was applied to the Mascarada River basin, located in northeastern Rio Grande do Sul State, Brazil. The results demonstrated strong predictive performance, particularly for forecast horizons equal to or shorter than 6 hours, for both training and validation datasets. Nash–Sutcliffe efficiency values exceeded 0.75, while RSR values remained below 0.50. For forecast horizons of 12 hours or longer, performance ranged from good to satisfactory. Overall, the proposed methodology and forecasting model demonstrated high potential for water level prediction in data-scarce river basins, representing a promising tool for the anticipation of extreme hydrological events under such conditions.

Keywords: Artificial Neural Network; Ungauged Catchments; Flood forecasting; Water level prediction; Machine learning

[1] Federal University of Rio Grande do Sul (UFRGS), Graduate Program in Water Resources and Environmental Sanitation, Porto Alegre, Brazil

1. Introduction

The increasing frequency and severity of extreme hydrological events, particularly floods, have driven the need for more advanced forecasting and early warning systems to reduce risks and impacts on vulnerable populations (PAINTER *et al.*, 2025; ZAHRI *et al.*, 2025). This issue is especially critical in the state of Rio Grande do Sul, which was recently affected by one of the most catastrophic hydrological events in its history (COLLISCHONN *et al.*, 2025; ROBAINA *et al.*, 2024). The prediction of hydrological events, particularly in Brazil, is often challenging, mainly due to limited data availability, insufficient investment in hydrometeorological monitoring, and discontinuous policies for hydrological data prevention and monitoring (RUCKERT; VICENTE; GOMES, 2024).

In smaller river basins and in regions of low economic relevance, operational and financial constraints hinder hydrological monitoring, resulting in a low density of hydrometeorological stations. Headwater catchments, with small to medium contributing areas, are characterized by abrupt changes in river water level over time scales ranging from a few minutes to several hours, depending on regional conditions (GOURLEY *et al.*, 2017). The forecasting of such events is therefore often limited and provides a short response time for preparedness and population evacuation (SARCHANI; TSANIS, 2019).

Hydrological modeling of events at the meso-scale can become costly and complex, particularly when traditional forecasting approaches are applied, as these methods require more robust input datasets and a detailed characterization of river basins (KEDAM *et al.*, 2024; LANGE; SIPPEL, 2020; LI *et al.*, 2023). Moreover, data scarcity in many small- and medium-sized basins represents a substantial challenge for the application of traditional hydrological models, whether empirical or physically based, since data limitations compromise calibration procedures and the generalization capability of these models (RUIZ-VILLANUEVA *et al.*, 2013; STAUDINGER *et al.*, 2025; XIMIN *et al.*, 2025). Data scarcity may be further exacerbated by the lack of detailed geomorphological information, such as high-resolution contour lines, elevation profiles, and slope parameters, as well as by the absence of information on land use and land cover, vegetation type and density, baseflow characteristics, and high-resolution climatic data. Some of this information constitutes initial conditions, which seriously compromises model calibration and validation processes (HUANG; BARDOSSY, 2020; SHADMEHRI TOOSI *et al.*, 2025).

In scenarios characterized by scarce or low-quality hydrological data, the uncertainties associated with forecasts limit the applicability of traditional models, making machine learning approaches a robust alternative to improve predictive accuracy and support monitoring systems, early warning, and disaster risk management (HASAN *et al.*, 2024; MOSAVI; OZTURK; CHAU, 2018). These techniques exhibit a greater capacity to adapt to data-scarce environments, provided that reliable time series of observed variables, such as precipitation and river water level, are available. In this context, within the field of water resources, applications that demand high computational cost stand out, including computational hydrological modeling (KEDAM *et al.*, 2024; MARCOS JUNIOR *et al.*, 2024; PATHAN *et al.*, 2024) and water distribution network optimization (EVANGELISTA *et al.*, 2024).

Recent studies predominantly employ machine learning models based on Recurrent Neural Networks, particularly Long Short-Term Memory (LSTM) architectures, which have demonstrated excellent performance in rainfall–stage (JHONG *et al.*, 2024), rainfall–peak discharge (HU *et al.*, 2018),and rainfall–runoff modeling (MALIK *et al.*, 2024). Despite the widespread adoption of these architectures, recent evidence indicates that non-recurrent models, such as Multilayer Perceptron (MLP) neural networks, can achieve performance comparable to LSTM networks in hydrological simulations, including ungauged basins, when properly parameterized (SHIN; JUNG, 2025). Moreover, recurrent approaches often require large volumes of data, as demonstrated by (PENG *et al.*, 2022), who reported significant performance degradation of LSTM networks under constraints in training data availability. In this context, recent studies have shown that feedforward architectures can yield satisfactory predictive performance even in the absence of explicit sequential modeling, suggesting that recurrence is not always indispensable in certain hydrological formulations (KEDAM *et al.*, 2024; SHIN; JUNG, 2025).

Nevertheless, the application of these approaches in meso-scale basins with limited data availability or discontinuous records remains insufficiently explored, particularly under data-scarce conditions in which deep learning models face additional challenges related to training and generalization (HASAN *et al.*, 2024; MOSAVI; OZTURK; CHAU, 2018). In this context, the present study reformulates flood level prediction as a conditional mapping between the observed initial water level, accumulated precipitation, and the forecast horizon, proposing a methodological framework particularly suited for basins with limited monitoring records.

Accordingly, with the objective of enabling hydrological forecasting in data-scarce regions and supporting the development of early warning systems, a peak water level prediction

model based on machine learning techniques was developed using Artificial Neural Networks (ANN). In this framework, direct prediction of water level plays a central role in early warning applications, as stage thresholds are sufficient to support response actions and impact mitigation without requiring precise discharge estimates. In river basins with limited data availability and unreliable rating curves, adopting water level as the primary variable reduces additional uncertainties and facilitates the implementation of simpler and more operationally feasible monitoring and warning systems. The proposed methodology was applied to the Mascarada River basin, located in the northeastern region of Rio Grande do Sul state, Brazil, which presents the data limitations and characteristics discussed above.

2. Methodology of the Artificial Neural Network–Based Peak Water Level Forecasting Model

2.1. Model Forecasting Architecture

The model aims to predict maximum water levels, or levels close to peak conditions, at a river cross section based on the previously observed water level at the same location, the cumulative precipitation up to the forecast horizon, and the lead time (Δt). The methodology is specifically designed to estimate water levels associated with flood conditions that may potentially trigger inundations, with particular emphasis on levels corresponding to peak stages. The direct use of the observed water level as the main model variable was adopted instead of discharge in order to avoid the uncertainties associated with stage–discharge relationships, especially in contexts characterized by discontinuous or poorly consistent time series.

The model framework is based on the identification of maximum water levels throughout the historical record, which are used exclusively as references for flood event definition and for dataset construction. From each identified peak, water levels observed at previous time steps are extracted and used as the initial water levels forming the model inputs. Given the issue of data scarcity, and in order to increase the number of samples available for model training and validation, the same procedure is also applied to water levels preceding the peak level. Thus, in addition to the peak level itself, water levels recorded at earlier times (e.g., from Δt + 1 hour to Δt + n hours) are considered, as illustrated in Figure 1 (a). For each of these reference levels, the initial water level is defined using the same lead time interval, with successive temporal increments relative to the initial level N_0.

As previously noted, the model was structured by considering the water level observed at a time prior to the event peak, referred to as the initial water level. The lead time Δt between the initial water level and the peak water level should be defined so as to ensure that the initial state represents, whenever possible, predominantly baseflow conditions, that is, prior to the onset of the rising limb of the hydrograph. This approach seeks to characterize conditions with minimal influence from runoff generated by precipitation occurring immediately before the initial water level observation.

For meso-scale catchments, it is suggested that this interval be on the order of three to four times the estimated time of concentration of the basin, which increases the likelihood that the initial water level is not influenced by runoff from previous rainfall events.

Based on the initial water levels defined by the lead time, preferably close to baseflow levels, the predicted water levels are generated for each of the previously defined forecast lead times, as illustrated in Figure 1 (b). Thus, each initial water level is associated with a set of forecast horizons ranging from time T_0 to T_n, with the same number of forecast horizons maintained for all initial water levels considered.

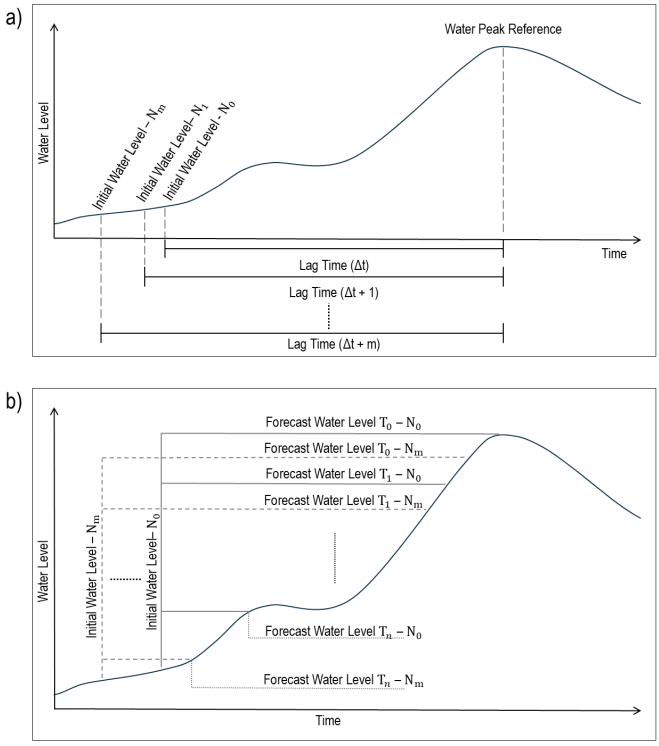


Figure 1. Definition of initial water levels for the construction of the training dataset; (b) Prediction of water levels based on the initial water level. Source: The authors.

In the context of the input variables of the forecasting model, three inputs were considered: the initial water level $N(T_0)$, the accumulated precipitation $PA(h)$ and the lead time

h associated with the water-level forecasting horizon. For each combination of these variables, the model provides as output the predicted water level at the future time $T_0 + h$, denoted as $NP(T_0 + h)$.

It is emphasized that, for a given simulation, the initial water level is kept constant throughout the dataset, while the accumulated precipitation and the forecasting horizon vary. In this way, a single initial water level gives rise to multiple future water-level estimates corresponding to the different forecasting horizons considered. Figure 2 illustrates the structure of the model, highlighting the organization of the input variables and the generation of the outputs associated with each forecasting horizon, considering, in this context, a single reference initial water level.

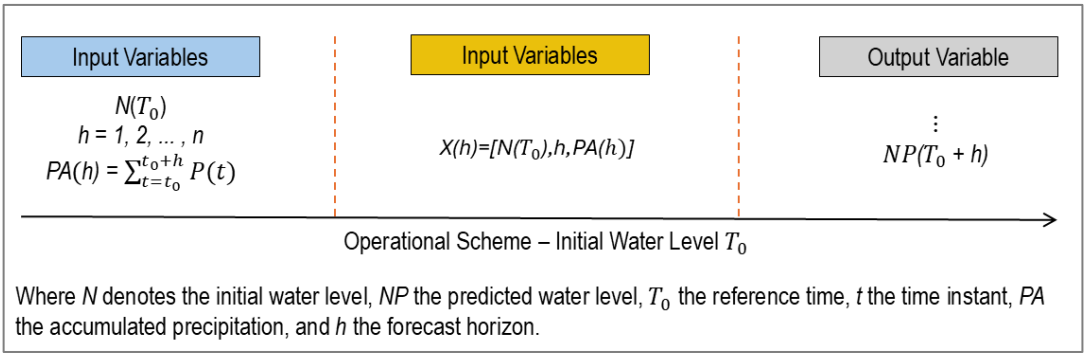


Figure 2. Operational framework of the proposed water level forecasting model. Source: The authors.

In summary, the proposed forecasting model adopts a multi-horizon strategy based on fixed initial water levels, from which independent predictions are generated for different forecast lead times. Each prediction is processed in a non-recurrent manner, using as input variables the initial water level, the forecast lead time, and the corresponding accumulated precipitation, without any feedback mechanism between the considered horizons.

Finally, it is important to emphasize that the model was structured to operate as a component of a real-time hydrological early warning system, in which water level observations are continuously updated and combined with accumulated precipitation and the selected forecast horizon to produce advance stage estimates. Therefore, the proposed methodology is not limited to retrospective modeling but is explicitly designed for operational application, enabling direct decision support under imminent risk scenarios.

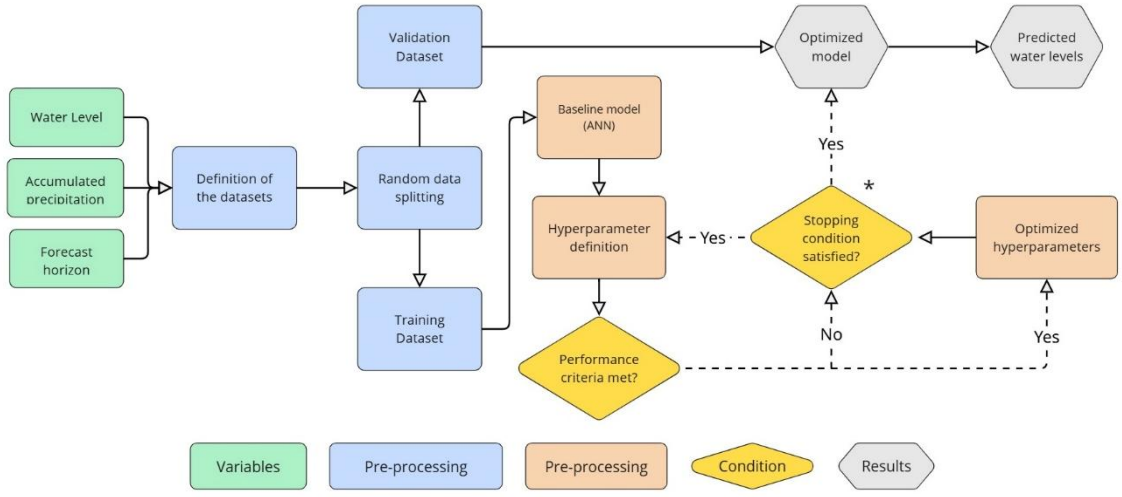
2.2. Machine Learning Model Parameterization

The model was structured with three input variables and one target variable. The network architecture includes a single output neuron in the final layer, corresponding to the

variable of interest predicted by the system. The final neural network was designed with one input layer, three hidden layers, and one output layer. The hidden layers employ the Rectified Linear Unit (ReLU) activation function, a linear activation model widely used for its effectiveness in mitigating the vanishing gradient problem and promoting faster learning (IJOMAH *et al.*, 2020).

The adoption of a feedforward architecture (Multilayer Perceptron) also contributes to the computational simplicity of the model, enabling fast training and low processing costs. This characteristic facilitates its implementation in small-scale monitoring systems or in environments with limited computational infrastructure, thereby broadening its applicability in regions facing technical and financial constraints.

To improve predictive performance and identify the optimal combination of parameters that maximizes model accuracy, the TensorFlow framework (KEETON, 2016) was used in conjunction with the Hyperband algorithm (LI *et al.*, 2018) for model optimization, based on a bandit-based hyperparameter optimization strategy. Unlike manual calibration approaches, Hyperband evaluates multiple hyperparameter combinations and iteratively adjusts the model according to validation performance. Moreover, the search is conducted efficiently by dynamically allocating computational resources based on real-time performance, enabling a broad and effective exploration of the hyperparameter space. Figure 3 presents the input–output flowchart of the model.



* A maximum number of epochs was predefined; however, training could be terminated early based on an early stopping criterion when no improvement in the validation metrics was observed over a predefined number of consecutive epochs.

Figure 3. Flowchart of the forecasting model. Source: The authors.

Model optimization was carried out by varying the number of units in the hidden layers, with values ranging from 16 to 128, using the Hyperband algorithm. This approach is based on

an adaptive search strategy with dynamic resource allocation, enabling the efficient evaluation of multiple configurations by prioritizing those that show better performance in the early stages of training. The learning rate was varied between $1 * 10^{-2}$ and $1 * 10^{-4}$ and a dynamic variation in the number of epochs was adopted, with a maximum of 2000 epochs. Training monitoring was performed through the evaluation of the loss function, based on the Mean Absolute Error (MAE).

Finally, in order to prevent overfitting—understood as the excessive adjustment of the model to the training data (SRIVASTAVA *et al.*, 2014) —an early stopping strategy was adopted, which is widely used to improve the generalization capability of neural networks (GOODFELLOW; BENGIO; COURVILLE, 2016). In this context, training is interrupted if no improvement in the loss function is observed after 50 consecutive epochs.

To assess the model's generalization capability, cross-validation was employed to evaluate its performance. The validation was based on the k-fold cross-validation method, in which the dataset is divided into *k* subsets (folds). The model is trained *k* times, each time using *k–1* folds for training and the remaining fold for validation, alternating the validation set at each iteration.

The use of this procedure provides a more reliable estimate of model performance and helps to assess potential overfitting. Cross-validation also enables a more robust comparison among different models or hyperparameter configurations. Consequently, this approach ensures that the selected model achieves an appropriate balance between bias and variance, making its predictions more reliable when applied to unseen data.

2.3. Case Study – Mascarada River Basin

The proposed methodology was applied to the Mascarada River basin (Figure 4), located in the northeastern portion of the state of Rio Grande do Sul, Brazil, and encompassing the municipalities of Rolante, Riozinho, and São Francisco de Paula. The basin has an approximate drainage area of 312 km², representing the largest portion of the Rolante River basin and constituting its main tributary. The Rolante River, and consequently the Mascarada River, are tributaries of the Sinos River, one of the most important river basins in the state of Rio Grande do Sul, which supplies a large portion of the Porto Alegre metropolitan region.

The Mascarada River basin has an approximate channel length of 32.4 km, with headwater elevations reaching about 981 m and the lowest portion of the basin at approximately

70 m. As a result, the basin presents an average slope of 2.8%. The basin time of concentration, calculated using the Kirpich equation (GREGORY, 1982), was estimated to be approximately 4.0 hours.

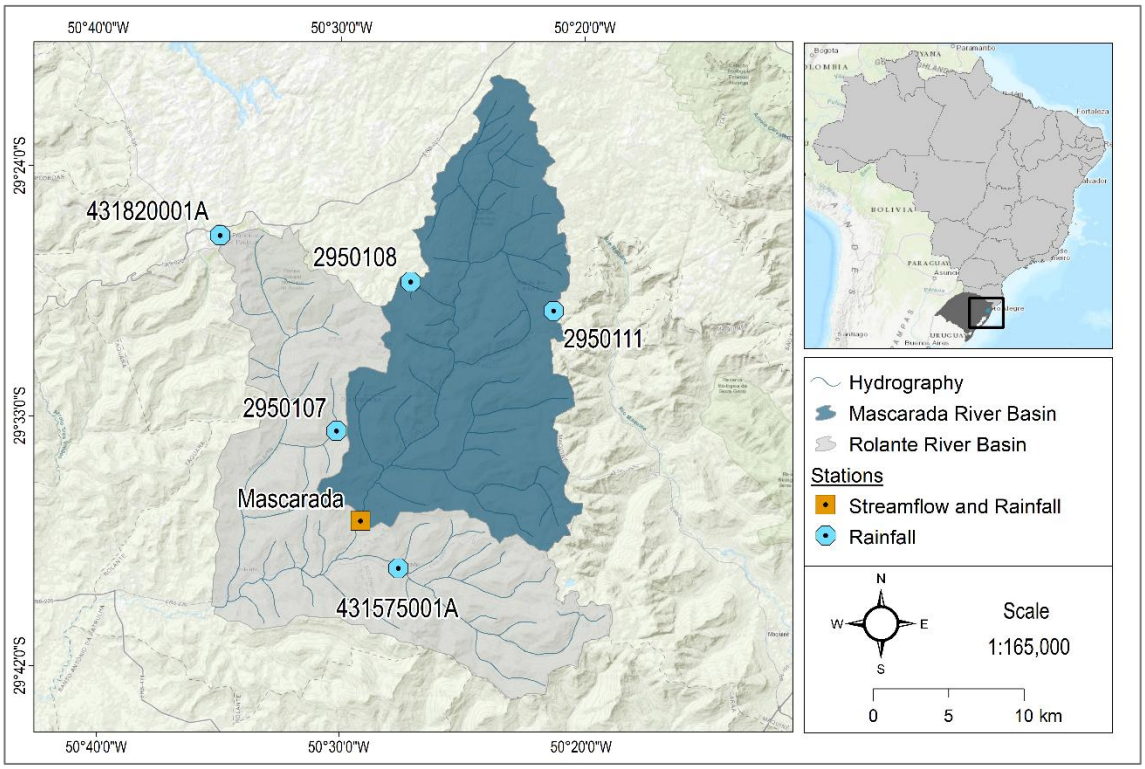


Figure 4. Automated hydrometeorological stations in the Rolante and Mascarada river basins. The authors.

The basin is classified as having a humid subtropical climate (Cfa), with mean annual temperatures ranging between 0 °C and 22 °C and no well-defined dry season. Furthermore, due to its geomorphological characteristics, the Rolante River basin may experience orographic rainfall, particularly in its headwater region Guirro (2020).

The Mascarada River basin is equipped with several hydrometeorological stations, predominantly rainfall gauges. However, most of these stations present data gaps or discontinuous periods of operation, which compromises the availability of long and consistent historical time series for the application of traditional hydrological models. In addition, the region has only one *Water Level* gauging station, which is currently inactive and whose data records are also discontinuous, with less than two years of operation. In this context, the adoption of *Water Level* as the modeling variable proved to be particularly appropriate, given the limitations associated with the consistency of the rating curves available for the Mascarada station.

Regarding the precipitation data used in the model, the average precipitation derived from at least three rain gauges available within the basin was adopted. Although the basin is equipped with more than three rain gauges, a prioritization criterion was established for station selection, assigning greater weight to precipitation data from the headwater regions of the basin. In the absence of data from the priority stations, information from the remaining stations located within and around the basin was used. Table 1 summarizes the operational periods of each station, as well as their respective priority levels for use in the analysis.

Table 1. Hydrometeorological station data available in the Rolante River Basin. Source: The Authors.

ID	Station Name	Station Type	Operator	Priority Order	Status	Start of Operation	End of Operation
2950108	Taipas Farm	Rain Gauge	SEMA	1	Active	12/31/2018	-
2950111	Rincão dos Kroeff	Rain Gauge	SEMA	2	Inactive	12/31/2018	01/31/2022
431820001A	Sao Francisco de Paula	Rain Gauge	CEMADEM	3	Active	01/01/2014	-
2950107	Boa Esperança Rolante	Rain Gauge	SEMA	4	Inactive	12/31/2018	31/03/2023
Mascarada	Mascarada	Rainfall/Streamflow Station	GPDEN	5	Inactive	02/04/2018	11/25/2020
431575001A	Riozinho	Rain Gauge	CEMADEM	6	Inactive	04/01/2018	02/01/2019

The analysis of the historical time series from the Mascarada streamflow station allowed the identification of hydrological events characterized by maximum water levels exceeding 1.0 m. These events were used as a reference for extracting the peak water levels employed in the development of the forecasting model for the Mascarada River basin. Based on the characterization of the basin baseflow conditions, a lead time of 14 h relative to the peak water level was adopted as a criterion. The initial water level used for model training was defined according to this temporal interval.

From the initial water level N_0 forecasting horizons of 1, 2, 6, 12, and 14 hours after the initial level were established, with the 14-hour horizon corresponding to the event peak water level, considering the reference initial level N_0. In order to increase the number of records available for model training, additional initial water levels (N_1, N_2, and N_3) were defined, corresponding to lead times of T_0 + 1h, T_0 + 2h and T_0 + 3h, respectively. The observed peak water levels ranged from 1.04 m to 3.26 m, totaling 38 event records. From these peak records, 750 individual water level samples were generated and subsequently used in the model training and validation processes.

Considering that the initial water level is fixed for each dataset—i.e., the same initial level is used to process forecasts with lead times of 1, 2, 6, 12, and 14 hours—datasets composed of five water level records were defined. This grouping strategy was applied during the randomization process of the model input data in order to ensure that, in both the training and validation phases, the same number of samples was available for each forecasting horizon. This procedure also prevents information leakage from a given initial water level between the training and validation datasets. The data were randomly divided into training and validation sets, comprising 525 records (70%) for training and 225 records (30%) for validation.

Figure 5 illustrates the input structure of the forecasting model applied to the Mascarada River basin case study. At least 150 datasets were defined, each organized based on different initial water level conditions. For each dataset, the initial water level is combined with multiple forecasting horizons and their corresponding accumulated precipitation values, resulting in multiple input vectors associated with the same dataset.

Although the data are organized into datasets associated with a common initial water level, each input vector is processed independently by the model, with no direct dependency between forecasts generated for different lead times. Thus, for each combination of $N(T_0)$, $PA(h)$, and forecasting horizon h, the artificial neural network produces a single estimate of the future water level, denoted as $NP_0(h)$.

After the hyperparameter optimization process, the final architecture of the forecasting model was defined, consisting of an input layer, three hidden layers, and an output layer. The first two hidden layers comprised 116 neurons each, while the third hidden layer was configured with 60 neurons. The output layer contained a single neuron, corresponding to the predicted variable. Model training was performed with a maximum of 2,000 epochs and a batch size of 64.

This configuration characterizes the model as a non-recurrent multi-horizon approach, enabling forecasts with different lead times to be generated from the same initial state while preserving the independence among the estimates produced for each prediction horizon.

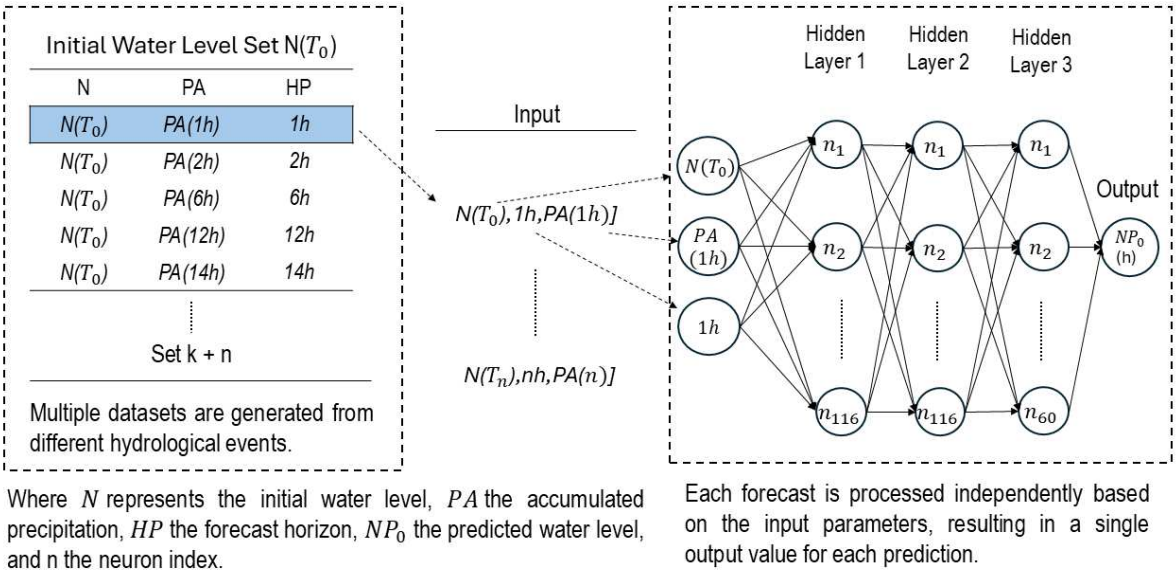


Figure 5. Water level forecasting model for the Mascarada River basin case study. Source: The Authors.

2.4. Evaluation Metrics

To evaluate the model performance, the Percent Bias (PBIAS), Nash–Sutcliffe efficiency coefficient (NSE), and the ratio of the root mean square error to the standard deviation of the observed data (RSR) were employed (MORIASI *et al.*, 2007). Table 2 presents the reference values used to assess the performance of the forecasting model.

Table 2. Reference values for PBIAS, RSR, and NSE. Source: Adapted from Moriasi et al. (2007).

Performance	PBIAS	RSR	NSE
Excellent	PBIAS $< \mp 10$	$0{,}00 \leq$ RSR $\leq 0{,}50$	$0{,}75 <$ NSE $\leq 1{,}00$
Good	$\mp 10 \leq$ PBIAS $< \mp 15$	$0{,}50 <$ RSR $\leq 0{,}60$	$0{,}65 <$ NSE $\leq 0{,}75$
Satisfactory	$\mp 15 \leq$ PBIAS $< \mp 25$	$0{,}60 <$ RSR $\leq 0{,}70$	$0{,}50 <$ NSE $\leq 0{,}65$
Unsatisfatory	PBIAS $\geq \mp 25$	RSR $> 0{,}70$	NSE $\leq 0{,}50$

3. Results and Discussion

The predictability of water levels in river basins is intrinsically associated with the complexity of hydrological processes, which are strongly influenced by the spatial heterogeneity of basin morphology, land use and land cover, and hydrodynamic characteristics. In meso-scale river basins, such variability results in nonlinear hydrological responses and imposes limitations on system predictability, particularly as the forecast horizon increases (BEVEN, 2012; SIVAPALAN *et al.*, 2003). Consequently, data-driven models are expected to perform better at shorter forecast horizons, for which the hydrological memory of the system still exerts a significant influence, with a progressive degradation of forecast accuracy at longer horizons, as widely discussed in the hydrological literature and in recent machine learning–

based modeling studies (BEVEN, 2012; HUANG; BARDOSSY, 2020; MOSAVI; OZTURK; CHAU, 2018). This behavior is consistent with evidence indicating that deep learning-based models are sensitive to the volume and representativeness of training data, exhibiting performance degradation under data-scarce conditions or increasing temporal complexity (PENG *et al.*, 2022).

The initial performance assessment of the model is based on the comparison between predicted and observed water levels for the training and validation datasets, as illustrated in Figure 6 (a) and (b). A strong correlation is observed between estimated and measured values, with coefficients of determination (R^2) ranging from 0.79 to 0.86, indicating adequate predictive capability of the model. Values of this magnitude are comparable to those reported in recent literature for different hydrological modeling approaches, including recurrent LSTM architectures applied to daily streamflow forecasting in long time series (RAHIMZAD *et al.*, 2021), as well as machine learning models developed for rainfall–stage prediction in data-limited basins, where R^2 values between 0.75 and 0.90 are considered indicative of good performance (KEDAM *et al.*, 2024; PATHAN *et al.*, 2024). results demonstrate that the conditional formulation adopted in this study achieves competitive performance even in the absence of explicit sequential modeling. Furthermore, comparative studies evaluating feedforward and recurrent architectures for hydrological simulation have shown that MLP-based models can achieve performance comparable to LSTM networks when properly parameterized (SHIN; JUNG, 2025).

It is also noted that the available historical record shows a higher concentration of water levels between 1.0 and 2.0 m, a range in which the model exhibited better agreement, a behavior also identified by Martel et al. (2024), who reported increased uncertainty associated with the prediction of extreme water levels.

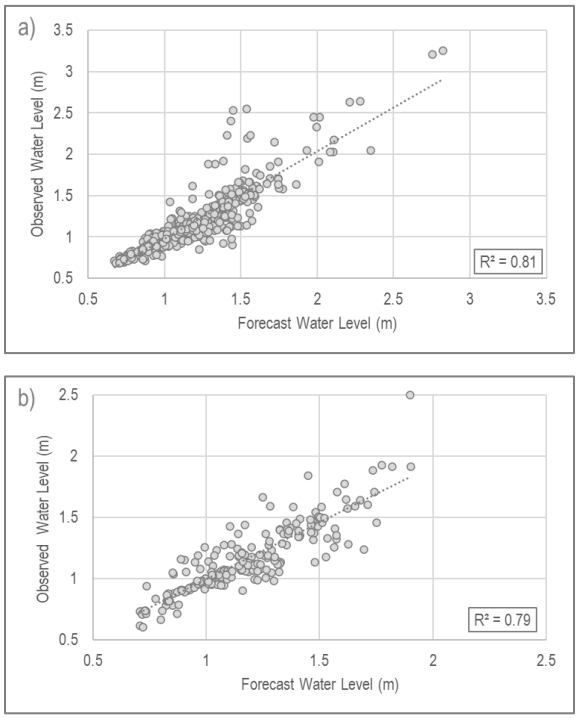


Figure 6. a) Correlation between observed and predicted water levels for the training dataset; b) Correlation between observed and predicted water levels for the validation dataset. Source: Author.

Regarding the model cross-validation, good performance was observed when comparing the mean values across the folds for the observed and predicted data. In this context, the mean Nash–Sutcliffe Efficiency (NSE) was estimated at 0.66, with a standard deviation of ±0.060. The mean RSR value was estimated at 0.57, with a standard deviation of ±0.047, while the mean PBIAS was calculated as −1.04, with a standard deviation of ±1.94, the latter being classified as excellent performance.

The performance assessment segmented by forecast horizon indicates that the best results are obtained for short lead times, particularly up to 6 hours ahead, with the 1-hour horizon standing out, as expected due to the shorter temporal distance between the initial conditions and the observed water level. This behavior is well documented in the literature, where a progressive decrease in predictive accuracy with increasing forecast horizon is commonly reported, regardless of the model architecture employed (HU *et al.*, 2018; PATHAN *et al.*, 2024; RAHIMZAD *et al.*, 2021).

In the present study, as show in Table 3, excellent results were obtained for both the training and forecasting scenarios up to a 6-hour lead time, whereas a loss of performance was observed for horizons longer than 12 hours, with results ranging from satisfactory to unsatisfactory, particularly for the 14-hour lead time. This behavior is expected and consistent with recent studies conducted in small basins with limited data availability, in which

hydrological predictability tends to decrease significantly as the forecast horizon increases (KEDAM *et al.*, 2024; XIMIN *et al.*, 2025). Nevertheless, the model's ability to provide reliable forecasts with lead times of up to 6 hours demonstrates strong potential for application in flood early warning systems, as highlighted in recent studies focused on hydrological risk management and decision-support frameworks (GOURLEY *et al.*, 2017; MARCOS JUNIOR *et al.*, 2024; PAINTER *et al.*, 2025).

Table 3. Model performance statistics for each forecast Horizon.

	Treinamento			Validação		
	PBIAS	RSR	NS	PBIAS	RSR	NS
Geral	-1,22	0,43	0,81	-2,79	0,48	0,77
1H	-1,02	0,18	0,97	-3,29	0,32	0,89
2H	-0,88	0,23	0,95	-2,26	0,42	0,82
6H	-1,83	0,29	0,92	-5,99	0,53	0,72
12H	-2,46	0,52	0,73	-3,45	0,70	0,51
14H	-0,11	0,59	0,65	0,12	0,64	0,58

From a hydrological perspective, the 14-hour forecast horizon may exceed the concentration time of the Mascarada River basin, particularly given its meso-scale characteristics and relatively short flow propagation times, which hampers reliable long-horizon predictions. In addition, the shift in model bias from overestimation to underestimation as the forecast horizon increases is a recurring behavior in hydrological forecasting models, commonly associated with increasing uncertainty at longer lead times and the limited representation of extreme events in the training dataset (PATHAN *et al.*, 2024).

The reduction in the model's predictive capability, particularly for the 14-hour forecast horizon, follows the widely documented tendency of performance degradation as the forecast horizon increases. At longer temporal scales, hydrological processes become more complex and nonlinear, increasing the system's sensitivity to uncertainties associated with initial conditions and input variables. This behavior is linked to the chaotic nature of hydrological systems, in which small initial errors tend to amplify over time, thereby limiting model predictability (HASAN *et al.*, 2024; LANGE; SIPPEL, 2020).

From the validation dataset, eight sets were randomly selected for the graphical representation of observed and predicted water levels, each comprising at least five forecast records, totaling 40 water-level observations. The set of plots presented in Figure 7 (a) and (b) illustrates the comparison between observed and predicted water levels for different forecast horizons, with Figure 7 (a) referring to the selected sets with initial water levels below 1.0 m

and Figure 7 (b) to cases with initial water levels above 1.0 m. This separation allows for the assessment of model performance under distinct initial hydrological conditions, enabling a clearer analysis of the influence of the initial water level on the model's predictive capability.

In both figures, a higher agreement between observed and predicted water levels is observed up to the 6 h forecast horizon, with increasing discrepancies at longer lead times. In Figure 7 (b) a closer correspondence between observed and predicted values can be noted, suggesting better model performance for initial water levels above 1.0 m. However, this behavior may have occurred randomly, as it was identified in only the eight selected datasets, which precludes its generalization to the remaining forecast data.

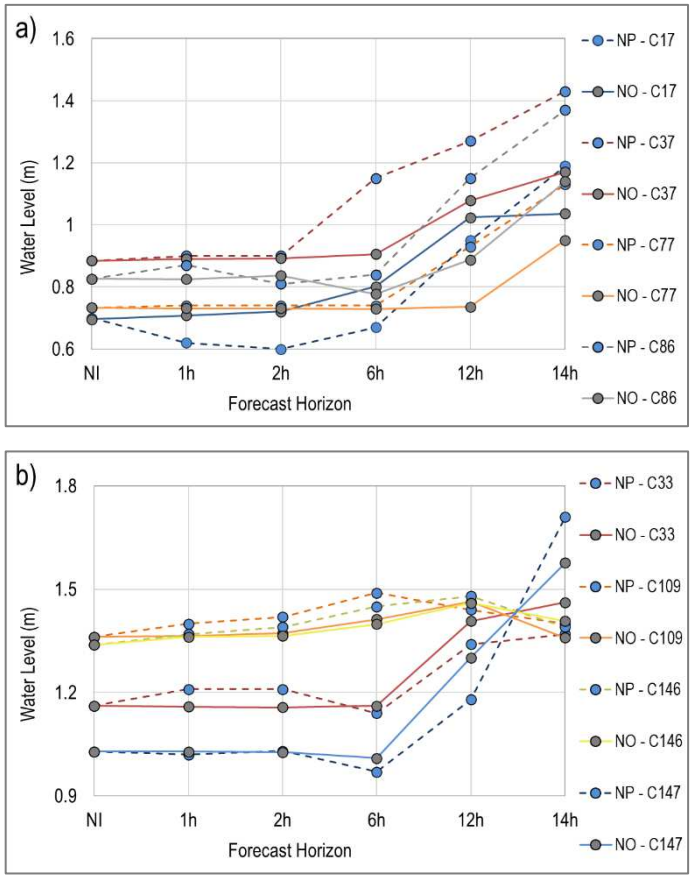


Figure 7. Comparison between observed (OL) and predicted (PL) water levels for four datasets across forecast horizons ranging from 1 h to 14 h: (a) initial water levels below 1.0 m; (b) initial water levels above 1.0 m.

Considering the complexity and spatial heterogeneity of the Mascarada River basin, whose morphological, hydrological, and land-use/land-cover characteristics vary from the headwater regions to the downstream reaches, the results obtained indicate a good model response for the prediction of hydrological events. Furthermore, the achieved performance is consistent with the results reported by (KEDAM *et al.*, 2024), who demonstrated the capability

of machine learning–based models to represent nonlinear hydrological processes and to provide satisfactory forecasts, particularly for shorter lead times, even in complex hydrological systems.

That said, it is noteworthy that the forecasting methodology developed in this study, even under conditions of limited data availability, demonstrated satisfactory performance in the estimation of hydrological events, highlighting its potential applicability in data-scarce contexts. These findings corroborate the results reported by JHONG et al. (2024), showing that, in addition to Long Short-Term Memory (LSTM)–based architectures, feedforward artificial neural network models can also be effective tools for forecasting abrupt flood events, particularly when appropriately structured according to the problem at hand.

Although recurrent models, such as LSTM, are widely used in hydrological forecasting, their application in data-scarce contexts is expected to be less effective in this case. In catchments with a limited number of extreme events and short historical records, more parsimonious architectures, such as feedforward artificial neural networks, tend to exhibit greater stability and improved generalization (HUANG; BARDOSSY, 2020; STAUDINGER *et al.*, 2025). Moreover, in the proposed method, the forecast lead time is explicitly incorporated into the input vector, imparting a conditional rather than sequential nature to the problem, thereby reducing the need for recurrent structures. Finally, unlike the studies conducted by MARTEL et al. (2024) and MALIK et al. (2024), which focus on continuous simulation of hydrological time series or temporally distributed forecasting, the approach adopted in this work is centered on estimating peak water levels at the Mascarada River gauging station, directly aligning with the objectives of supporting early warning systems.

3.1. Advantages and Limitations of the Proposed Methodology

The proposed methodology proved to be particularly suitable for application in data-scarce river basins, as it employs a feedforward artificial neural network in which the forecast lead time is explicitly included in the input vector. This structure enables multi-horizon forecasting in a simple and efficient manner, without requiring long time series, and demonstrates good performance especially for short- and medium-term lead times, which are particularly relevant for flood warning and response actions.

From an operational standpoint, the proposed approach offers significant advantages over traditional hydrological and hydrodynamic models, which typically require larger data volumes, more complex calibration procedures, and greater computational effort. Direct water

level forecasting enables a simpler and more agile implementation, making it particularly suitable for river basins with limited monitoring and short historical records. In this context, even with a restricted data collection period, combined with the installation of real-time water level sensors, the development of local hydrological warning systems based exclusively on the proposed methodology becomes feasible, highlighting the high replicability potential of the method in other data-constrained river basins.

Additionally, the direct use of water level as an input variable contributed to simplifying the forecasting process based on precipitation by reducing the uncertainties associated with the conversion from stage–discharge relationship through rating curves, a procedure commonly adopted in traditional hydrological studies and often prone to inconsistencies under data-scarce conditions. This methodological choice enhances the robustness and operational applicability of the model, particularly for near–real-time decision-support applications.

As limitations, a reduction in prediction accuracy for longer forecast horizons is highlighted, likely associated with the increased complexity of hydrological processes and the consequent decrease in system predictability. It is also emphasized that the methodology was not trained to generate forecasts from initial water levels close to or corresponding to the event peak, which may restrict its applicability under such specific conditions. Furthermore, as with other machine learning–based models, the proposed approach exhibits limited extrapolation capability for extreme water levels that are poorly represented in the training dataset, a condition that should be considered when interpreting the results.

Finally, the potential for expanding the applicability of the proposed method should be emphasized through the adoption of additional training and evaluation strategies, such as the targeted inclusion of new extreme events in the dataset, segmentation of the learning process according to water level ranges, and integration with hybrid approaches that combine data-driven models and hydrological concepts. Recent studies have explored hybrid frameworks that integrate deep neural networks, signal decomposition techniques, and evolutionary optimization algorithms to enhance predictive performance in complex hydrological scenarios (FENG *et al.*, 2025). The incorporation of similar strategies may represent a promising avenue for future research.

4. Conclusion

The results of this study indicate that the proposed peak forecasting methodology achieved satisfactory performance in estimating hydrological events in the Mascarada River basin, even under conditions of limited data availability. In this context, the proposed approach enables the implementation of simplified early warning systems, in which direct prediction of water level eliminates the need for discharge estimation and reduces uncertainties associated with rating curves.

The model demonstrated better performance for shorter forecast horizons, with progressive degradation in predictive accuracy as the lead time increased, a behavior consistent with recent literature on machine learning applications in hydrology (KEDAM *et al.*, 2024; PATHAN *et al.*, 2024). From an operational perspective, the methodology presents advantages by requiring a low volume of data, simplified calibration procedures, and reduced computational effort compared to traditional hydrological approaches. In this regard, unlike physically based hydrological or hydrodynamic models—which require detailed basin characterization, geomorphological data, and extensive historical records for robust calibration the proposed methodology essentially relies on continuous monitoring of water level and precipitation, thereby broadening its applicability in regions with limited hydrometeorological infrastructure.

Overall, the model proved to be more reliable for forecasts with lead times of up to 6 hours, while longer horizons showed reduced performance, reflecting the increasing complexity of hydrological processes and the inherent decrease in system predictability. Nevertheless, the findings reinforce the potential of the proposed approach as a viable and rapidly deployable alternative for hydrological forecasting and extreme event management in meso-scale basins with limited observational data.

Finally, the results obtained for the Mascarada River basin demonstrate that, even with a relatively simple modeling structure and limited data availability, the proposed methodology can support hydrological early warning systems and local risk management, particularly for short-term forecasts, thereby contributing in a practical manner to reducing regional vulnerability to flood events. As future perspectives, further application of the methodology to different river basins and evaluation of model performance under more severe extreme events or changing hydrological conditions are recommended. Additional investigations may include

the incorporation of complementary variables and the assessment of predictive uncertainties in order to further enhance the robustness and reliability of the proposed framework.

Declarations

Author Contributions

Vinícius Santanna Castiglio: conceived the study, developed the methodology, performed the analysis, and wrote and revised the manuscript.

Gean Paulo Michel: supervision, conceptual guidance, writing – review and editing.

Funding

This study was supported by the Coordination for the Improvement of Higher Education Personnel (CAPES), Brazil, under Ordinance No. 120/2024 – Program for Supporting the Dissemination of Scientific and Technological Information (PADICT).

Competing Interests

The authors declare that they have no competing interests.

References

BEVEN, K. J. **Rainfall-runoff modelling : the primer**. [s.l.] Wiley-Blackwell, 2012. 457 p.

COLLISCHONN, W.; FAN, F. M.; POSSANTTI, I.; DORNELLES, F.; PAIVA, R.; MEDEIROS, M. S.; MICHEL, G. P.; FILHO, F. J. C. M.; MORAES, S. R.; MARCUZZO, F. F. N.; MICHEL, R. D. L.; BESKOW, T. L. C.; BESKOW, S.; FERNANDES, E. H. L.; DOS SANTOS, L. L.; RUHOFF, A.; KOBIYAMA, M.; COLLARES, G. L.; BUFFON, F.; DUARTE, E.; LIMA, S.; MEIRELLES, F. S. C.; PICCILLI, D. G. A. The exceptional hydrological disaster of April-May 2024 in southern Brazil. **Revista Brasileira de Recursos Hidricos**, v. 30, 2025.

EVANGELISTA, L.; MÓLLER, D.; BRENTAN, B.; MEIRELLES, G. Application of Feedforward Artificial Neural Networks to Predict the Hydraulic State of a Water Distribution Network. **Engineering Proceedings**, p. 49, 4 set. 2024.

FENG, Z. K.; LIU, W. J.; NIU, W. J.; YANG, T.; WANG, W. C.; WANG, S. Compound Hydrological Forecasting Model by Long Short-term Memory Network Coupled with Adaptive Mode Decomposition and Evolutionary Algorithm. **Water Resources Management**, v. 39, n. 6, p. 2651–2672, 1 abr. 2025.

GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. **Deep Learning**. [s.l: s.n.]

GOURLEY, J. J.; FLAMIG, Z. L.; VERGARA, H.; KIRSTETTER, P. E.; CLARK, R. A.; ARGYLE, E.; ARTHUR, A.; MARTINAITIS, S.; TERTI, G.; ERLINGIS, J. M.; HONG, Y.; HOWARD, K. W. The flash project

Improving the Tools for Flash Flood Monitoring and Prediction across the United States. **Bulletin of the American Meteorological Society**, v. 98, n. 2, p. 361–372, 1 fev. 2017.

GREGORY, J. M. **Time of Concentration Determination for Small Watersheds**. [s.l: s.n.].

HASAN, F.; MEDLEY, P.; DRAKE, J.; CHEN, G. Advancing Hydrology through Machine Learning: Insights, Challenges, and Future Directions Using the CAMELS, Caravan, GRDC, CHIRPS, PERSIANN, NLDAS, GLDAS, and GRACE Datasets. **Water**, v. 16, n. 13, p. 0–31, 3 jul. 2024.

HU, C.; WU, Q.; LI, H.; JIAN, S.; LI, N.; LOU, Z. Deep learning with a long short-term memory networks approach for rainfall-runoff simulation. **Water**, v. 10, n. 11, 30 out. 2018.

HUANG, Y.; BARDOSSY, A. Impacts of data quantity and quality on model calibration: Implications for model parameterization in data-scarce catchments. **Water (Switzerland)**, v. 12, n. 9, 1 set. 2020.

IJOMAH, W. L.; NWANKPA, C. E.; MARSHALL, S.; IJOMAH, W. L.; GACHAGAN, A. Activation Functions: Comparison of trends in Practice and Research for Deep Learning. Em: 2nd International Conference on Computational Sciences and Technologies, 2020, [...]. 2020. p. 17–19.

JHONG, Y. Da; CHEN, C. S.; JHONG, B. C.; TSAI, C. H.; YANG, S. Y. Optimization of LSTM Parameters for Flash Flood Forecasting Using Genetic Algorithm. **Water Resources Management**, v. 38, n. 3, p. 1141–1164, 1 fev. 2024.

KEDAM, N.; TIWARI, D. K.; KUMAR, V.; KHEDHER, K. M.; SALEM, M. A. River stream flow prediction through advanced machine learning models for enhanced accuracy. **Results in Engineering**, v. 22, 1 jun. 2024.

KEETON, Kimberly. **Proceedings of the 12th USENIX conference on Operating Systems Design and Implementation**. [s.l.] USENIX Association, 2016. 786 p.

LANGE, H.; SIPPEL, S. Machine Learning Applications in Hydrology. *Em*: **Forest-Water Interactions**. [s.l.] Springer, 2020. p. 233–257.

LI, J.; CHEN, Y.; ZHU, Y.; LIU, J. Study of Flood Simulation in Small and Medium-Sized Basins Based on the Liuxihe Model. **Sustainability (Switzerland)**, v. 15, n. 14, 2023.

LI, L.; JAMIESON, K.; ROSTAMIZADEH, A.; TALWALKAR, A. Hyperband: A Novel Bandit-Based Approach to Hyperparameter Optimization. **Journal of Machine Learning Research**, v. 18, p. 1–52, 2018.

MALIK, H.; FENG, J.; SHAO, P.; ABDULJABBAR, Z. A. Improving flood forecasting using time-distributed CNN-LSTM model: a time-distributed spatiotemporal method. **Earth Science Informatics**, 1 ago. 2024.

MARCOS JUNIOR, A. D.; SILVEIRA, C. da S.; DA COSTA, J. M. F.; GONÇALVES, S. T. N. Combining traditional hydrological models and machine learning for streamflow prediction. **Revista Brasileira de Recursos Hidricos**, v. 29, 2024.

MARTEL, J.-L.; ARSENAULT, R.; TURCOTTE, R.; CASTAÑEDA-GONZALEZ, M.; BRISSETTE, F.; ARMSTRONG, W.; MAILHOT, E.; PELLETIER-DUMONT, J.; LACHANCE-CLOUTIER, S.; RONDEAU-GENESSE, G.; CARON, L.-P. Exploring the ability of LSTM-based hydrological models to simulate streamflow time series for flood frequency analysis. **EGUsphere**, p. 2024–2134, 20 ago. 2024.

MORIASI, D. N.; ARNOLD, J. G.; LIEW, M. W. Van; BINGNER, R. L.; HARMEL, R. D.; VEITH, T. L. **MODEL EVALUATION GUIDELINES FOR SYSTEMATIC QUANTIFICATION OF ACCURACY IN WATERSHED SIMULATIONSTransactions of the ASABE**. [s.l: s.n.].

MOSAVI, A.; OZTURK, P.; CHAU, K. W. **Flood prediction using machine learning models: Literature reviewWater (Switzerland)**MDPI AG, 27 out. 2018.

PAINTER, P.; SEMMENS, K.; MAXFIELD, K.; CATTOËN, C.; CARR, R. H. **Designing Effective Flood Early Warning Systems: A Review of Barriers, Best Practices, and Key CharacteristicsJournal of Flood Risk Management**John Wiley and Sons Inc, 1 dez. 2025.

PATHAN, A. I.; SIDEK, L. B. M.; BASRI, H. B.; HASSAN, M. Y.; KHEBIR, M. I. A. Bin; OMAR, S. M. B. A.; KHAMBALI, M. H. bin M.; TORRES, A. M.; NAJAH AHMED, A. Comparative assessment of rainfall-based water level prediction using machine learning (ML) techniques. **Ain Shams Engineering Journal**, v. 15, n. 7, 1 jul. 2024.

PENG, A.; ZHANG, X.; XU, W.; TIAN, Y. Effects of Training Data on the Learning Performance of LSTM Network for Runoff Simulation. **Water Resources Management**, v. 36, n. 7, p. 2381–2394, 1 maio 2022.

RAHIMZAD, M.; MOGHADDAM NIA, A.; ZOLFONOON, H.; SOLTANI, J.; DANANDEH MEHR, A.; KWON, H. H. Performance Comparison of an LSTM-based Deep Learning Model versus Conventional Machine Learning Algorithms for Streamflow Forecasting. **Water Resources Management**, v. 35, n. 12, p. 4167–4187, 1 set. 2021.

ROBAINA, E. de S. L.; TRENTINI, A. A. V. S. R.; VALLI, A. de N. C. V. M. B.; PEREIRA, S. **Desastres Hidrológicos: Levantamento para o estado do Rio Grande do Sul, Brasil**. [s.l: s.n.].

RUCKERT, A.; VICENTE, F. J.; GOMES, L. F. **A Tragédia Climática No Rio Grande Do Sul Em 2024. Anotações Sobre Uma Catástrofe Anunciada**. [s.l: s.n.].

RUIZ-VILLANUEVA, V.; DÍEZ-HERRERO, A.; BODOQUE, J. M.; BALLESTEROS CÁNOVAS, J. A.; STOFFEL, M. Characterisation of flash floods in small ungauged mountain basins of Central Spain using an integrated approach. **Catena**, v. 110, p. 32–43, nov. 2013.

SARCHANI, S.; TSANIS, I. Analysis of a flash flood in a small basin in Crete. **Water (Switzerland)**, v. 11, n. 11, 2019.

SHADMEHRI TOOSI, A.; BATELAAN, O.; SHANAFIELD, M.; GUAN, H. **Land Use-Land Cover and Hydrological Modeling: A ReviewWiley Interdisciplinary Reviews: Water**John Wiley and Sons Inc, 1 mar. 2025.

SHIN, M. J.; JUNG, Y. Comparison of Multilayer Perceptron with an Optimal Activation Function and Long Short-Term Memory for Rainfall-Runoff Simulations and Ungauged Catchment Runoff Prediction. **Water Resources Management**, v. 39, n. 6, p. 2479–2502, 1 abr. 2025.

SIVAPALAN, M.; TAKEUCHI, K.; FRANKS, S. W.; GUPTA, V. K.; KARAMBIRI, H.; LAKSHMI, V.; LIANG, X.; MCDONNELL, J. J.; MENDIONDO, E. M.; O'CONNELL, P. E.; OKI, T.; POMEROY, J. W.; SCHERTZER, D.; UHLENBROOK, S.; ZEHE, E. IAHS Decade on Predictions in Ungauged Basins (PUB), 2003-2012: Shaping an exciting future for the hydrological sciences. **Hydrological Sciences Journal**, v. 48, n. 6, p. 857–880, dez. 2003.

SRIVASTAVA, N.; HINTON, G.; KRIZHEVSKY, A.; SALAKHUTDINOV, R. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. **Journal of Machine Learning Research**, v. 15, p. 1929–1958, 2014.

STAUDINGER, M.; HERZOG, A.; LORITZ, R.; HOUSKA, T.; POOL, S.; SPIELER, D.; WAGNER, P. D.; MAI, J.; KIESEL, J.; THOBER, S.; GUSE, B.; EHRET, U. How well do process-based and data-driven hydrological models learn from limited discharge data? **Hydrology and Earth System Sciences**, v. 29, n. 19, p. 5005–5029, 8 out. 2025.

XIMIN, Y.; TIANZHI, W.; FUCHANG, T.; XIAOJIAO, W. A novel hybrid SCS-LSTM model for intelligent flood forecasting in data-scarce small mountainous catchments. **Journal of Hydrology: Regional Studies**, v. 61, 1 out. 2025.

ZAHRI, R.; MD ALI, A. H.; MATSUURA, S.; RAMBAT, S. Enhancing Youth Resilience in Disaster Risk Reduction: Insights from a systematic literature review. **Environment-Behaviour Proceedings Journal**, v. 10, n. SI33, p. 63–70, 15 jun. 2025.