Check for updates

Multi-step-ahead water level forecasting for operating sluice gates in Hai Duong, Vietnam

Hung Viet Ho · Duc Hai Nguyen ·
Xuan-Hien Le [ORCID] **· Giha Lee**

Abstract Recently, machine learning (ML) is being applied to various fields, including hydrology and hydraulics. The numerical models based on ML algorithms have been widely used for forecasting water levels or flowrate in different timescales. Especially in estuary areas where the hydrodynamic regime becomes complicated, the water level forecast information in this area plays an essential role in the operation of tidal sluices. This study proposes an efficient approach using an ML model, long short-term memory (LSTM), to predict short-term water levels in tidal sluice gates from 6 to 48 hours ahead. The An Tho culvert located in the Bac Hung Hai irrigation system, the most extensive irrigation system in Vietnam, was selected as a case study station. The high accuracy of predictive results reveals LSTM models' effectiveness in different forecasting scenarios. In the first scenario using just water level data at the prediction station, the Kling–Gupta efficiency (KGE) coefficient ranges from nearly 0.89 to 0.96. Meanwhile, in the second scenario, the combination of observed data of three gauge stations exhibited better performance with KGE coefficients ranging from just under 0.93 to 0.98 for eight forecasted cases. The findings of this study highlight the performance of LSTM models in providing high-accuracy short-period water level forecasts for areas near estuaries. These obtained results can play a vital role in the management and operation of tidal sluices in the Bac Hung Hai irrigation system, as well as a reference for the operation of other irrigation systems around the world.

Introduction

In recent years, artificial intelligence and machine learning (ML) which have been on the rise as apparent proof of the fourth industrial revolution are applied to many different fields, including hydrology and hydraulics (Aghelpour et al., 2021; Le et al., 2020b; Shen, 2018; Sit et al., 2020). Numerous neural network models based on ML algorithms have been widely used to predict important characteristics of flows such as discharge, river water level (Ardabili et al., 2020; Phan & Nguyen, 2020; Yaseen et al., 2015), or flow in coastal estuaries (Granata & Di Nunno, 2021; Hidayat et al., 2014). In the agricultural sector, particularly irrigation systems, the forecasted

H. V. Ho · D. H. Nguyen · X.-H. Le (✉)
Faculty of Water Resources Engineering, Thuyloi University, 175 Tay Son, Dong Da District, Hanoi 10000, Vietnam
e-mail: hienlx@tlu.edu.vn

G. Lee
Department of Advanced Science and Technology Convergence, Kyungpook National University, 2559 Gyeongsang-daero, Sangju 37224, Gyeongsangbuk, South Korea

water level at the sluice gates is one of the crucial pieces of information to establish a real-time sluice operation rule toward the provide water for irrigation of crops in the agricultural field, as well as serving other fields.

Most of the commercially available models are numerical models where factors such as hydrodynamics and hydrology are of primary interest (Devia et al., 2015; Jaiswal et al., 2020). These models tend to exploit a large number of data types such as digital elevation model, water level, discharge, land cover, or precipitation as input data. In some instances, it may be challenging to collect these types of data fully, or they may not even be available in all locations (Eldho & Kulkarni, 2017; Le et al., 2019). Besides that, process-based methods have weaknesses in terms of real-time flow prediction because a significant amount of their time is spent interpreting input data (Thirumalaiah & Deo, 2000). An alternative approach is using models based on artificial neural network (ANN), especially ML models based on recurrent neural networks (RNN), for predicting river discharge and water level (Bai et al., 2021; Chen et al., 2012; Masrur Ahmed et al., 2021). ANN-derived models belong to the data-driven models, not the physics-based models, that have been studied for flood water level forecasting since the early 1990s (Cloke & Pappenberger, 2009; Young, 2002). A variety of studies have been performed suggesting that ANN-based models (or ML models) may offer a promising alternative to traditional process-based models for forecasting several of the specific hydrological properties of river flows (Adnan et al., 2021; Papacharalampous et al., 2019; Zounemat-Kermani et al., 2021). Furthermore, researchers and scholars have applied and combined various ML algorithms for the purpose of enhancing the hydrological prediction performance of data-driven models. (Nguyen & Bae, 2020; Ni et al., 2020; Yuan et al., 2018). Several studies on river water level prediction using ML algorithms are briefly reviewed below.

Mosavi et al. (2018) implemented a review of studies on flood forecasting using ML algorithms. This study identified key trends in exploiting data-driven models and provided an overview of the performance of various ML algorithms that are prevalently used in hydrology. Besides, the applications of ML models in general hydrological sciences in recent years are summarized in the study implemented by Xu and Liang

(2021). In addition to the aforementioned achievements, this paper also points out challenges in terms of physical interpretability and the influence of small sample sizes in hydrological applications. An application of ML algorithms, in particular the GRU model (belonging to the RNN family), has been studied by Le et al. (2021a) for short-time water level forecasting and applied to the Geum River, one of the largest river basins in South Korea. The Nash–Sutcliffe efficiency coefficient in this study is up to well above 99% in the case of one-hour prediction, which proves the effectiveness of RNN-based models when working with sequential data, for example, water level data. Another study on the long short-term memory (LSTM) neural network was successfully performed in the research of Yang et al. (2020) to predict the tidal level of harbors in Taiwan. This study also indicates the superiority of predictive performance of the LSTM model compared to other models such as ANN, support vector regression (SVR), and convolutional neural network (CNN). Several listed studies above have demonstrated the effectiveness of ML models in terms of water level prediction, especially LSTM or GRU models (RNN-based models).

In Vietnam, several recent studies on the application of ML models in streamflow forecasting have also been performed on the river basins. It can be mentioned as the case of Le et al. (2020a), a GRU model has been proposed to forecast river water levels on the Luoc River, Vietnam; Le and Ho (2018) applied the LSTM model to generate the water level prediction at the Quang Phuc and Cua Cam stations in Hai Phong, Vietnam, or research on the comprehensive evaluation of deep learning algorithms in streamflow prediction with a case study for the Red River basin carried out by Le et al. (2021b). The aforementioned studies indicate that LSTM models can be applied to predict the water level in rivers near the estuary in Vietnam. For coastal estuarine areas where the hydrodynamic regime is complicated by the interference between the river flow regime and the tidal regime, the forecasting of water levels at hydrological stations in these areas has not been significantly exploited.

In this study, we toward utilizing the LSTM architecture in predicting the water level multi-hour in advance for the tidal region to serve as a reference information for the management agencies in operating tidal sluices effectively. The tidal area selected

as the case study is situated in the lower delta of the Red-Thai Binh river system, Vietnam. Several LSTM models, a special kind of ML algorithm, have been developed with the aim of predicting multiple-timestep-ahead water levels at An Tho hydrological station, in particular from one to eight timesteps of lead time. The An Tho station (or tidal sluice) is under the direct management of the Bac Hung Hai irrigation system, the largest irrigation and drainage system in Vietnam and located in Hai Duong Province. Besides mitigating the negative impacts that may be caused by sea-level rise and excess salinity for agricultural crops, An Tho sluice gate is also designed to operate during the rainy season with the aim of flood reduction for Bac Hung Hai irrigation system with a maximum designed flood discharge capacity of 105 m^3/s (Le et al., 2020a). Hence, short-term water level forecast information (from 6 to 48 h in advance) at An Tho sluice gate plays a vital role in monitoring and paving the way for tidal sluice gate operation to ensure freshwater demand for agricultural activities of this area.

Material and methods

Study area

The study area is located in the lower part of the delta region of the Red-Thai Binh river basin, one of the largest river basin systems in Vietnam and ranked second nationwide. This area is also known as the northern coastal region of Vietnam, where the river system is adjacent to the estuary area. Therefore, the dynamic flow regime of tributaries in this area is relatively complicated due to the influence of the tidal regime. In this study, the selected target-forecast station is An Tho gauge station, or more precisely downstream of An Tho sluice gate. The An Tho sluice belongs to the management of the Bac Hung Hai irrigation system, which was constructed in 1958. This is a monitoring station with a crucial position because a few monitoring factors such as water level and salinity will determine the operation rule of the sluice gate to serve the irrigation purposes of the Bac Hung Hai system. In addition to the data at the An Tho sluice, the water level data at the Chanh Chu hydrological station on the Luoc River (next to the An Tho sluice) and the sea level data at the Hon Dau gauge station

(on the Hon Dau island) were also collected. The geographical location of the study site and the information of the gauge stations are depicted in Fig. 1 and Tables 1 and 2.

Water level data were collected from three gauge stations over 15 years, from January 2002 to December 2016. The observation timestep of the collected water level data is six hours, which means that the data information at the monitoring stations is observed every six hours. In addition, the collected water level data and the forecasted results in this study were measured in meters (m).

LSTM network

The proposed models are based on the LSTM neural networks, which belong to the category of RNN introduced by Hochreiter and Schmidhuber (1997), to solve multi-step time series forecasting problems. According to Olah (2015), all RNNs have the form of a chain of repeating memory blocks, and these blocks each have a simple structure, namely a single hyperbolic tangent layer. The LSTM models are also organized as a series of modules, but these interact in a special way because there are four layers instead of one in the standard RNN.

The typical structure of an LSTM network is an ordered arrangement of a sequence of memory blocks called cells. Accordingly, the output of each previous block of memory, consisting of two crucial components referred to as the cell state and the hidden state, is used as the input information of the current memory block. The cell state is also understood as a place to store vital long-term information (called long-term memory), while the hidden state is considered as a store of short-term information (called short-term memory) which is transmitted from the previous memory block. The main components of an LSTM memory block are depicted in Fig. 2.

At each given time step, the LSTM cells will interact with the input data through mechanical gates known as the forget gate layer, input gate layer, and output gate layer. In the stage of the forget gate layer, the data at the current time (x_t) and the short-term memory information from the previous LSTM cell (h_{t-1}) are the input data. The sigmoid activation function is operated in this mechanical gate, and output in this step (f_t) determines which input information from

(b) Lower delta of Red-Thai Binh River system   **(a)** Location of study area in Vietnam

(c) Location of gauge station in the study area
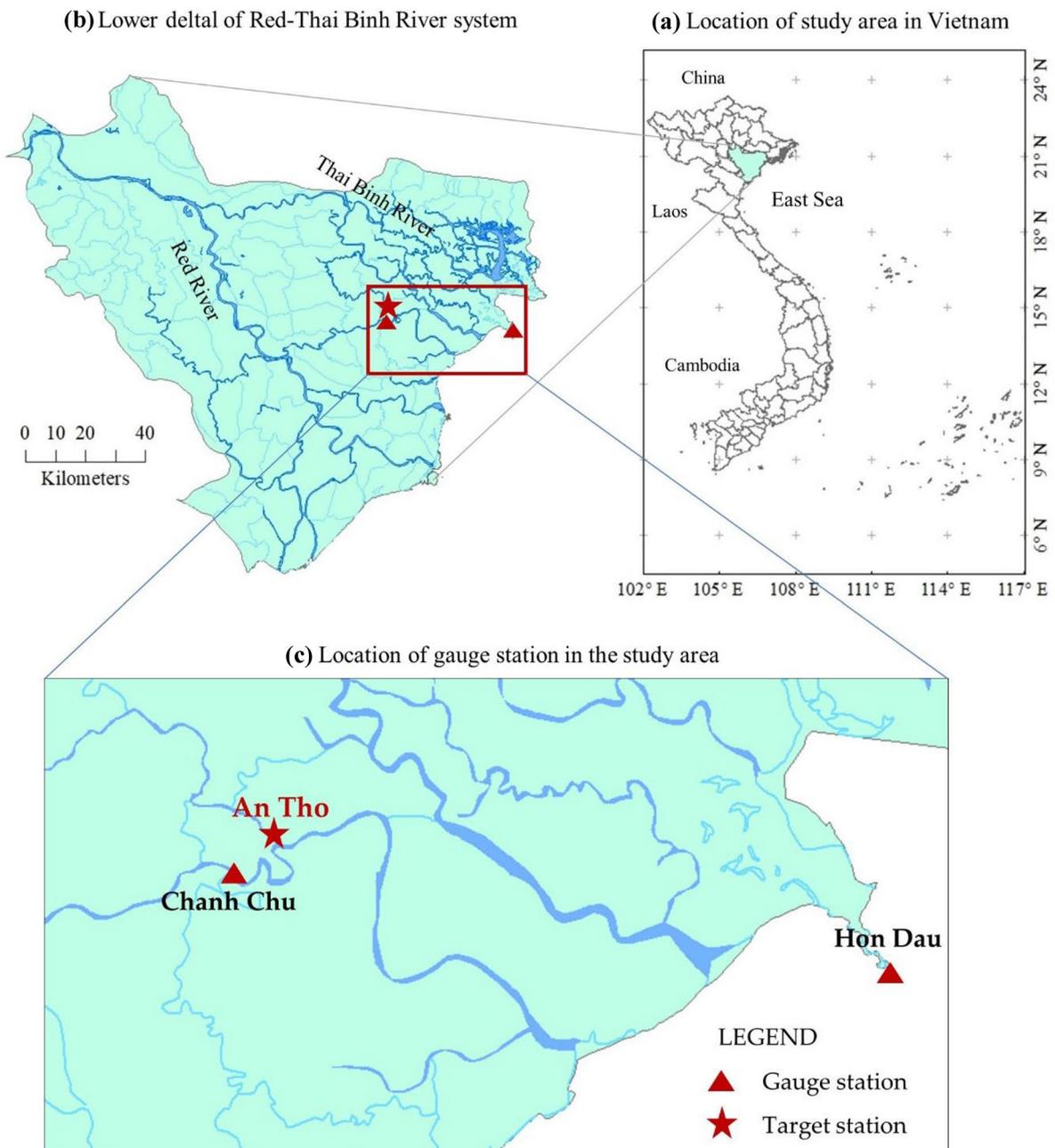
Fig. 1 Location information of gauge stations in the study area

the current time step should be omitted. In the next steps, long-term memory information or cell state will be stored and updated based on the influence of the input data. First, the sigmoid activation function in the input gate layer is applied to decide whether to update new information or not. Then, the hyperbolic tangent

(*tanh*) activation function evaluates the importance of this information. The new information of the long-term memory (C_t) will be updated from the previous cell state (C_{t-1}) based on the reported values of the sigmoid functions (f_t, i_t) and *tanh* function ($\overline{C_t}$). In the state of the output gate layer, the outcome value of the memory

Table 1 Information on the geographical location of the gauge stations

No	Stations	Item	Latitude	Longitude	Period	Note
1	Hon Dau	Water level	20.667	106.816	Jan 2002- Dec 2016	Sea level station
2	Chanh Chu	Water level	20.733	106.411	Jan 2002- Dec 2016	Hydrology station
3	An Tho*	Water level	20.749	106.455	Jan 2002- Dec 2016	Hydrology station

* denotes the target station

block (h_t) is calculated based on the combination of the outputs obtained from the current input via the sigmoid activation function (O_t) and the long-term memory through the *tanh* function. The outputs of each LSTM memory cell, including long-term memory (C_t) and short-term memory (h_t), are summarized by the following equations:

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

$$i_t = \sigma\left(W_i\left[h_{t-1}, x_t\right]\right) \tag{2}$$

$$\overline{C_t} = \tanh\left(W_c\left[h_{t-1}, x_t\right]\right) \tag{3}$$

$$C_t = f_t \otimes C_{t-1} + i_t \otimes \overline{C_t} \tag{4}$$

$$O_t = \sigma\left(W_o\left[h_{t-1}, x_t\right]\right) \tag{5}$$

$$h_t = O_t \otimes \tanh(C_t) \tag{6}$$

where σ is the sigmoid function; *tanh* is the hyperbolic tangent function; W_f, W_i W_o are the weight matrices of the forget gate layer, input gate layer, and output gate layer, respectively; C_t is the cell states at time t; and $\otimes$ represents the element-wise multiplication.

Model design

The ML modes have developed in this study based on the TensorFlow framework—an open-source software library provided by Google (Abadi et al., 2015).

Table 2 Correlation matrix of the collected data

Stations	Hon Dau	Chanh Chu	An Tho
Hon Dau	1	0.42	0.57
Chanh Chu	0.42	1	0.96
An Tho	0.57	0.96	1

Besides, Python (Rossum, 1995) is the dominant programming language throughout this research. Several available libraries such as Numpy, Pandas, and Matplotlib are exploited for data management and visualization (Hunter, 2007; McKinney, 2010; Van Der Walt et al., 2011).

Scenarios

For data-based predictive models, the characteristic of the observed dataset is one of the quantities that have a close influence on model performance besides geographical location and hydrodynamic factors. The input data are collected from three various stations, however, they belong to two different governing agencies. While An Tho station is under the management center of the Bac Hung Hai irrigation and drainage system, the other two measuring stations—Chanh Chu and Hon Dau—are under the management of the national hydro-meteorological center. On the other hand, the location data of the hydro-meteorological station network of Vietnam also indicate that Chanh Chu is the closest hydrological station and has a certain correlation to the An Tho sluice. In addition, the target station of the study area could be affected by several factors such as tidal regime and flow regulation of the Bac Hung Hai irrigation system (Le et al., 2020a). Therefore, two different scenarios of the LSTM model are proposed to predict the water level at An Tho gate in this paper. The data characteristics of interest comprise the length of the measured data series and the quantity of input data. These two scenarios are explained in detail as follows.

For the first scenario, the input data of the ML model only have information about the measured data at the target forecast station (downstream of An Tho station—x_t) with the frequency of monitoring every 6 h. The ML model is set up for training and can produce forecasts from 6 to 48 h in advance based on observed water level data at present and in the previous. The data operation structure of the first scenario is defined as follows:
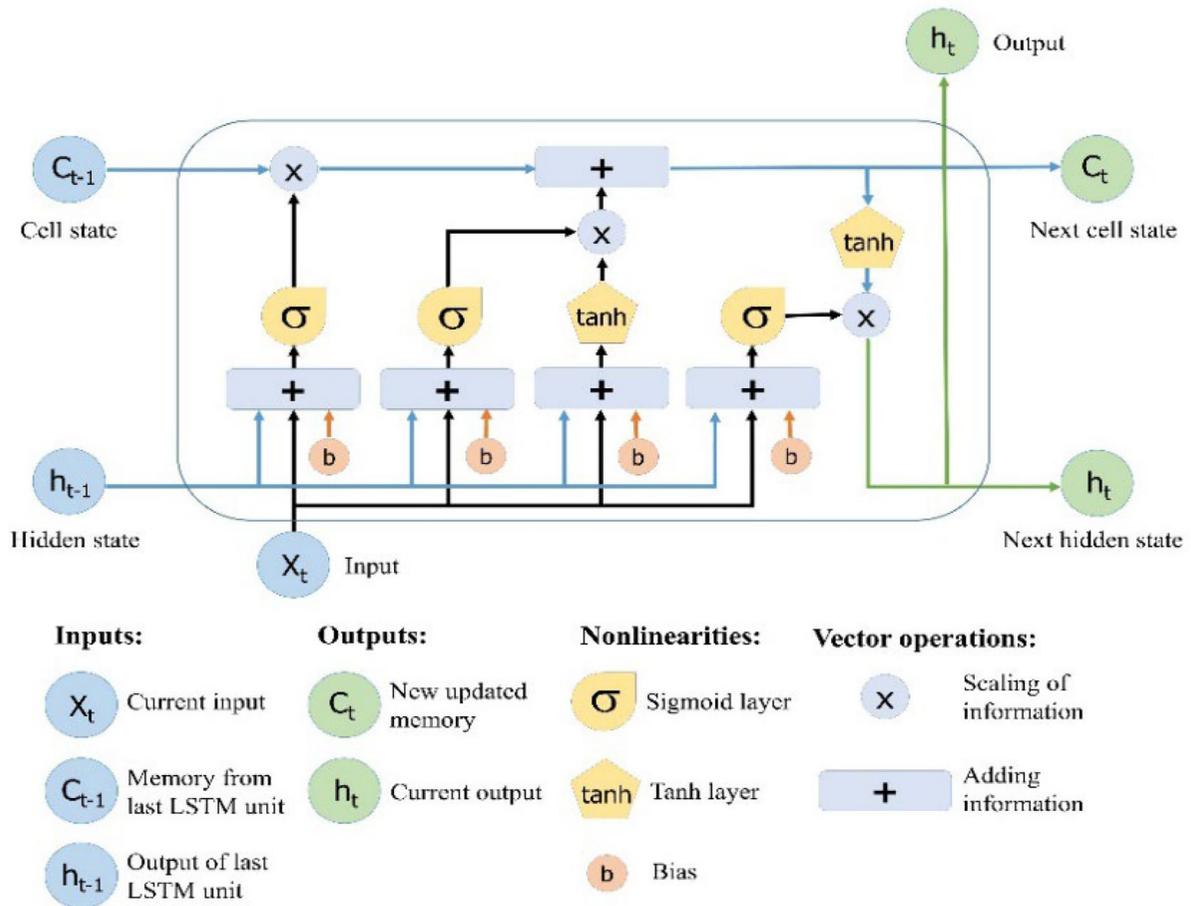
Fig. 2 Main components of an LSTM cell (Le et al., 2019)

$$\begin{bmatrix} x_t \\ x_{t-1} \\ ... \\ x_{t-n} \end{bmatrix} \rightarrow y_{t+i} \qquad (7)$$

In the second scenario, we considered water levels measured at three stations, including the target station (Hon Dau—x^1, Chanh Chu—x^2, downstream of the An Tho sluices—x^3) as the input information of the model. The operation structure of the second scenario in vector form is as follows:

$$\begin{bmatrix} x_t^1 & x_t^2 & x_t^3 \\ x_{t-1}^1 & x_{t-1}^2 & x_{t-1}^3 \\ & ... & \\ x_{t-n}^1 & x_{t-n}^2 & x_{t-n}^3 \end{bmatrix} \rightarrow y_{t+i} \qquad (8)$$

In the above formulas, x is the input data variable, y is the output data variable, t is the current time, n is mentioned as the number of timesteps in the past, and i is the term to refer to the number of forecasted timesteps ahead (the value of i in this study ranges from 1 to 8, corresponding to the forecasted time from 6 to 48 h in advance).

In order to be able to evaluate the performance of the predictive models objectively, the available dataset was divided into three different parts for training, validation, and testing purposes, respectively. The first dataset used for training purposes consists of a 14-year observations series from 01 January 2002 to 31 December 2013. The second dataset from 01 January 2014 to 30 November 2016 (approximately two years of data) is exploited to

validate model performance based on the selected parameters from the training process. The remaining part is a dataset of observed values in December 2016 that is applied to assess the forecasting model performance objectively.

Parameter setting

Each LSTM model was trained, validated, and tested only to forecast the water level at one particular lead time. Several training options and model hyperparameters have been varied to ensure that the model yields high stability and efficiency. Accordingly, the parameters of the number of hidden layers and the number of units per hidden layer are determined based on the trial process because there are no specific guideline documents for choosing these values. The number of hidden layers and the number of units recommended for this study are 1 and 16, respectively. This is because, according to Le et al. (2021a), increasing the number of hidden layers as well as the number of units per hidden layer may not improve model performance, especially for streamflow prediction problems. For this study, research experiments also reveal that changing the above-suggested values does not significantly increase the model performance. Meanwhile, complicating the architecture of predictive models can consume significant training time for ML models and in certain cases can even lead to over-fitting problems (Chollet, 2017; LeCun et al., 2015).

With respect to the optimization process, the Adam algorithm with the default learning rate of 0.001 is recommended for this study. This can be understood by the Adam is one of the most preferred optimization algorithms by scholars and researchers because of its effectiveness not only in the field of computer science (Kingma & Ba, 2014; Ruder, 2016) but also in hydrological applications (Le et al., 2020b, 2021b). The optimization function and the learning rate remain constant in all forecast cases corresponding to the two proposed scenarios. In addition, the proposed model is set up for training and validation processes with a maximum number of epochs of 50,000 for all forecasted cases, and the necessary information of the model is recorded until the training process ends. Besides, the computation process will be stopped when the error value of the loss function on

the validation data set does not decrease after 1,000 epochs (early stopping technique).

Evaluation criteria

The effectiveness of predictive models is assessed in terms of the root mean square error (RMSE), Nash–Sutcliffe efficiency (NSE) (Nash & Sutcliffe, 1970), and Kling—Gupta efficiency (KGE) (Gupta et al., 2009). These criteria are applied to compare the predicted value series and the observed value series. The proposed model presents a good result when the RMSE value is small (close to 0); the NSE and KGE values are approximately 1.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(O_i - P_i\right)^2} \qquad (9)$$

$$NSE = 1 - \frac{\sum_{i=1}^{n}\left(O_i - P_i\right)^2}{\sum_{i=1}^{n}\left(O_i - \overline{O_i}\right)^2} \qquad (10)$$

where O_i and P_i are observed and predicted values at time t; $\overline{O_i}$ is the mean of observed values; and n is the total number of observations.

$$KGE = 1 - \sqrt{(r-1)^2 + \left(\frac{\sigma_{sim}}{\sigma_{obs}} - 1\right)^2 + \left(\frac{\mu_{sim}}{\mu_{obs}} - 1\right)^2} \qquad (11)$$

where r is the linear correlation between simulation and observations; σ_{sim} and σ_{obs} are the standard deviation of simulations and observations; μ_{sim} and μ_{obs} are the simulation mean and observation mean, respectively.

Results and discussion

Validation phase

The proposed ML models corresponding to the two recommendation scenarios were validated on previously unseen data. The qualitative and quantitative performance of these models was determined using the criteria as RMSE, NSE, and KGE. After the training and validating processes were finished, the best

Table 3 Statistics on model parameters and performance in the validation period

Forecasted time (hours)	1st scenario				2nd scenario			
	Number of epochs	RMSE (m)	NSE	KGE	Number of epochs	RMSE (m)	NSE	KGE
6	35,087	0.133	0.929	0.965	31,889	0.104	0.957	0.971
12	34,929	0.144	0.917	0.947	33,885	0.123	0.939	0.966
18	13,414	0.148	0.912	0.934	41,913	0.130	0.933	0.950
24	28,677	0.146	0.915	0.938	35,898	0.130	0.932	0.959
30	24,890	0.174	0.879	0.920	13,512	0.148	0.913	0.945
36	47,885	0.186	0.862	0.919	36,728	0.158	0.900	0.937
42	16,200	0.194	0.850	0.900	24,952	0.168	0.887	0.925
48	30,442	0.194	0.850	0.890	26,697	0.173	0.881	0.925

version of identified predictive models regarding the epoch parameter is summarized in Table 3 and Figs. 3, 4, 5.

The figures in Table 3 demonstrate that the proposed LSTM models can provide impressive results for both forecasting scenarios. The difference in the forecast performance of the two scenarios in the prediction cases from 6 to 48 h ahead is about 0.3% on average, with the higher value belonging to the second scenario. This result indicates that

combining with water level data from Chanh Chu and Hon Dau stations can produce better performance than using just target station water level data (at the An Tho culvert).

In addition, the accuracy of forecasting models has a tendency to reduce as the forecasted time increases, which can be clearly identified through the NSE and KGE values (see Figs. 3 and 4). The KGE and NSE values express a similar development in the forecast cases. Although the KGE index is improved based on

(a) RMSE   **(b)** KGE
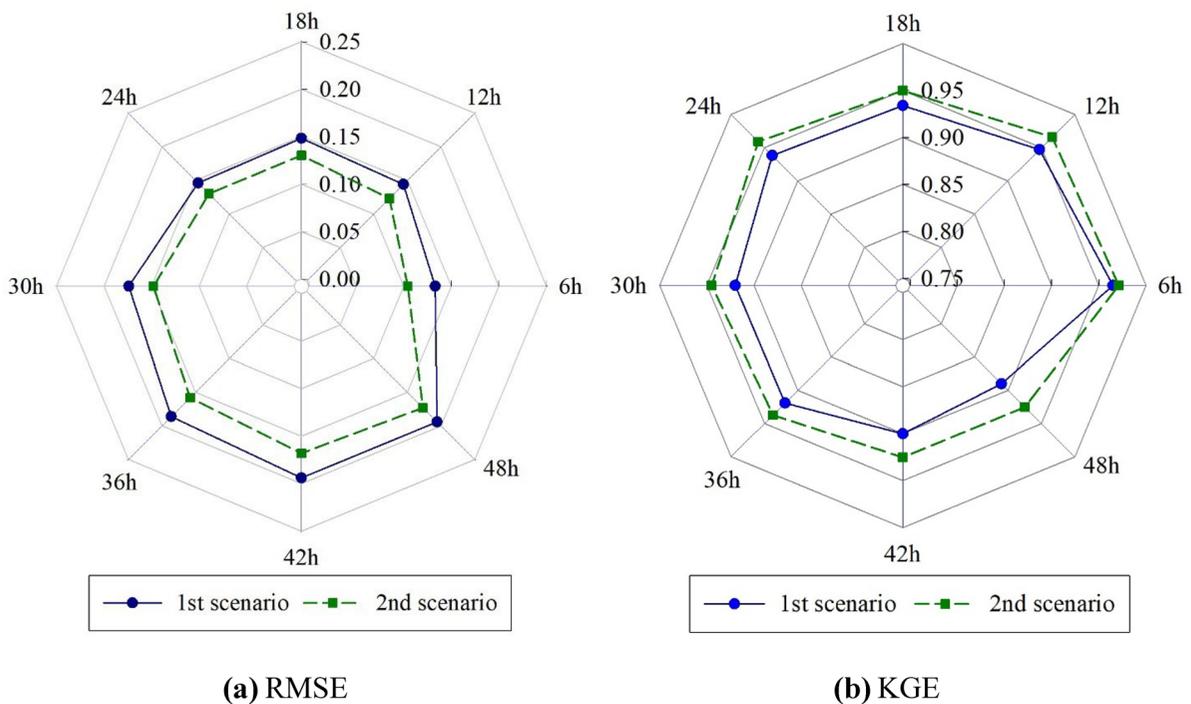
Fig. 3 Radar chart of two prediction scenarios in the validation period for **a** RMSE and **b** KGE
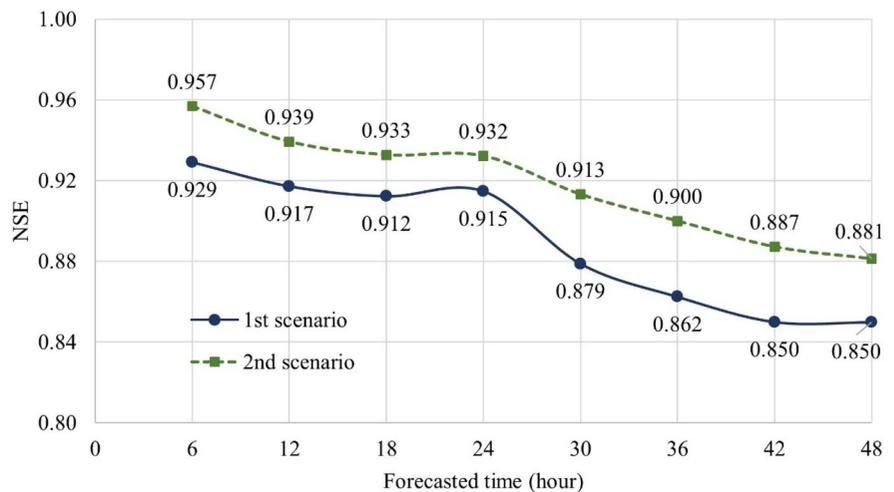
the NSE index and is expected to become an evaluation criterion in hydrological problems instead of using NSE, the effectiveness of KGE was not clarity demonstrated in this study as the calculated results indicated that the KGE values were higher than the NSE values in all reported cases.

In the first scenario, where the input data are just the water level data at An Tho culvert, the efficiency of eight forecast cases ranges from 85 to 93% (with NSE index), and the RMSE values are lower than 0.194 m. The range corresponding to the KGE coefficient in this case is from 89% to just under 97%. A similar trend was identified in the second scenario, where the performance of the predictive model ranges from 88% to nearly 96% and from roughly 93% to 97% for the NSE and KGE coefficients, respectively. In this scenario, the RMSE values are lower than 0.173 m. These errors are acceptable for both scenarios.

Figure 3 visually compares the water level forecast results of eight specific cases corresponding to the two proposed scenarios via the radar chart. In the event of prediction 6 h in advance (or one timestep ahead), the root mean square error of the second scenario (RMSE) on the whole validation dataset is 0.104 m, which is smaller than those of the first scenario of 0.133 m. Figure 5 depicts the scatter plot for the eight forecast cases of the second scenario. It can be seen that the predicted values and the observed values describe a high correlation in most cases. However, there are still a few points where the observed values are higher than the forecasted value

(some points are below and away from the 45^0 lines). This may be one of the reasons for the reduced performance of the ML models. Additionally, Table 3 provides information on the number of epochs of suggested models. Despite the fact that the predictive models are set up with a maximum number of computation iterations of 100,000 times, their convergence speed is not the same. The application of the early stopping technique is the cause of the difference in the epoch values in Table 3.

Testing phase

Two scenarios of the ML models were assessed by using a testing dataset with the aim of objectively verifying the predictive performance of the models and their accuracy. The results of the testing process for eight forecast cases using two different scenarios are reported in Table 4. Additionally, Figs. 6 and 7 illustrate a comparison of RMSE and NSE values between two scenarios in the testing step.

In general, a similarity is noted based on the figures in Tables 3 and 4, which describe the results of the validation and testing phases. The numbers in Table 4 demonstrate insignificant differences in RMSE, NSE, and KGE values during the testing phase in comparison with the validating one. The test results are slightly better than the validation results, which certify that the proposed models can produce stable, highly accurate results in both phases. Besides, the second scenario still outperforms the first scenario in providing forecasts, with higher NSE values and
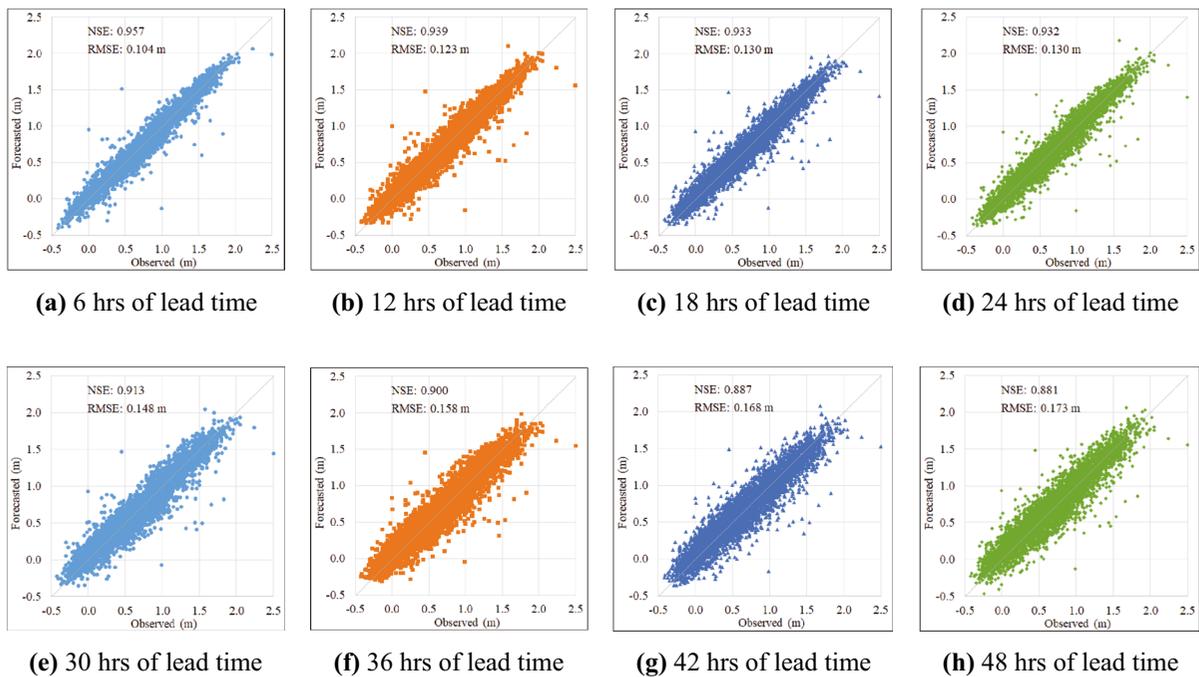
Fig. 5 Scatter plot of 2nd scenario in the validation period for forecasting: **a** one timestep, **b** two timesteps, **c** three timesteps, **d** four timesteps, **e** five timesteps, **f** six timesteps, **g** seven timesteps, and **h** eight timesteps ahead

lower RMSE values. The average difference in the performance of eight prediction cases between these two scenarios is about 3% for the NSE coefficient and about 2–4% for the KGE coefficient.

With respect to the second scenario, the KGE coefficient fluctuates from 0.98 to 0.926 for eight forecast cases, from 6 to 48 h ahead. The NSE coefficient for this scenario is always greater than 0.88. In contrast, the RMSE error value increased from 0.09 m to 0.18 m, corresponding to these cases. Figure 6 illustrates a similar trend in the graph of the RMSE and NSE values during the testing period as compared

with the validation phase (Fig. 4). The calculation results from the testing phase confirmed again that the prediction accuracy in the second scenario, which utilizes three input data series, is significantly higher than in the first scenario, which is just mining an input data series.

In addition, the information on the graph of the NSE coefficient in Fig. 6b reveals that the difference in the performance of the two scenarios has a tendency to increase as the forecasted time increases. This trend is noted both during the validation phase (Fig. 4) and the testing phase (Fig. 6). In the case of

Table 4 Performance of ML models in the testing process

Forecasted time (hours)	1st scenario			2nd scenario		
	RMSE (m)	NSE	KGE	RMSE (m)	NSE	KGE
6	0.127	0.945	0.961	0.089	0.973	0.980
12	0.143	0.930	0.946	0.118	0.953	0.969
18	0.144	0.929	0.942	0.119	0.952	0.959
24	0.146	0.927	0.940	0.122	0.949	0.955
30	0.184	0.883	0.919	0.150	0.922	0.952
36	0.195	0.867	0.912	0.166	0.904	0.906
42	0.205	0.854	0.892	0.172	0.897	0.913
48	0.214	0.840	0.887	0.183	0.884	0.926
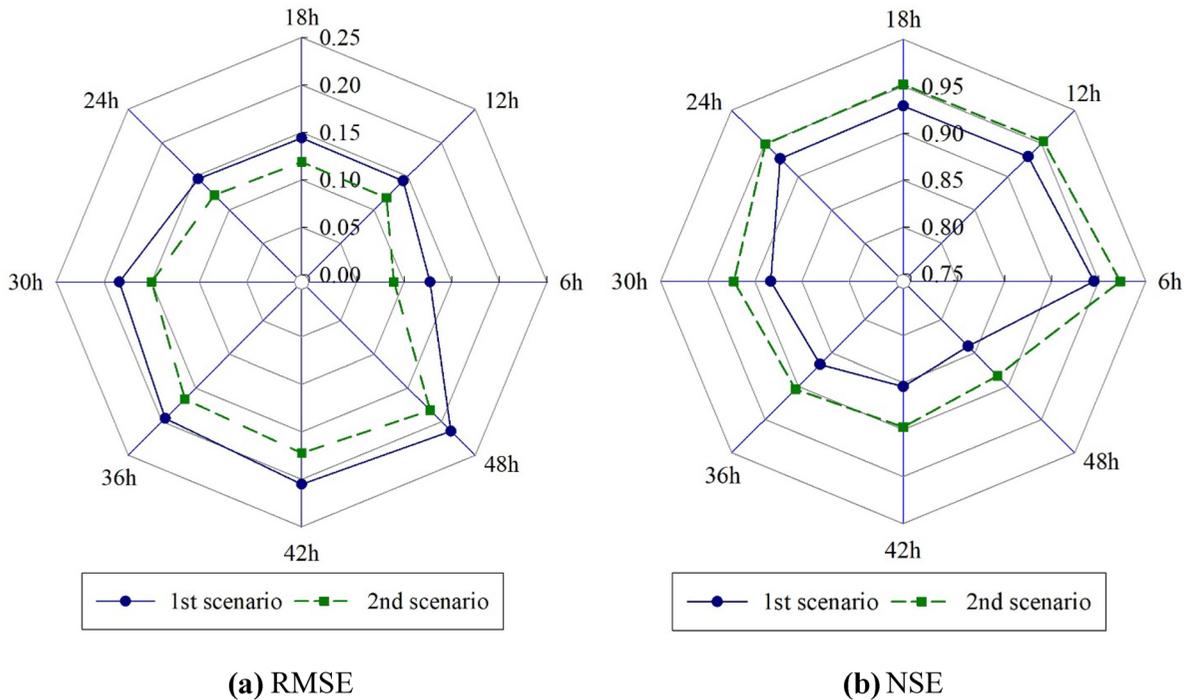
(a) RMSE

(b) NSE

Fig. 6 Radar chart of two prediction scenarios in the testing period for **a** RMSE and **b** NSE

one-time-step forecasting (or 6 h of lead time), the performance difference between the two scenarios is only 2.8%; however, this value has risen to 4.4%

when the model produces a forecast of eight-time-step (48 h ahead). The NSE values of the two scenarios in the case of forecasting the water level one

(a) 6 hrs of lead time

(b) 12 hrs of lead time

(c) 18 hrs of lead time

(d) 24 hrs of lead time

(e) 30 hrs of lead time

(f) 36 hrs of lead time

(g) 42 hrs of lead time

(h) 48 hrs of lead time
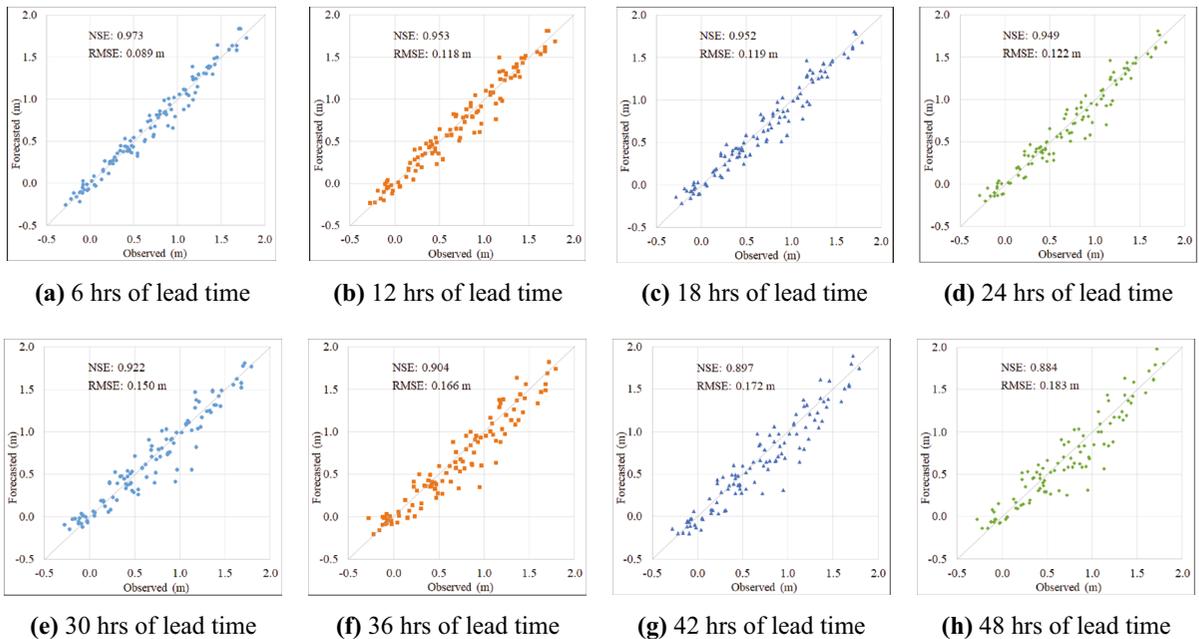
Fig. 7 Scatter plot of 2nd scenario in the test period for forecasting: **a** one timestep, **b** two timesteps, **c** three timesteps, **d** four timesteps, **e** five timesteps, **f** six timesteps, **g** seven timesteps, and **h** eight timesteps ahead

day ahead (or 24 h in advance) reached values of 94.9% and 92.7% and reached values of 88.4% and 84% when forecasting two days in advance (48 h), respectively. The RMSE values in the first scenario are higher than in the second scenario, but they are still less than 0.214 m for all forecast cases. Besides, paired data pairs are located close to the 45^0 lines exhibiting a high correlation between the predicted and observed values in the eight forecast cases (see Fig. 7). Regarding the second scenario, an evaluation in terms of temporal correlation between the forecasted values and the observed values is performed and summarized in Figs. 8, 9, 10.

Figures 7–10 clearly illustrate the effectiveness of the proposed LSTM models and the results on the testing dataset to express a high agreement between the outcome of the model and the observation data. Furthermore, the LSTM models could predict peak water level (the highest water level of the day) with remarkable accuracy in both magnitude and occurrence time. The graphs in Figs. 8, 9, 10 indicate that the predicted peak water level is higher than the actual water level measured at 7 am on 16 December 2016 from 0.04 m to 0.26 m depending on the forecast case; the forecasted time of occurrence of the maximum and minimum water levels coincides with the actual time, in all forecast cases.

There are certain cases where the predicted values from the model do not satisfy the observed value.

One of the possible explanations for this is that the streamflow downstream of An Tho culvert is simultaneously influenced by the river hydrodynamic and tidal regimes. In addition, it is found that Chanh Chu station affects the forecast results more than Hon Dau station because Chanh Chu station is closer to An Tho culvert than Hon Dau station. However, when forecasting long-term, Hon Dau station can promote its effects. From the results of this study, we can see that the data series with a measurement time of every 6 h per day is the most effective for prediction because it allows forecasting 6 h in advance and ensures the accuracy of long-term forecasting results.

Comparisons with the literature

With the advancement of modern technology, there have been several studies on using ML algorithms in water level forecasting. It is worth noting that Nguyen et al. (2022) established an hourly water level forecasting model for the Jungrang metropolitan region, which is located on the Han River in Korea. When compared to other ML models such as the multiple linear regression model or SVR model, the integration of the genetic algorithm with the Bayesian additive regression tree model boosts the water level prediction efficiency. However, as the forecasted period rises, the efficiency of the suggested model tends to
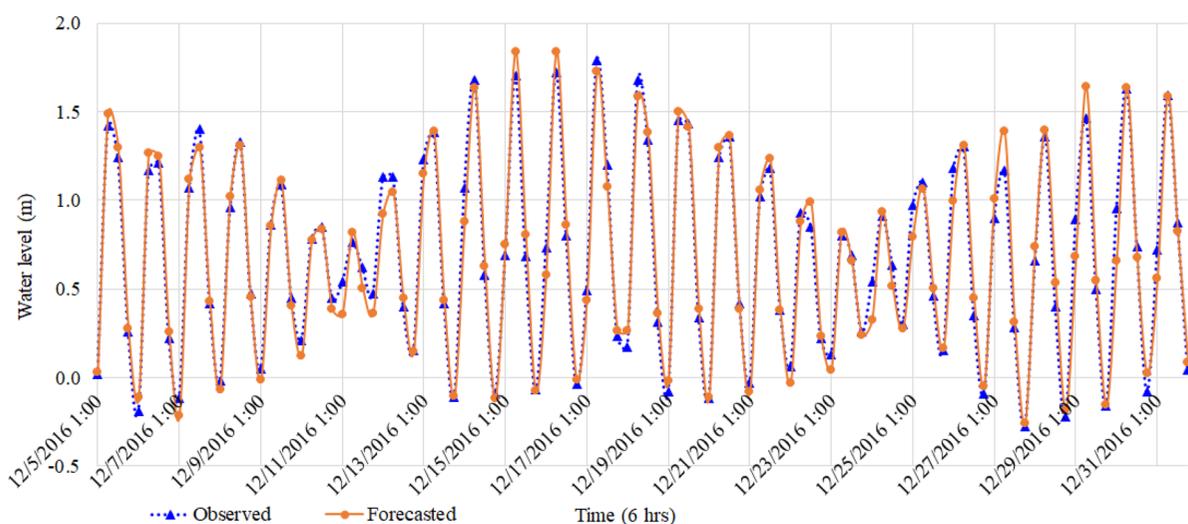
Fig. 8 Comparison of forecasted results and observed data for the 6-h prediction case of the second scenario during the test period
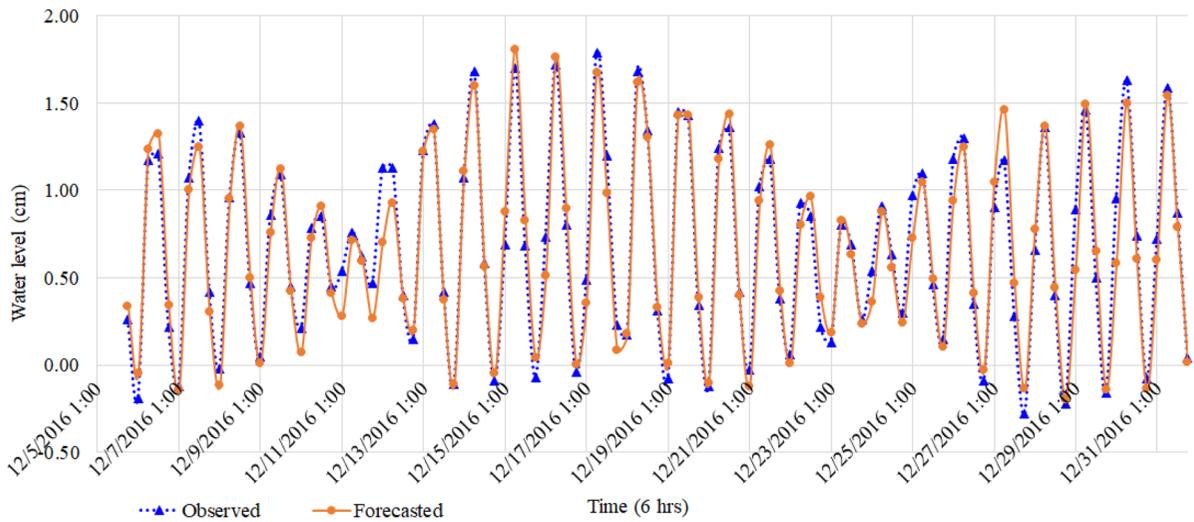
Fig. 9 Comparison of forecasted results and observed data for the 24-h prediction case of the second scenario during the test period

decline considerably; the corresponding NSE coefficient in this example drops from 0.96 when predicting one timestep ahead to just 0.47 when forecasting six timesteps in advance. In another investigation, Nguyen et al. (2021) pointed out that the XGBoot algorithm proved to be superior to tree-based models such as random forests or classification and regression trees. The NSE value in this scenario varies between [0.97–0.66], matching water level forecasts from one hour to six timesteps ahead (six hours). In addition, Yang et al. (2020) examined the efficacy of

ML models in predicting hourly tidal water levels in Taiwanese ports. According to this study, the LSTM model has the lowest prediction error among the other approaches listed, including ARIMA, SVR, ANN, and CNN. Another model belongs to the RNN family, Gated Recurrent Unit—GRU, which was also examined by Le et al. (2020a) in predicting water levels at tidal-affected sites. Even though it only uses data from the target station, this model can offer forecasts with high accuracy of up to [94%-96%] when predicting up to a 24-h lead time.
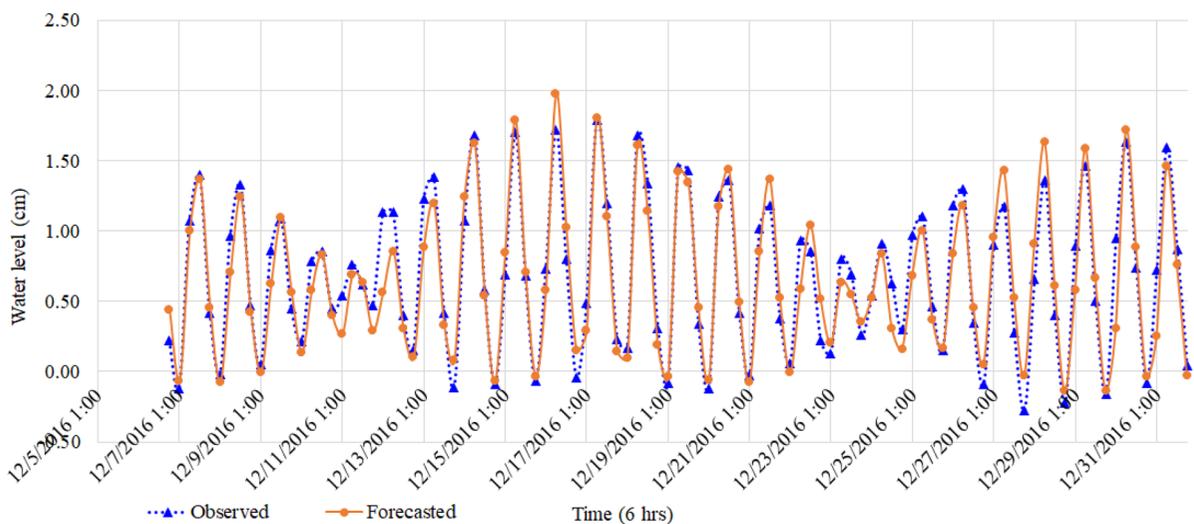
Fig. 10 Comparison of forecasted results and observed data for the 48-h prediction case of the second scenario during the test period

The majority of the aforementioned studies demonstrate the great capacity of ML (particularly LSTM-based) in providing high-accuracy short-term water level forecasts for rivers or tides only (Granata & Di Nunno, 2021; Kao et al., 2020; Li et al., 2020). However, in fact, the interference between river flow and tidal regimes is fairly prevalent in coastal areas where there is a direct influence on the agricultural industry. In the meanwhile, there haven't been many studies on the effect of input variables on model performance, as well as the development of an opening-closing procedure to minimize saline intrusion into the field. This study was conducted to address the above-mentioned issue, and the efficiency of the LSTM model in forecasting the water level in the tidal-affected sites has been verified. The NSE coefficients varied from [0.97–0.88] for prediction events from one to eight timesteps in advance, which was higher than in previous investigations. Furthermore, the significance of the input variables that directly impact predictability is thoroughly reviewed. This work has enriched solutions for high-accuracy flow forecasting for irrigation sluices in the intertidal zone.

Conclusions

In this paper, several LSTM models based on ML algorithms have been established with the aim of forecasting river water levels in tidal affected areas. From the above analyses, both scenarios illustrate the similarity of the forecast results to the observed data in the validation and testing phases. The prediction accuracy tends to reduce when forecasting multiple timesteps in advance slightly. In which, the first scenario with just exploiting water level information at the target station (An Tho culvert) has slightly lower efficiency than the second scenario using a combination of the water level data at the target station and two other stations upstream and downstream from the target station. This means that the increase in the number of input data series results in a more accurate forecast and vice versa.

On the other hand, for tidal rivers, the duration of the observed data series has a significant influence on model performance in addition to the quantity of the collected data. In several special cases when the number of input data series is limited, but the collected data series is long enough, the recorded data of past water levels at the target station can be utilized as input of the ML models to forecast the water level for itself at the next time-steps. This means that it is possible to use the observed water level data at a gauge station to provide water level predictions at this station.

Additionally, rainfall-runoff models with the advantage of simulating and interpreting flows on the river often face difficulties in short-term forecasting. Furthermore, these models have not exhibited their advantages in places where there is an interference of flow regimes, such as coastal estuaries. Therefore, ML models can completely be applied as an alternative approach to being able to make short-term estimations with high accuracy in these areas.

Although it is possible to produce short-term water level forecasts with high efficiency, ML models in this study still have disadvantages, such as providing estimates for only a limited number of stations and not simulating the physical process of river flow. Thus, studies on the combination between ML models and physical models should be of more interest.

This work is the first step that aims to improve our knowledge of applying ML algorithms in multi-timestep-ahead water level forecasting for tidal-affected areas. The present results can contribute to the management and operation procedures of sluice gates in the Bac Hung Hai irrigation and drainage system in Vietnam, as well as being applicable to other intertidal zones in the world.

Declarations

References

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., et al. (2015). TensorFlow: Large-scale machine learning on heterogeneous distributed systems. ArXiv:abs/1603.04467

Adnan, R. M., Petroselli, A., Heddam, S., Santos, C. A. G., & Kisi, O. (2021). Short term rainfall-runoff modelling using several machine learning methods and a conceptual event-based model. *Stochastic Environmental Research and Risk Assessment, 35*(3), 597–616. https://doi.org/10.1007/s00477-020-01910-0

Aghelpour, P., Bahrami-Pichaghchi, H., & Varshavian, V. (2021). Hydrological drought forecasting using multi-scalar streamflow drought index, stochastic models and machine learning approaches, in northern Iran. *Stochastic Environmental Research and Risk Assessment, 35*(8), 1615–1635. https://doi.org/10.1007/s00477-020-01949-z

Ardabili, S., Mosavi, A., Dehghani, M., Várkonyi-Kóczy, A. R. (2020). Deep Learning and Machine Learning in Hydrological Processes Climate Change and Earth Systems a Systematic Review, Cham; pp. 52–62. https://doi.org/10.1007/978-3-030-36841-8_5

Bai, P., Liu, X., & Xie, J. (2021). Simulating runoff under changing climatic conditions: A comparison of the long short-term memory network with two conceptual hydrologic models. *Journal of Hydrology, 592*, 125779. https://doi.org/10.1016/j.jhydrol.2020.125779

Chen, W.-B., Liu, W. C., & Hsu, M. H. (2012). Comparison of ANN approach with 2D and 3D hydrodynamic models for simulating estuary water stage. *Advances in Engineering Software, 45*(1), 69–79. https://doi.org/10.1016/j.advengsoft.2011.09.018

Chollet, F. (2017). *Deep learning with python*. Manning Publications.

Cloke, H. L., & Pappenberger, F. (2009). Ensemble flood forecasting: A review. *Journal of Hydrology, 375*(3), 613–626. https://doi.org/10.1016/j.jhydrol.2009.06.005

Devia, G. K., Ganasri, B. P., & Dwarakish, G. S. (2015). A Review on Hydrological Models. *Aquatic Procedia, 4*, 1001–1007. https://doi.org/10.1016/j.aqpro.2015.02.126

Eldho, T. I., & Kulkarni, A. T. (2017). Conceptual and Physically Based Hydrological Modeling, Sustainable Water Resources Management. 81–118. https://doi.org/10.1061/9780784414767.ch04

Granata, F., & Di Nunno, F. (2021). Artificial Intelligence models for prediction of the tide level in Venice. *Stochastic Environmental Research and Risk Assessment*. https://doi.org/10.1007/s00477-021-02018-9

Gupta, H. V., Kling, H., Yilmaz, K. K., & Martinez, G. F. (2009). Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *Journal of Hydrology, 377*(1), 80–91. https://doi.org/10.1016/j.jhydrol.2009.08.003

Hidayat, H., Hoitink, A. J. F., Sassi, M. G., & Torfs, P. J. J. F. (2014). Prediction of discharge in a tidal river using artificial neural networks. *Journal of Hydrologic Engineering, 19*(8), 04014006. https://doi.org/10.1061/(ASCE)HE.1943-5584.0000970

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

Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *Computing in Science & Engineering, 9*(3), 90–95. https://doi.org/10.1109/MCSE.2007.55

Jaiswal, R. K., Ali, S., & Bharti, B. (2020). Comparative evaluation of conceptual and physical rainfall–runoff models. *Applied Water Science, 10*(1), 48. https://doi.org/10.1007/s13201-019-1122-6

Kao, I. F., Zhou, Y., Chang, L. C., & Chang, F. J. (2020). Exploring a Long Short-Term Memory based Encoder-Decoder framework for multi-step-ahead flood forecasting. *Journal of Hydrology, 583,* 124631. https://doi.org/10.1016/j.jhydrol.2020.124631

Kingma, D. P., Ba, J. (2014). Adam: A method for stochastic optimization. ArXiv:abs/1412.6980

Le, X. H., Ho, H. V., & Lee, G. (2020a). Application of Gated Recurrent Unit (GRU) Network for Forecasting River Water Levels Affected by Tides, In Proceedings of APAC 2019, Hanoi, Vietnam; pp. 673–680. https://doi.org/10.1007/978-981-15-0291-0_92

Le, X. H., Ho, H. V., Lee, G., & Jung, S. (2019). Application of Long Short-Term Memory (LSTM) Neural Network for Flood Forecasting. *Water, 11*(7), 1387. https://doi.org/10.3390/w11071387

Le, X. H., & Ho, V. H. (2018). Using long short-term memory neural network to forecast water level at the Quang Phuc and the Cua Cam stations in Hai Phong Vietnam. *Journal of Water Resources & Environmental Engineering, 62*, 9–16.

Le, X.H., Jung, S., Yeon, M., & Lee, G. (2021a). River Water Level Prediction Based on Deep Learning: Case Study on the Geum River, South Korea, In Proceedings of Lecture Notes in Civil Engineering, Singapore. 319–325. https://doi.org/10.1007/978-981-16-0053-1_40

Le, X. H., Lee, G., Jung, K., An, H.-U., Lee, S., & Jung, Y. (2020b). Application of Convolutional Neural Network for Spatiotemporal Bias Correction of Daily Satellite-Based Precipitation. *Remote Sensing, 12*(17), 2731. https://doi.org/10.3390/rs12172731

Le, X. H., Nguyen, D. H., Jung, S., Yeon, M., & Lee, G. (2021b). Comparison of Deep Learning Techniques for River Streamflow Forecasting. *IEEE Access, 9*, 71805–71820. https://doi.org/10.1109/ACCESS.2021.3077703

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature, 521*, 436–444. https://doi.org/10.1038/nature14539

Li, Y., Shi, H., & Liu, H. (2020). A hybrid model for river water level forecasting: Cases of Xiangjiang River and Yuanjiang River China. *Journal of Hydrology, 587*, 124934. https://doi.org/10.1016/j.jhydrol.2020.124934

Masrur Ahmed, A. A., Deo, R. C., Feng, Q., Ghahramani, A., Raj, N., Yin, Z., et al. (2021). Deep learning hybrid model with Boruta-Random forest optimiser algorithm for streamflow forecasting with climate mode indices, rainfall, and periodicity. *Journal of Hydrology, 599*, 126350. https://doi.org/10.1016/j.jhydrol.2021.126350

McKinney, W. (2010). Data structures for statistical computing in Python, In Proceedings of 9th Python in Science Conference, Austin, TX, USA, 28 June – 3 July; pp. 51–56.

Mosavi, A., Ozturk, P., & Chau, K.-W. (2018). Flood prediction using machine learning models: Literature review. *Water, 10*(11), 1536. https://doi.org/10.3390/w10111536

Nash, J. E., & Sutcliffe, J. V. (1970). River flow forecasting through conceptual models Part I - A discussion of

principles. *Journal of Hydrologic Engineering, 10*(3), 282–290. https://doi.org/10.1016/0022-1694(70)90255-6

Nguyen, D. H., & Bae, D.-H. (2020). Correcting mean areal precipitation forecasts to improve urban flooding predictions by using long short-term memory network. *Journal of Hydrology, 584*, 124710. https://doi.org/10.1016/j.jhydrol.2020.124710

Nguyen, D. H., Le, X. H., Anh, D. T., Kim, S. H., & Bae, D.-H. (2022). Hourly streamflow forecasting using a Bayesian additive regression tree model hybridized with a genetic algorithm. *Journal of Hydrology, 606,* 127445. https://doi.org/10.1016/j.jhydrol.2022.127445

Nguyen, D. H., Le, X. H., Heo, J. Y., & Bae, D. H. (2021). Development of an Extreme Gradient Boosting Model Integrated With Evolutionary Algorithms for Hourly Water Level Prediction. *IEEE Access, 9,* 125853–125867. https://doi.org/10.1109/ACCESS.2021.3111287

Ni, L., Wang, D., Singh, V. P., Wu, J., Wang, Y., Tao, Y., et al. (2020). Streamflow and rainfall forecasting by two long short-term memory-based models. *Journal of Hydrology, 583*, 124296. https://doi.org/10.1016/j.jhydrol.2019.124296

Olah, C. (2015). Understanding LSTM networks. Available at: http://colah.github.io/posts/2015-08-Understanding-LSTMs/ (accessed on: 28 June 2018).

Papacharalampous, G., Tyralis, H., & Koutsoyiannis, D. (2019). Comparison of stochastic and machine learning methods for multi-step ahead forecasting of hydrological processes. *Stochastic Environmental Research and Risk Assessment, 33*(2), 481–514. https://doi.org/10.1007/s00477-018-1638-6

Phan, T. T. H., & Nguyen, X. H. (2020). Combining statistical machine learning models with ARIMA for water level forecasting: The case of the Red river. *Advances in Water Resources, 142,* 103656. https://doi.org/10.1016/j.advwatres.2020.103656

Rossum, G. (1995). Python tutorial, CWI (Centre for Mathematics and Computer Science), Amsterdam, The Netherlands.

Ruder, S. (2016). An overview of gradient descent optimization algorithm*s*, Availabe at: https://ruder.io/optimizing-gradient-descent/ (accessed on: 2020 Jun 6).

Shen, C. (2018). A Transdisciplinary Review of Deep Learning Research and Its Relevance for Water Resources Scientists. *Water Resources Research, 54*(11), 8558–8593. https://doi.org/10.1029/2018wr022643

Sit, M., Demiray, B. Z., Xiang, Z., Ewing, G. J., Sermet, Y., & Demir, I. (2020). A Comprehensive Review of Deep Learning Applications in Hydrology and Water Resources. ArXiv:abs/2007.12269

Thirumalaiah, K., & Deo, M. C. (2000). Hydrological forecasting using neural networks. *Journal of Hydrologic Engineering, 5*(2), 180–189. https://doi.org/10.1061/(ASCE)1084-0699(2000)5:2(180)

Van Der Walt, S., Colbert, S. C., & Varoquaux, G. (2011). The NumPy array: A structure for efficient numerical computation. *Computing in Science & Engineering, 13*(2), 22–30. https://doi.org/10.1109/mcse.2011.37

Xu, T., & Liang, F. (2021). Machine learning for hydrologic sciences: An introductory overview. *Wires Water, 8*(5), e1533. https://doi.org/10.1002/wat2.1533

Yang, C. H., Wu, C. H., & Hsieh, C. M. (2020). Long Short-Term Memory Recurrent Neural Network for Tidal Level Forecasting. *IEEE Access, 8*, 159389–159401. https://doi.org/10.1109/ACCESS.2020.3017089

Yaseen, Z. M., El-shafie, A., Jaafar, O., Afan, H. A., & Sayl, K. N. (2015). Artificial intelligence based models for streamflow forecasting: 2000–2015. *Journal of Hydrology, 530*, 829–844. https://doi.org/10.1016/j.jhydrol.2015.10.038

Young, P. C., (2002). Advances in real–time flood forecasting. *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, *360*(1796), 1433.

Yuan, X., Chen, C., Lei, X., Yuan, Y., & Muhammad Adnan, R. (2018). Monthly runoff forecasting based on LSTM–ALO model. *Stochastic Environmental Research and Risk Assessment, 32*(8), 2199–2212. https://doi.org/10.1007/s00477-018-1560-y

Zounemat-Kermani, M., Batelaan, O., Fadaee, M., & Hinkelmann, R. (2021). Ensemble machine learning paradigms in hydrology: A review. *Journal of Hydrology, 598*, 126266. https://doi.org/10.1016/j.jhydrol.2021.126266