Proceeding Series of the Brazilian Society of Computational and Applied Mathematics, v. 12, n. 1, 2026.

Trabalho apresentado no XLIV CNMAC, Fundação Getulio Vargas - Rio de Janeiro - RJ, 2025.

Proceeding Series of the Brazilian Society of Computational and Applied Mathematics

River Level Forecasting Using a Multilayer Perceptron: A Case Study in the Tamanduateí Basin

Rogério C. O. Junior[1], Giovanni G. Soares
INPE, Instituto Nacional de Pesquisa Espacial, São José dos Campos, São Paulo
Fernando L. S. Filho
Unifesp, Universidade Federal de São Paulo, São Paulo
Rogério G. Negri
Unesp, Universidade Estadual Paulista, São José dos Campos, São Paulo
Leonardo B. L. Santos, Elton V. E. Silva
Cemaden, Centro Nacional de Monitoramento e Alertas de Desastres Naturais, São José dos Campos, São Paulo

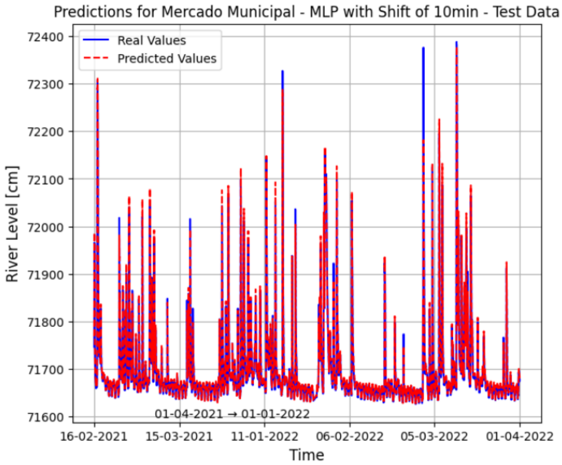
Accurate river level forecasting is crucial for effective hydrological management, particularly in urban basins susceptible to flooding. This study employs a *Multi-Layer Perceptron* (MLP) neural network to predict water levels in the Tamanduateí Basin, Brazil, using high-frequency data collected every 10 minutes from six monitoring stations. The dataset spans the first three months of each year from 2018 to 2022, totaling 63,959 data points per station. The model focuses on forecasting water levels at the Mercado Municipal station, a key location downstream, across various time horizons (10, 60, and 120 minutes). The MLP-based model demonstrates strong short-term predictive capabilities, but its performance diminishes for longer-term forecasts, revealing limitations in modeling complex hydrological dynamics. These findings align with previous studies, such as [2], which highlighted the challenges of long-term forecasting in complex basins like the Rio Negro. The MLP model was chosen for its ability to capture nonlinear relationships in hydrological data, as demonstrated in [3]. The proposed model uses water level data from six monitoring stations to predict the water level at the Mercado Municipal station. Separate MLP models were developed for different forecasting horizons, with the network architecture consisting of an input layer with six neurons, two hidden layers with 36 neurons each, and a single output neuron. The ReLU activation function was used in the hidden layers, and the training process was optimized using the OPTUNA framework. The MLP model demonstrated high effectiveness in short-term river level forecasting at the Mercado Municipal station. Figure 1(a) summarizes the model evaluation according to different forecasting performance measures at distinct forecast horizons. Figure 1(b) shows the observed and predicted levels at this station. For the 10-minute prediction horizon, the model achieved a Mean Absolute Error (MAE) of 4.9041, a Root Mean Squared Error (RMSE) of 8.5724, and a Nash-Sutcliffe Efficiency (NSE) of 0.9903, indicating excellent predictive performance. However, as the forecasting horizon increased, the model's performance declined. For the 120-minute prediction, the MAE rose to 22.7925, the RMSE increased to 49.3505, and the NSE dropped to 0.6775, highlighting the challenges of long-term forecasting. These results are consistent with the findings of [1], who emphasized the limitations of data-driven models in capturing long-term hydrological dynamics. Regardless of the horizons studied, the loss resulted in values close to zero.

The decline in forecasting performance for longer forecasting horizons suggests that the model does not fully capture the intricate interactions between hydrological processes and external factors, such as urban drainage and rainfall. To address these limitations, future improvements could

[1]rogerio.costa@inpe.br

Mercado Municipal station			
Metrics	10 min	60 min	120 min
MAE	4.9041	16.9261	22.7925
MSE	73.4857	946.3468	2435.4744
RMSE	8.5724	30.7628	49.3505
NSE	0.9903	0.8747	0.6775
Loss	0.0002	0.0015	0.0035

(a)



(b)

Figure 1: (a) Model evaluation and (b) the level observed and predicted at the Mercado Municipal station. Source: Authors.

involve incorporating additional environmental variables or adopting hybrid modeling approaches that combine data-driven techniques with physical hydrological models. This approach aligns with the recommendations of [2], who proposed integrating physical laws into neural networks to improve long-term forecasting performance.

In conclusion, the study demonstrates that MLP models are effective for short-term river level forecasting but face challenges in longer-term predictions. The integration of additional environmental variables and the adoption of hybrid modeling approaches, such as *Physics-Informed Neural Networks* (PINNs), could significantly improve the performance and robustness of long-term forecasts. Future research will focus on incorporating physical laws, such as the Saint-Venant equations, into the learning process to enhance the model's predictive capabilities and provide more reliable tools for hydrological forecasting in urban basins like the Tamanduateí.

Acknowledgements: This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) – Brasil – Finance Code 001 and by the São Paulo Research Foundation (FAPESP) under grant number 2024/02748-7.

References

[1] C. W. Dawson and R. L. Wilby. "Hydrological modelling using artificial neural networks". In: **Progress in Physical Geography** 25.1 (2001), pp. 80–108.

[2] J. S. Duque, R. Santos, J. Arteaga, R. S. Oyarzabal, and L. B. L. Santos. "Nonlinear hydrological time series modeling to forecast river level dynamics in the Rio Negro Uruguay basin". In: **Chaos: An Interdisciplinary Journal of Nonlinear Science** 34.5 (2024), p. 053132. DOI: 10.1063/5.0201784.

[3] F. Kratzert, D. Klotz, G. Shalev, G. Klambauer, S. Hochreiter, and G. Nearing. "Towards learning universal, regional, and local hydrological behaviors via machine learning applied to large-sample datasets". In: **Hydrology and Earth System Sciences** 23.12 (2019), pp. 5089–5110.