Asian Journal of Mathematics and Computer Research

Volume 31, Issue 4, Page 10-20, 2024; Article no.AJOMCOR.12387
ISSN: 2456-477X

Multi Time Series WA-LSTM-Adam for Water Level Forecasting in Center Vietnam

Nguyen Duc Khoa [a], Nguyen Quang Dat [a*], Vo Quang Linh [b], Nguyen Ha Vy [a], Vu Hoang Nam Khanh [b] and Phan Viet Hoang [b]

[a] *High School for Gifted Students, Hanoi University of Science - VNU, Hanoi, Vietnam.*
[b] *Hanoi-Amsterdam High School for Gifted Students, Hanoi, Vietnam.*

Authors' contributions

This work was carried out in collaboration among all authors. All authors read and approved the final manuscript.

Article Information

DOI: https://doi.org/10.56557/ajomcor/2024/v31i48891

Open Peer Review History:
This journal follows the Advanced Open Peer Review policy. Identity of the Reviewers, Editor(s) and additional Reviewers, peer review comments, different versions of the manuscript, comments of the editors, etc are available here: https://prh.ikprress.org/review-history/12387

Original Research Article

Received: *20/07/2024*
Accepted: *24/09/2024*
Published: *10/10/2024*

Abstract

The central region of Vietnam suffers from floods almost every year as a result of a combination of frequent storms, heavy rainfall, and short, steep rivers in the region. This is a big problem because they can negatively affect the economy of the region as well as people's lives when not managed properly. Therefore,

---

**Corresponding author: E-mail: nguyenquangdat@hus.edu.vn;*

Cite as: *Khoa, Nguyen Duc, Nguyen Quang Dat, Vo Quang Linh, Nguyen Ha Vy, Vu Hoang Nam Khanh, and Phan Viet Hoang. 2024. "Multi Time Series WA-LSTM-Adam for Water Level Forecasting in Center Vietnam". Asian Journal of Mathematics and Computer Research 31 (4):10-20. https://doi.org/10.56557/ajomcor/2024/v31i48891.*

it is important to have a reliable forecasting method for flooding in order to ensure effective natural disaster management. In this research, we aim at addressing this issue by introducing a multi time series hybrid deep learning model that combines WA (wavelet analysis) and LSTM (long-short-term memory) optimized with the Adam algorithm and uses water level and rainfall data as the input variables. Compared to other traditional methods and some recent models, our WA-LSTM-Adam method shows better results overall.

Keywords: rainfall; water level; time series forecasting; wavelet; LSTM; adam.

1 Introduction

In Vietnam, storms tend to concentrate more in the central compared to other regions. This, along with the fact that the characterized short, steep rivers in the rock mountain region cause floods to happen more frequently there, severely damaging property and threatening human lives. Therefore, having a reliable prediction method is needed. Due to the complexities of flood forecasting, we decided that a deep learning method is a good model to solve this problem.

In the study, we realized using only the time series of water level is not enough information due to the frequent storms, heavy rains, and the rock mountain featured in the region. So for the input data, we use both the time series of rainfall and that of water level.

Traditional RNN models have some problems when dealing with the vanishing gradient problem. LSTM partially solves this problem, and therefore can learn more complex learning patterns. GRU is also a type of LSTM which requires less material and training time to run, but it is not as effective when dealing with larger problems with high accuracy requirements.

Wavelet Analysis (WA) is a mathematical method dividing time signal, which is excellent for removing noise.

Adam is a good optimizing algorithm when dealing with large problem, providing fast convergence and having little memory requirement.

Therefore, we use a hybrid model between the WA combined with LSTM, using Adam as the optimizer to retain all the advantage the aforementioned models have.

2 Related Works

The prediction of water level of rivers can be applied in a lot of cohorts and plays a vital roles in flood management, especially for the countries relying heavily on agriculture.

2.1 Wavelet Analysis - WA

The study in 2023, written by Ehsan Azizi et al. [1] , presented a machine learning model using wavelet transform analysis to predict groundwater levels. The method's results fluctuated around 3% and 10%, illustrating the high accuracy of the new model.

In 2023, Youming Li et al. [2] tested a method based on the LSTM neural network and enhanced with Bayesian optimization and wavelet decomposition (BO-WD-LSTM) for water level prediction of the reservoir in Liuxihe River Basin. The prediction of BO-WD-LSTM achieved higher accuracy compared to other single and optimization methods, as the errors of this model are 0.4084m and 0.1987m.

Table 1.Wavelet analysis

Research objective	Forecasting method	Year	
Modelling and prediction of groundwater level using wavelet transform and machine learning methods	Machine learning methods using wavelet transform (WT)	2023	[1]
Water level prediction of Liuxihe Reservoir	BO-WD-LSTM	2023	[2]
Machine learning models combined with wavelet transform and phase space reconstruction for groundwater level forecasting	WT-PSR-ANN	2023	[3]

In 2023, Aihua Wei et al. [3] presented a model that combined WT (wavelet transform) with PSR (phase space reconstruction) before being implemented with ANN. This WT-PSR-ANN model had achieved the best accuracy among all other testing methods, showing its reliability to predict groundwater levels.

2.2 LSTM

Table 2. LSTM

Research objective	Forecasting method	Year	
Prediction of flood by deep learning model	LSTM	2022	[4]
Prediction of water levels using machine learning model in Red River	LSTM	2022	[5]
Water level forecasting on the Bangladesh river network	PSO-LSTM	2023	[6]

In 2022, Selle Nevo et al. [4] tested 4 different ML methods for flood forecasting the river system in India and Bangladesh, and it seems that the LSTM with extra precipitation input model is better in comparison with the remaining methods, as the NSE (Nash-Sutcliffe efficiency) and persistent-NSE medians for this forecast model are 0.99 and 0.69, respectively.
In the same year, Vida Atashi et al. [5] tried three different methods for flood prediciton, namely LSTM, SARIMA, and RF, which showed that LSTM outperformed the other two, as the RMSE value in all stations is much lower, at 0.190, 0.151, and 0.107.

In 2023, Jannatul Ferdous Ruma et al. [6] presented a new method to forecast water levels in Bangladesh, which is a LSTM model using particle swarm optimization. After testing with several other models, this PSO-LSTM method had been recorded to have the highest NSE value, at about 0.95 after 15 days.

2.3 Hybrid models

Table 3. Some hybrid models

Research objective	Forecasting method	Year	
Satellite-based LSTM for Jakarta's water level forecasting	LSTM, RNN	2022	[7]
Water level forcasting	GRU-LSTM	2022	[8]
Prediction of water level in reservoir	SARIMA-ANN	2022	[9]
Forecasting daily river runoff	VMD-CNN-AM-BOA-BiLSTM	2023	[10]
Prediction of water level in reservoir	BO-WD-LSTM	2023	[2]
Monthly Water Inflow Forecasting	CNN-GRU	2024	[11]

In 2022, Hadi Kardhana, Faizal Immaddudin Wira Rohmat and their companions [7] used a hybrid model for flood forecasting. To do that, they combined LSTM and RNN and then optimized it using the ADAM algorithm. The LSTM-RNN had some promising results, with an R^2 of 0.98 while training and R^2 of 0.86 while testing.

In 2022, Minwoo Cho et al. [8] used a GRU-LSTM model for water level forecasting. This method showed a good result and had better accuracy than GRU or LSTM models, as the MSE, NSE and MAE were 3.92, 0,942 and 2.22 respectively.

In 2022, Abdus Samad Azad, Rajalingam Sokkalingam and their companions [9] presented a SARIMA-ANN hybrid-time series model for reservoir water level forecasting. The model showed a good result, as it outperformed the ANN and RNN models (R^2=0.84,MAE=328.69, MAPE=32,868.51%)

In 2023, Junhao Wu et al. [10] showed a deep learning model used for prediction of daily river runoff. To do that, they combined VMD for data decomposition, CNN for feature extraction, BiLSTM with attention mechanism for temporal modeling, and Bayesian optimization for hyperparameter tuning, which called VMD-CNN-AM-BOA-BiLSTM. The accuracy was reliable, with MAE of 12.74 and MAPE of 0.03 when the model was used to forecast flood peak.

In 2023, Youming Li and his companion [2] tested a method by using LSTM neural network and optimized it with Bayersian and wavelet analysis (BO-WD-LSTM) to forecast water level of a reservoir in Liuxihe river. The model BO-WD-LSTM had achieved higher accuracy compared to other single and optimization methods, as the errors of this model are 0.4084m and 0.1987m.

In 2024, Wenxin Xu et al. [11] introduced a hybrid method which can be applied to monthly water inflow forecasting. The model used the a simplified version of VIC as the hydrological model and combined the CNN (convolutional neural network) and GRU (Gated recurrent network) to be a deep learning method. The model showed great ability to make monthly streamflow predictions with the lack of training data, with test results of KGE, R^2, and WI (Willmott's Index of Agreement) of about 0.53, 0.32, and 0.74, respectively.

3 Methodology

3.1 Wavelet analysis - WA

Wavelet analysis is a frequency-analysis mathematical tool developed by Mallat that has many applications, including machine learning. It is used to analyze and transform data. By splitting the data into different frequency parts, this reveals noise at higher frequencies, and then noise is suppressed to produce a cleaner version of the original data, therefore effectively eliminating noise. There exist three types of wavelet analysis, namely the CWT (continuous wavelet transform), DWT (discrete wavelet transform), and SWT (stationary wavelet transform), each of them can be applied to different fields. In this particular study, we will focus only on DWT, a typical method for noise removal.

The following can be used to show the DWT:

$$W_{(m,n)} = 2^{-\frac{m}{n}} \sum_{t=0}^{N-1} \psi\Big(\frac{t - 2^m \cdot n}{2^m}\Big) \tag{3.1}$$

in which:

- $2^m = s$: Scaling parameter
- $2^m \cdot n = \tau$: Translation parameter

When $|s| > 1$, the wavelet signal expanded for the initial signal. Oppositely, when $0 < |s| < 1$, the wavelet signal has been compacted from the original one.

Fig. 1 shows that a first-order wavelet is used to split the input data into two parts: low-frequency D_1 and high-frequency A_1. We use a 2^{nd}-order wavelet to continue splitting the high-frequency A_1 into the low-freq D_2 and the high-freq A_2. This process can be repeated to further decompose the high frequency part. After performing the process N time, we obtain a set of part (D_1, D_2, D_3 ,. .., D_N, A_N). The high-frequency component A_N tend to be noise after repeating the splitting process several times and should be removed. We can find the the wavelet level N by this following equation:

$$N = log_{10}(n) \tag{3.2}$$

with n stands for the original size of the time series.

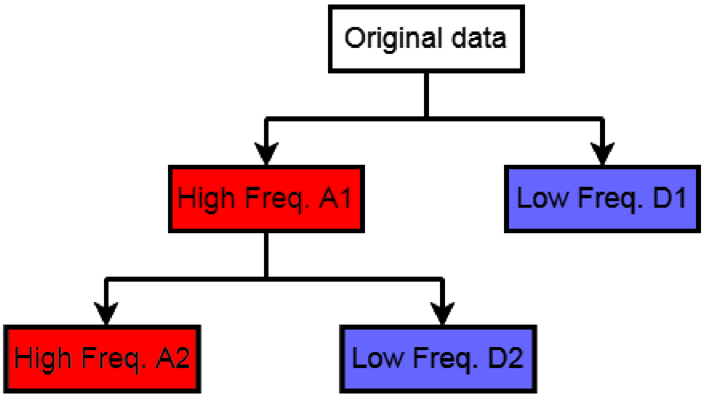

Fig. 1. Wavelet analysis: 2 level

3.2 RNN

When using feedforward neural networks (FNN), the $N - 1$ earlier value (step) stands for the present result.

However, the recurrent neural network is different, as it contain extra connection between two consecutive time steps. The RNN can also compact its entire history in the lower dimensional space, make it superiour in comparison with the FNN which can only compress one past value. Moreover, the RNN can form short-term memory, enable it to adapt with different position variances, while that does not true for FNN. Also, nodes in RNN are abel to have self-connection overtime and share their weights between different time steps. Therefore, adjusting the input data becomes much simpler and more effective when using RNN.

There are several problems with this model, as it shows its inability while capturing and handling long-term dependencies [12, 13].

3.3 Long-Short term memory - LSTM

To address the aforementioned problem, in 1997, [12, 13] Hochreiter and Schmidhuber [12, 14] came up with a new variant of RNN, called LSTM. To do this, a LSTM unit contained 3 gates and memory cells (to store data during training). With this design, LSTM has the capability to handle long-term dependencies more effectively than traditional RNN, partially solve one of RNN's main drawback. The following figure illustrates how the LSTM unit works:

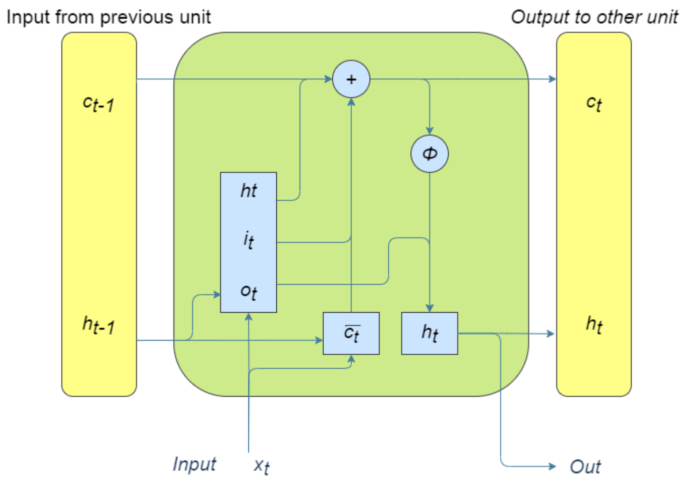

Fig. 2. How a LSTM unit works

A LSTM works according to the following steps:

- Step 1: The forgotten gates eliminated information from the cell of state:

$$f_t = \sigma(W_f.[h_{t-1}, x_t] + b_f) \tag{3.3}$$

- Step 2: *Cell state* gets more new information.

$$i_t = \sigma(W_i.[h_{t-1}, x_t] + b_i) \tag{3.4}$$
$$\tilde{c}_t = \tanh(W_{\tilde{c}}.[h_{t-1}, x_t] + b_{\tilde{c}}) \tag{3.5}$$

- Step 3: *state cells* are changed from C_{t-1} to C_t:

$$c_t = f_t \otimes c_{t-1} \oplus i_t \otimes \tilde{c}_t \tag{3.6}$$

- Step 4: Generating output

$$o_t = \sigma(W_o.[h_{t-1}, x_t] + b_o) \tag{3.7}$$
$$h_t = o_t \otimes \tanh(h_t) \tag{3.8}$$

with $W_f, W_i, W_{\tilde{c}}, W_o$ are parameters of LSTM and $b_f, b_i, b_{\tilde{c}}, b_o$ are biases of LSTM model.

3.4 Adam-optimization

ADAM optimization is a popular optimizing method when used in machine learning, especially in the deep learning field. It is a combination of AdaGrad and RMSProp and therefore has both of the advantages of adaptive learning rates for each parameter based on the first (mean) and second (variance) moments of the gradients. With this, ADAM is efficient, requiring little memory while maintaining robustness in various settings. Also, its ability to converge quickly makes it suitable for machine learning tasks.
The following pseudo-code is shown below:

Algorithm 1 ADAM-optimization

Input n: Stepsize
Input γ_1, γ_2: Exponential decay rates for the moment estimates
Input $f(\theta)$: Stochastic objective function with parameters θ
Input θ_0: Initial parameter vector
  $m_0 := 0;$
  $v_0 := 0;$
  $i := 0;$
  while θ_t not converged **do**
    $i := i + 1$
    $g_i := \nabla_\theta f_i(\theta_{i-1})$
    $m_i := \gamma_1 \cdot m_{i-1} + (1 - \gamma_1) \cdot g_i$
    $v_t := \gamma_2 \cdot v_{i-1} + (1 - \gamma_2) \cdot g_i^2$
    $\hat{m}_i := m_i/(1 + \gamma_1^t)$
    $\hat{v}_i := v_i/(1 + \gamma_2^t)$
    $\theta_i := \theta_{i-1} - n \cdot \hat{m}_i/(\sqrt{\hat{v}_i} + \epsilon)$
  end while
  return θ_i

4 Experiment and results

4.1 Data and area

4.1.1 Data

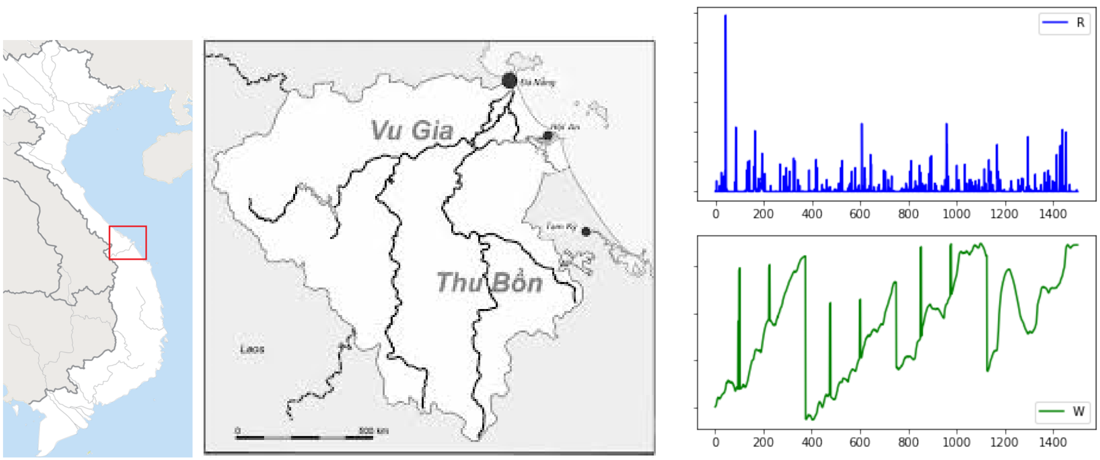

Fig. 3. The Vu Gia - Thu Bon river system (center Vietnam) and the 2 datasets at first station of Da Nang station. Chart (with R character) is rainfall and chart (with W character) is water level. Rainfall is measured by millimeter and water level is measured by centimeter

Because of the characterized mountain terrain in Central Vietnam, the rain falls straight to the rivers, directly affect the water level of those rivers. Therefore, the rainfall dataset is necessary for higher accuracy.

2 datasets have been used in this paper: water level data is used along with the rainfall data. Both datasets have been measured in rain season in Central Vietnam in 4 years from 2017 to 2020 at the station in Da Nang.

Every year, data on water levels is gathered between June 15 and September 15. Each data is collected with 2-hours interval.

Similar with the water levels data, rainfall data is also collected from June 15 to September 15 each year. But each data is collected every 6 hours.

Rainfall is measured by millimeter and water level is measured by centimeter.

4.1.2 Criteria for comparison

We used 2 types of value to evaluate the error and compare the results between the considered models: Mean Squared Error (MSE) and Mean Absolute Percentage Error (MAPE):

$$MSE = \frac{1}{n} \cdot \sum_{i=1}^{n} (\hat{y}_i - y_i)^2 \qquad MAPE = \sum_{i=1}^{n} \left| \frac{\hat{y}_i - \bar{\hat{y}}_i}{n} \right|$$

4.2 Model

In this model we will use the following algorithms:

- The WA model will use: N = 3 levels to analyze the data into High frequency and Low frequency components because the data has n ≈ 4000 data values. We will calculate the number of levels according to the following formula: $N = log_{10}(n)$
- The LSTM model is trained with 128 hidden layers, 10 batches, 10 eposh.
- We run the model on a computer with the configuration: Chip I7-10510U CPU 1.80GHz and RAM 8GB on Window 10 64bit.

4.3 Results

The given data is utilized for training of our proposed model along with the severan models: traditional RNN, traditional LSTM, LSTM with 2 input, LSTM-Adam and PSO-LSTM method presented by Jannatul Ferdous Ruma and his companions.

The Table 4 show the results of the models after we applied them to our data.

Table 4. Result of data

Model	Compare		Input	Cite	Note
	MSE	***MAPE***			
RNN	0.056110	0.088231	1 data input	[6]	
LSTM	0.054988	0.087021	1 data input	[15]	
LSTM	0.055096	0.086411	2 data input		
LSTM-Adam	0.043746	0.072289	2 data input		
PSO-LSTM	0.046021	0.075001	1 data input	[6]	*Rebuild*
WA-LSTM-Adam	**0.043668**	**0.072176**	2 data input		

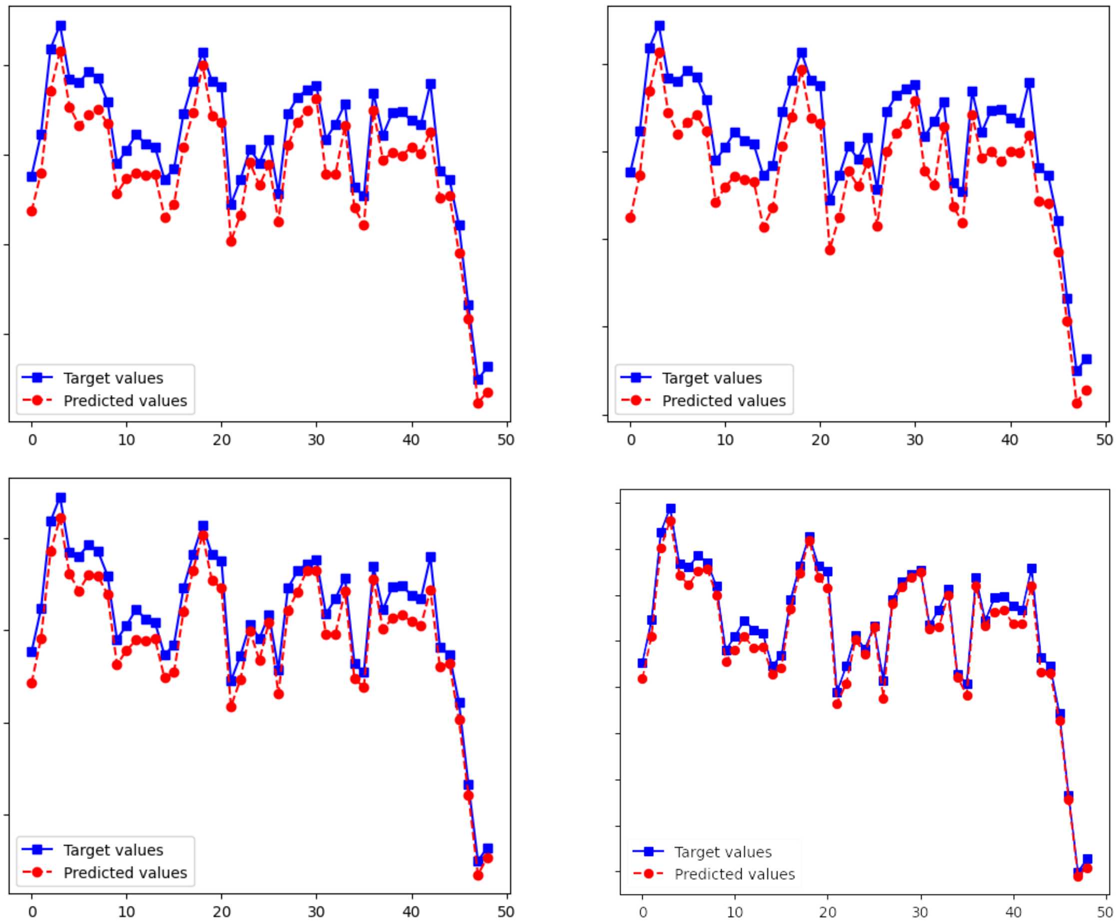

Fig. 4. Results graph of tested models for water level forecasting in Da Nang: LSTM (2 data input) (upper-lelf), LSTM-Adam (2 data input) (upper-right), LSTM-PSO by Ruma[6] (1 data input) and the proposed model with 2 inputs (lower-right)

The Table 4 had illustrated clearly that our WA-LSTM-Adam model with two inputs has the best accuracy among all tested models in this research. Based on our criteria to evaluate error, it seems that the original RNN method get the worst results: MSE=0.056110 and MAPE=0.088231. Furthermore, Ruma's model with only water level input had worse performance, with the MSE of 0.046021 and MAPE of 0.075001 than the WA-LSTM-Adam 2 inputs model (MSE=0.043668 and MAPE=0.072176).

5 Conclusion and Future Work

In this research, a new, different approach from Ruma's model is proposed when utilizing LSTM model on water level prediction, and based on some experiments, our 2 inputs WA-LSTM-Adam had outperformed that when using the given data. With the result, we can compare with other method to find the optimal option for each region and use it to minimize the damage caused by flood.

In the future, we will try to develop this model further with some other approach. Moreover, adding more inputs can be an option for further improvement.

Disclaimer (Artificial Intelligence)

Author(s) hereby declare that NO generative AI technologies such as Large Language Models (ChatGPT, COPILOT, etc) and text-to-image generators have been used during writing or editing of this manuscript.

Acknowledgement

This research was supported by HSGS (the Vietnam National University, Hanoi university of Science, High School for Gifted Students).

We want to thank of HSGS for their support during various stages while writing this paper.

Competing Interests

Authors have declared that no competing interests exist.

References

[1] Azizi E, Yosefvand F, Yaghoubi B, Izadbakhsh MA, Shabanlou S. Modelling and prediction of groundwater level using wavelet transform and machine learning methods: A case study for the Sahneh Plain, Iran. Irrigation and Drainage. 2023;72(3):747-762. DOI:10.1002/ird.2794

[2] Li Y, Qu J, Zhang H, Long Y, Li S. Water level prediction of Liuxihe Reservoir based on improved long short-term memory neural network. Water Supply. 10 2023;23(11):4563-4582. DOI:10.2166/ws.2023.282

[3] Wei A, Li X, Yan L, Wang Z, Yu X. Machine learning models combined with wavelet transform and phase space reconstruction for groundwater level forecasting. Computers Geosciences. 2023;177:105386. DOI:10.1016/j.cageo.2023.105386

[4] Nevo S, Morin E, Gerzi Rosenthal A, et al. Flood forecasting with machine learning models in an operational framework. Hydrology and Earth System Sciences. 2022;26(15):4013-4032. DOI:10.5194/hess-26-4013-2022

[5] Atashi V, Gorji HT, Shahabi SM, Kardan R, Lim YH. Water Level Forecasting Using Deep Learning Time-Series Analysis: A Case Study of Red River of the North. Water. 2022;14(12). DOI:10.3390/w14121971

[6] Ruma JF, Adnan MSG, Dewan A, Rahman RM. Particle swarm optimization based LSTM networks for water level forecasting: A case study on Bangladesh river network. Results in Engineering. 2023;17:100951. DOI:10.1016/j.rineng.2023.100951

[7] Kardhana H, Valerian JR, Rohmat FIW, Kusuma MSB. Improving Jakarta's Katulampa Barrage Extreme Water Level Prediction Using Satellite-Based Long Short-Term Memory (LSTM) Neural Networks. Water. 2022;14(9). DOI:10.3390/w14091469

[8] Cho M, Kim C, Jung K, Jung H. Water Level Prediction Model Applying a Long Short-Term Memory (LSTM)–Gated Recurrent Unit (GRU) Method for Flood Prediction. Water. 2022;14(14). DOI:10.3390/w14142221

[9] Azad AS, Sokkalingam R, Daud H, et al. Water Level Prediction through Hybrid SARIMA and ANN Models Based on Time Series Analysis: Red Hills Reservoir Case Study. Sustainability. 2022;14(3). DOI:10.3390/su14031843

[10] Wu J, Wang Z, Hu Y, Tao S, Dong J. Runoff Forecasting using Convolutional Neural Networks and optimized Bi-directional Long Short-term Memory. Water Resources Management. 2023;37(2):937-953. DOI:10.1007/s11269-022-03414-8

[11] Xu W, Chen J, Corzo G, et al. Coupling Deep Learning and Physically Based Hydrological Models for Monthly Streamflow Predictions. Water Resources Research. 2024;60(2):e2023WR035618. DOI:10.1029/2023WR035618

[12] Hochreiter S, Schmidhuber J. Long Short-Term Memory. Neural Computation. 1997;9(8):1735-1780.

[13] Pascanu R, Mikolov T, Bengio Y. On the difficulty of training recurrent neural networks. Published online 2013:1310-1318.

[14] Gers FA, Schmidhuber J, Cummins F. Learning to Forget: Continual Prediction with LSTM. Neural Computation. 2000;12(10):2451-2471. DOI:10.1162/089976600300015015

[15] Vizi Z, Batki B, Rátki L, et al. Water level prediction using long short-term memory neural network model for a lowland river: a case study on the Tisza River, Central Europe. Environmental Sciences Europe. 2023;35. DOI:10.1186/s12302-023-00796-3

Peer-review history:
The peer review history for this paper can be accessed here (Please copy paste the total link in your browser address bar)
https://prh.ikprress.org/review-history/12387