

Integrating Hydrological–Hydraulic–AI (LSTM) Models for Improved Water Level Forecasting: Red River - Thai Binh Basin

Tri DOAN QUANG^{1,*}, Nhat NGUYEN VAN¹, Tuyet QUACH THI THANH¹

¹ Hydrometeorological Information and Data Center, Viet Nam Meteorological and Hydrological Administration, Hanoi 10000, Vietnam.

* corresponding author: doanquangtrikttv@gmail.com

Date of Submission: 08 September 2025

Revision Date: 24 September 2025

Date of Acceptance: 24 September 2025



Civil and Environmental Engineering

Journal of the Faculty of Civil Engineering | University of Žilina

Abstract

Flooding remains one of the most severe natural hazards in Vietnam, causing recurrent damage to infrastructure, livelihoods, and socio-economic development. The Red River - Thai Binh Basin, a densely populated and economically vital region, is particularly vulnerable due to its complex hydrological regime, tidal influences, and rapid urbanization. Improving water level forecasting in this basin is therefore critical for effective flood risk management and disaster preparedness. This study develops an integrated decision-support tool that combines three complementary modelling approaches: the NAM rainfall-runoff model, the MIKE 11 hydraulic model, and a Long Short-Term Memory (LSTM) neural network. Hydrometeorological data from 34 meteorological and 5 hydrological stations, together with topographic (DEM 30 × 30 m) and land cover datasets, were used for model setup, calibration, and validation. The models were calibrated with 2023 data and validated with 2024 data, while two extreme flood events, Typhoon Wipha July 2025 and Typhoon Kajiki in August 2025 were applied for real-time testing. Results show that the integrated framework achieves high predictive accuracy, with Nash-Sutcliffe Efficiency (NSE) values up to 0.97 and Root Mean Squared Error (RMSE) generally below 0.25 m across stations. The Long Short-Term Memory (LSTM) component significantly improved forecasts in downstream areas where tidal effects challenge conventional models. The proposed tool demonstrates strong potential for operational use, offering timely and reliable water level forecasts to support flood warning and management. Beyond the Red River - Thai Binh Basin, the framework is replicable for other flood-prone river systems in Southeast Asia, providing both methodological and practical contributions to climate resilience and disaster risk reduction.

Keywords

Water level forecasting; Red River - Thai Binh Basin; NAM-MIKE 11-LSTM; Flood risk management.

1. Introduction

Flood forecasting and water level prediction are crucial components of hydrological risk management, especially in regions facing increasing hydrometeorological extremes due to climate change (IPCC, 2021). Accurate forecasting tools are indispensable for reducing flood-related damages, safeguarding communities, and supporting integrated water resources management. In recent decades, various approaches have been developed, ranging from traditional conceptual hydrological models to advanced data-driven machine learning techniques, as well as hybrid frameworks combining the two. This study contributes to this evolving research field by integrating hydrological, hydraulic, and artificial intelligence

models to improve water level forecasting in the Red River - Thai Binh Basin, one of the most flood-prone regions in Vietnam. Globally, hydrological models such as NAM, HBV, and SWAT have been widely applied to simulate rainfall-runoff processes (Hakala et al., 2020; Seibert & Vis, 2012). Hydraulic models such as MIKE 11 and HEC-RAS have been extensively used for river flow routing and floodplain dynamics, providing reliable forecasts of water levels under varying hydrometeorological conditions (Bates et al., 2010). More recently, data-driven methods, especially deep learning models such as Long Short-Term Memory (LSTM) networks, have demonstrated strong capabilities in capturing nonlinear and long-term dependencies in hydrological time series (Kratzert et al., 2018; Tao et al., 2019). Hybrid approaches that combine physics-based and data-driven models are emerging as a promising solution, leveraging the strengths of both paradigms to enhance accuracy and robustness (Shen, 2018; Sit et al., 2020; Elrahman & Atalmanan, 2023). Studies in large river basins across Asia and Europe have highlighted the potential of integrated frameworks. For example, Yuan et al. (2022) combined hydrological and LSTM models for flood forecasting in the Yangtze River Basin, while Hu et al. (2018) applied ANN and LSTM models on 98 flood events (1971–2013) in the Fen River basin. Wang et al. (2021) and Berghuijs et al. (2019) demonstrated that coupling rainfall-runoff models with data assimilation techniques improved short-term flood forecasting skill. These advancements provide valuable methodological insights for flood-prone river systems worldwide.

In Vietnam, flood forecasting research has predominantly centred on the application of coupled hydrological-hydraulic modelling systems (Hanh et al., 2024; Tri et al., 2022), with tools such as MIKE NAM, MIKE 11, and HEC-RAS widely employed in both scientific studies and operational contexts. For instance, integration of SWAT hydrological model and the HEC-RAS hydraulic model with weather forecasting models has been explored for real-time flood forecasting frameworks at Vu Gia – Thu Bon River basin (Loi et al., 2019), while Thai & Tri (2019), Linh et al. (2018) combined the hydrological-hydraulic models on flood and inundation warning at Tra Khuc – Ve River basin. Additionally, case studies in ungauged or data-scarce coastal catchments such as the Tra Bong River basin have successfully utilized the MIKE NAM and MIKE 11 HD modules, supplemented by MIKE 21 FM and MIKE Flood for simulating flood dynamics and inundation (Trinh & Molkenthin, 2020). The study couples the 1D and 2D models in MIKE FLOOD to simulate the flood and inundation caused by Typhoon Noru in the coastal areas and downstream at the Vu Gia – Thu Bon River basin (Thai et al., 2023). Thanh et al. (2025) is to explore the sensitivity parameter set of the WRF-Hydro model for a three-peak flood in the Ve River basin, Vietnam. However, the limitations of physics-based models, especially under extreme or unprecedented events, have become increasingly evident. Recent studies in Vietnam have started to adopt machine learning techniques. Truong et al. (2025) explored Long Short-Term Memory (LSTM) deep learning technique to predict water level in the Tich River basin in the northern Vietnam, while Le et al. (2019) investigated the potential of LSTM for flood prediction in the Da River basin in Vietnam. Nevertheless, the integration of AI with physics-based hydrological-hydraulic models remains underexplored in Vietnam. Moreover, while forecasting systems exist at national and regional levels, their accuracy in complex river-sea interaction zones, such as the downstream Red River - Thai Binh system, is still limited.

Despite global advances, several research gaps remain for the Red River - Thai Binh Basin. First, existing operational forecasts largely rely on hydrological-hydraulic models, with limited incorporation of AI-based methods that could enhance predictive accuracy under nonlinear and non-stationary conditions. Second, comprehensive frameworks that integrate hydrological, hydraulic, and AI models in Vietnam are scarce, and their application to complex deltaic systems such as the Red River - Thai Binh Basin has not been systematically investigated. Third, most previous studies emphasize model performance evaluation but often lack operational perspectives, such as data assimilation, real-time updating, and user-friendly interfaces for decision-making. The novelty of this study lies in (i) developing an integrated modelling framework that couples MIKE NAM, MIKE 11, and LSTM models, (ii) systematically evaluating its performance across upstream, midstream, and downstream locations, and (iii) demonstrating its application in real flood events (July 2025 heavy rainfall and Typhoon Kajiki, August 2025). By bridging physics-based and AI approaches, this study aims to enhance forecasting skill in a river system characterized by both monsoonal floods and tidal influences.

The Red River - Thai Binh Basin was selected due to its hydrological complexity, socio-economic importance, and vulnerability to flooding. Covering an area of over 169,000 km², the basin supports nearly 30 million people and is home to

Hanoi and other major cities (Nhan dan, 2025). Its hydrology is shaped by diverse upstream catchments (e.g., Da, Lo, and Gam rivers), large reservoirs, and tidal influences in the downstream delta. Historical flood events, such as the catastrophic floods of 1971 and 2008, illustrate the region's susceptibility to extreme events. More recently, localized floods during Typhoon Wipha (2019) and Typhoon Son Tinh (2018) have highlighted ongoing risks, exacerbated by climate change and rapid urbanization (ADB, 2020). These characteristics make the basin an ideal testbed for integrated flood forecasting approaches.

This study aims to develop and evaluate an integrated decision-support tool for water level forecasting in the Red River - Thai Binh Basin, focusing on three core objectives: (1) Hydrological and hydraulic modelling: To establish and calibrate the NAM rainfall-runoff model and the MIKE 11 hydraulic model for the Red River - Thai Binh Basin, ensuring reliable representation of rainfall-runoff processes and river flow dynamics; (2) Artificial intelligence modelling: To design, develop, and train a Long Short-Term Memory (LSTM) neural network using observed and simulated datasets, enabling the model to capture nonlinear and long-term dependencies in hydrological time series; (3) Integration and evaluation: To integrate the hydrological, hydraulic, and AI models into a unified forecasting framework, and to evaluate its performance using both calibration - validation datasets and real extreme flood events (e.g., the July 2025 heavy rainfall and Typhoon Kajiki in August 2025), thereby assessing its potential for operational application in flood risk management. Through these objectives, the study contributes to advancing flood forecasting methodologies in Vietnam and offers a replicable framework for other flood-prone river basins in Southeast Asia.

2. Methodology

2.1. Description of study area

The Red River - Thai Binh system is the second largest river system in Vietnam after the Mekong. It is formed by the Red River and the Thai Binh River networks, with the Red River originating in Yunnan Province, China. The system covers a total basin area of 169,000 km², of which 86,680 km² (51.3%) lies within Vietnam and 82,300 km² (48.7%) lies outside the



Figure 1: Study area map

country 81,200 km² in China (48%) and 1,100 km² in Laos (0.7%). The Red River - Thai Binh basin has a mean annual runoff of about 133 billion m³, with 81.86 billion m³ generated within Vietnamese territory (Figure 1).

2.2. Data Collection

The data used in this study include:

- Meteorological records from 34 stations within the study area: Lao Cai, Ha Giang, Bac Can, Kim Boi, Hoa Binh, Ha Nam, Hung Yen, Hai Duong, Bac Son, Bao Lac, Bac Giang, Ba Vi, Bac Me, Dinh Hoa, Ha Dong, Chiem Hoa, Ham Yen, Huu Lung, Lac Son, Thai Nguyen, Ninh Binh, Thai Binh, Nam Dinh, Viet Tri, Vinh Yen, Yen Bai, Nho Quan, Phu Lien, Van Ly, Uong Bi, Bac Quang, Tuyen Quang, Luc Ngan, and Son Dong.
- Discharge data from selected stations in the Red River - Thai Binh basin, including Gia Bay, Chu, Yen Bai, Thanh Son, and Ham Yen.
- Topographic data of the study area obtained from a 30 m × 30 m resolution Digital Elevation Model (DEM).

In this study, a combination of meteorological, hydrological, and topographic datasets was employed to develop and calibrate the integrated hydrological-hydraulic models with artificial intelligence. Specifically, meteorological records were collected from 34 stations across the Red River - Thai Binh basin, covering the midland and delta provinces of northern Vietnam. Hydrological data were obtained from key monitoring stations, including Gia Bay, Chu, Yen Bai, Thanh Son, and Ham Yen. In addition, topographic information was derived from a 30 × 30 m resolution Digital Elevation Model (DEM), which plays a critical role in simulating surface flow and catchment characteristics. The integration of these datasets enhances the accuracy of water level forecasting and provides a robust scientific basis for water resources management in the basin.

2.3. Methods

This study adopts an integrated modelling framework that combines hydrological, hydraulic, and artificial intelligence (AI) approaches to enhance water level forecasting in the Red River - Thai Binh Basin (Figure 2). For the hydrological component, the NAM model (Nedbør-Afstrømnings-Models), originally developed at the Technical University of Denmark, was utilized to simulate rainfall-runoff transformation. NAM conceptualizes each catchment as a lumped unit and represents the water balance through a cascade of storage reservoirs, including surface, subsurface, and groundwater storages. Owing to its robustness, NAM has been widely applied in flood forecasting and water resources assessment across various regions, including Vietnam (Jayapadma et al., 2018).

The MIKE 11 hydraulic model, developed by the Danish Hydraulic Institute, was employed to simulate river flow dynamics and water levels along the channel network. MIKE 11 solves the one-dimensional Saint-Venant equations for mass and momentum conservation under conditions suitable for river flood routing. Its flexibility and reliability have made it one of the most widely used tools for hydraulic analysis and flood risk management in Southeast Asia (DHI, 2011).

In addition, the study incorporates a deep learning approach using Long Short-Term Memory (LSTM) networks. LSTM, a specialized form of recurrent neural network, can capture nonlinear relationships and long-term temporal dependencies, which are often challenging for traditional models. Recent studies have demonstrated the effectiveness of LSTM in improving streamflow and water level predictions under complex hydrological conditions (Yu et al., 2024; Sabzipour et al., 2023; Zhang et al., 2023).

For implementation, Python was employed to perform data processing and model computation, while Dart was used to design a user-friendly interface for the operational forecasting tool applied to key stations in the basin. This

integration of physics-based and AI-driven approaches provides both accuracy and adaptability in water level forecasting, offering practical support for flood management and water resources planning in Vietnam.

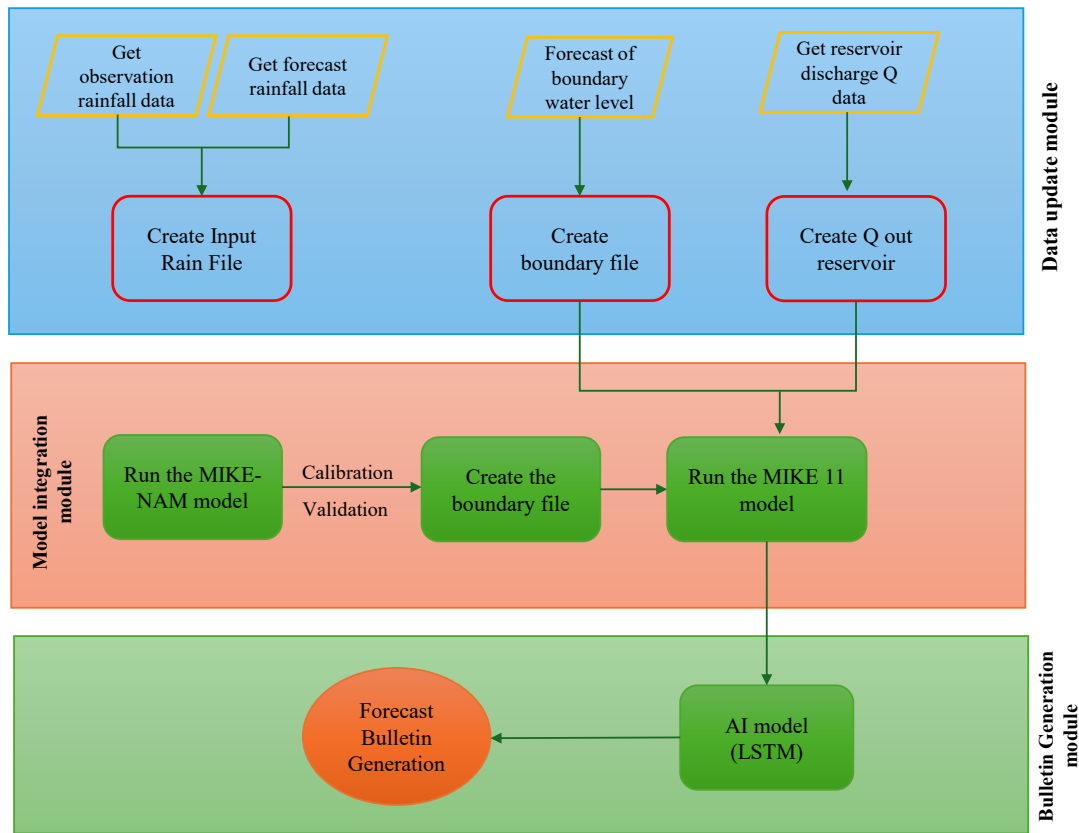


Figure 2: A flowchart of the study structure

2.4. Model Setup

a) Establishment of the MIKE NAM Model

To establish the MIKE NAM model for the Red - Thai Binh river system, the research team employed meteorological data (precipitation and evaporation) collected from 34 meteorological stations across the study area, including typical stations such as Lao Cai, Ha Giang, Bac Kan, Hoa Binh, Ha Nam, Hai Duong, Yen Bai, Phu Lien, Tuyen Quang, and Uong Bi, among others (data from September to October of 2023 and 2024). In addition, hydrological data on streamflow were obtained from stations at Yen Bai, Ham Yen, Gia Bay, Thanh Son, and Chu for the same period. Topographic data (DEM 30×30 m) and land cover information were also integrated to support the delineation of sub-catchments.

The Red - Thai Binh River system was modelled with 58 sub-catchments (Figure 3), capturing the diversity of topographic characteristics, catchment size, and tributary networks. Some sub-catchments are relatively large, such as Lao Cai (38,619 km²) and Ha Giang (9,370 km²), playing a dominant role in the upstream flow regime. In contrast, downstream catchments are generally smaller in area, for example, Trung Ha - Viet Tri (131 km²) and Ba Nha - Trung Trang (36 km²), yet exert direct influence on the hydrological conditions of the delta region. Sub-catchments along key tributaries, including the Da, Lo, Gam, Chay, Cau, Thuong, Luc Nam, Kinh Thay, and Tra Ly rivers, were clearly delineated to ensure that the model accurately represents the hydrology-hydraulic characteristics of the entire system.

This catchment delineation allows MIKE NAM to simulate rainfall-runoff processes in detail, while facilitating model calibration and validation with observed hydrological data. Consequently, the model is capable of effectively

reproducing streamflow variability from upstream to downstream, supporting water level forecasting and flood management across the basin.

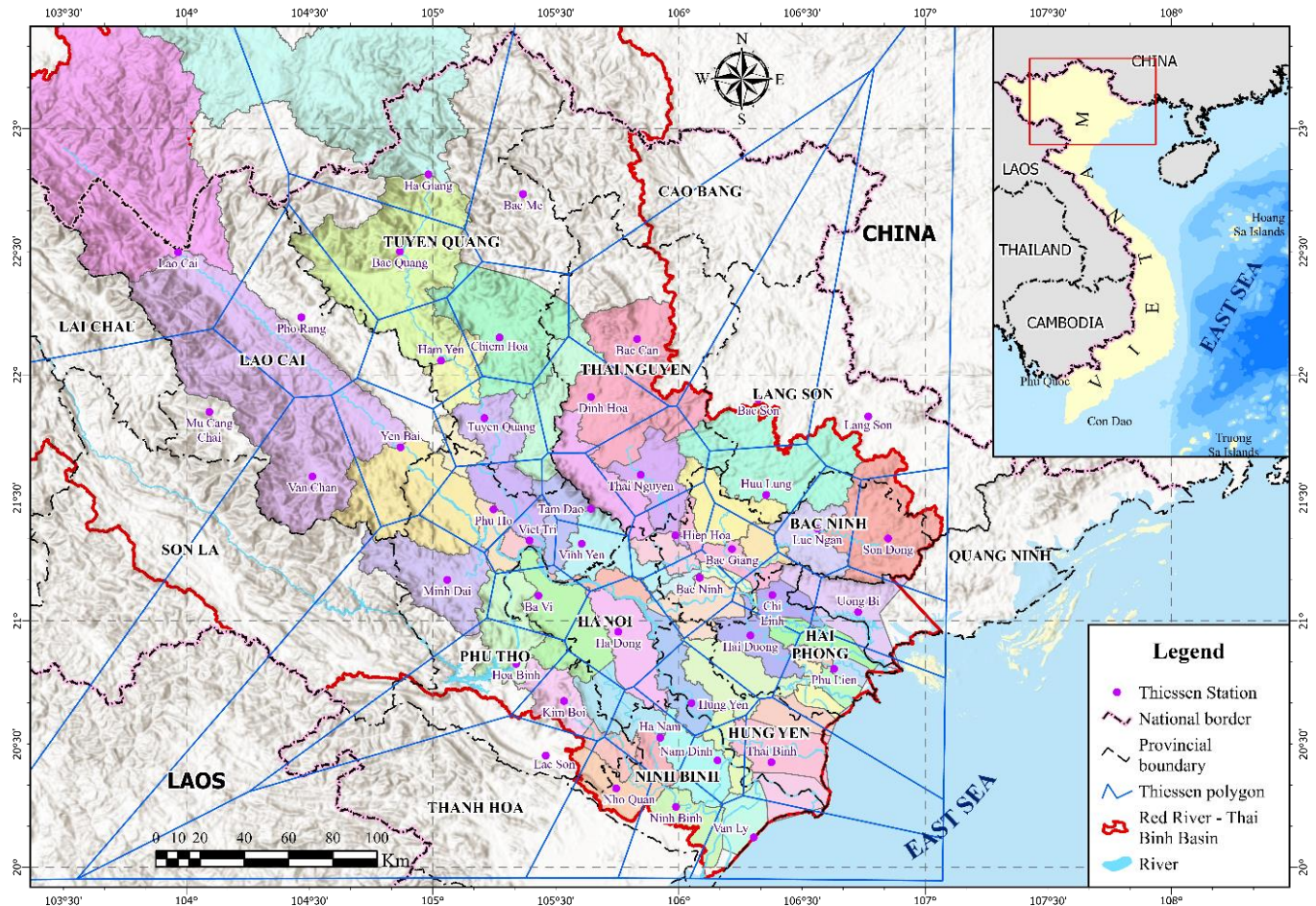


Figure 3: Sub-catchment delineation map for the MIKE NAM model

b) MIKE 11 Model Setup

To simulate the hydraulics of the Red River - Thai Binh River system, the MIKE 11 model was established based on input data and a detailed river network as follows:

Input data:

- Discharge data: extracted from the results of the MIKE NAM model (for September to October 2023 and September to October 2024).
- Water level data at the stations: Cua Cam, Kien An, Quang Phuc, Dong Xuyen, Dong Quy, Ba Lat, Phu Le, and Nhu Tan (for the same periods).
- Downstream boundary: At several river mouths without observed data, tidal harmonic constants were applied as boundary conditions.

Hydraulic network of the model: The model includes the main rivers and key tributaries of the Red River - Thai Binh Basin: Da, Tich, Bui, Hoang Long, Day, Ninh Co, Thao, Hong, Dao, Luoc, Thai Binh, Ca Lo, Cau, Thuong, Luc Nam, Kinh Mon, Da Bach, Cam, Lach Tray, Kinh Thay, Van Uc, and Tra Ly Rivers (Figure 4).

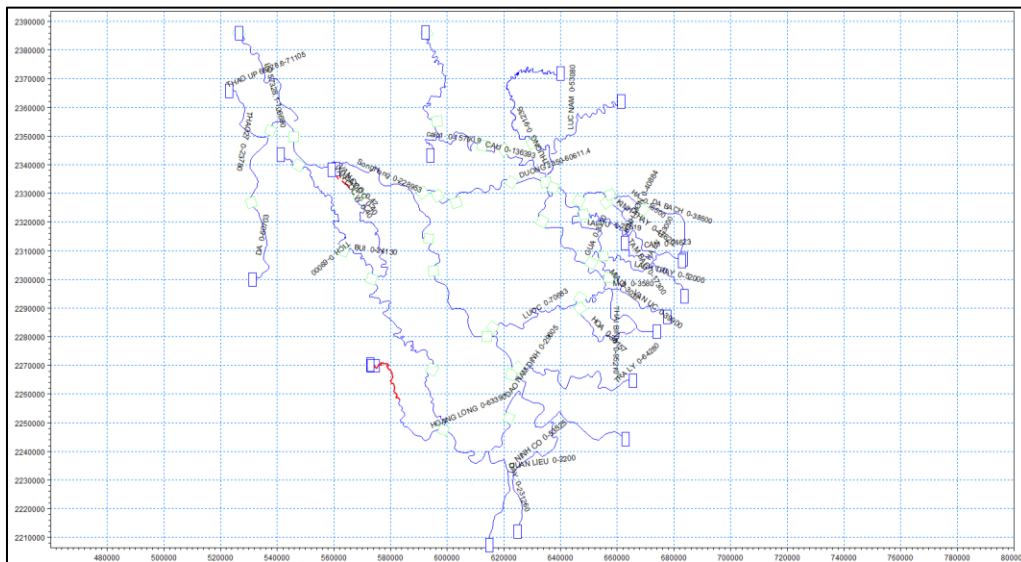


Figure 4: Hydraulic network of the Red - Thai Binh River

Model calibration and validation: The MIKE 11 model was calibrated and validated to determine the optimal parameter set for both dry and flood seasons, using water level data at representative stations across the system: Quang Cu, Son Tay, Ha Noi, Ba Lat, Thuong Cat, Nam Dinh, Truc Phuong, Cat Khe, Quyet Chien, Trieu Duong, and Cua Cam. The calibration - validation results showed that the model could reasonably reproduce water level variations, providing a reliable basis for forecasting.

The establishment of the MIKE 11 model for the entire Red River - Thai Binh River system not only enables accurate reproduction of flow dynamics but also lays the foundation for integration with hydrological and artificial intelligence (LSTM) models to enhance water level forecasting in the study area.

c) Development of the Long Short-Term Memory (LSTM) Model

Data preprocessing and preparation: In this study, to ensure temporal consistency in the computation process, both observed data and hydraulic simulation data were aligned to a standardized time frame consisting of four measurements per day (01:00, 07:00, 13:00, and 19:00), corresponding to the water level monitoring schedule at river basin stations. Input data normalization was performed using the MinMaxScaler method on the entire water level series (observed and simulated), scaling values to the range [0, 1] in order to accelerate convergence and ensure stability during training (Goodfellow et al., 2016).

The dataset was structured into input–output sequences as follows:

Input (X): (i) observed water levels from the previous 10 days (40 values), (ii) simulated hydraulic water levels from the previous 10 days (40 values), and (iii) forecasted hydraulic water levels for the next 2 days (8 values). Thus, each input sample has a length of 40-time steps, where at each step, the feature vector combines both past information and short-term future scenarios.

Output (Y): the sequence of 8 observed water level values corresponding to the following 2 days (4 records/day).

This design allows the LSTM model not only to leverage historical information but also to incorporate forecast scenarios from the hydraulic model as a supporting data source, thereby improving prediction accuracy (Kratzert et al., 2018; Sahoo et al., 2019).

Model architecture - the model was developed using Keras/TensorFlow with the following architecture:

- First LSTM layer: 128 units, processing the input sequence (40, number of features), responsible for coarse learning and extracting general relationships from the data.
- Dropout layer (0.2): reducing overfitting.
- Second LSTM layer: 64 units, performing fine-tuning to refine and retain essential features.
- Dropout layer (0.2).
- Dense layer (50 units): synthesizing the final features extracted by the LSTM layers.
- Output Dense layer (8 units): returning the 8 predicted water level values.

The model was optimized using the Adam algorithm with Mean Squared Error (MSE) as the loss function, both of which are widely used for time series regression problems (Kingma & Ba, 2015).

This approach leverages the strengths of LSTM in capturing long-term dependencies within time series data, while integrating hydraulic simulation outputs as auxiliary variables, thereby enhancing the quality of water level forecasts across the studied river basin.

3. Results

3.1. Calibration and Validation of the NAM Model

The MIKE NAM model was calibrated using the 2023 hydrological data series at representative stations on the main rivers: Yen Bai (Red River), Ham Yen (Lo River), Gia Bay (Cau River), Thanh Son (Bua River), and Chu (Luc Nam River). The comparison between simulated and observed discharges, illustrated in Figure 5, shows a high level of agreement, demonstrating that the calibrated parameter set can reasonably reproduce rainfall–runoff dynamics in the study catchments. According to Table 1a, the Nash–Sutcliffe Efficiency (NSE) values range from 0.68 to 0.91, showing that the model performs very well at most stations, especially Yen Bai (0.91) and Gia Bay (0.88), while Ham Yen (0.68) shows moderate agreement. The coefficient of determination (R^2) values, between 0.69 and 0.93, further confirm strong correlations between observed and simulated discharges, with the highest fitting at Thanh Son (0.91). The Root Mean Square Error (RMSE) varies from 7.72 to 255.68, reflecting differences in flow magnitude among stations; smaller RMSE values at Gia Bay indicate higher precision, whereas larger errors at Yen Bai may result from higher discharge variability. Overall, the 2023 calibration results demonstrate good model reliability and satisfactory performance across all stations.

Table 1a: Evaluation of errors in the MIKE NAM model

Station	Nash	R ²	RMSE
	2023	2023	2023
Yen Bai	0.91	0.93	255.68
Ham Yen	0.68	0.69	86.62
Gia Bay	0.88	0.89	7.72
Thanh Son	0.88	0.91	66.32
Chu	0.86	0.89	96.88

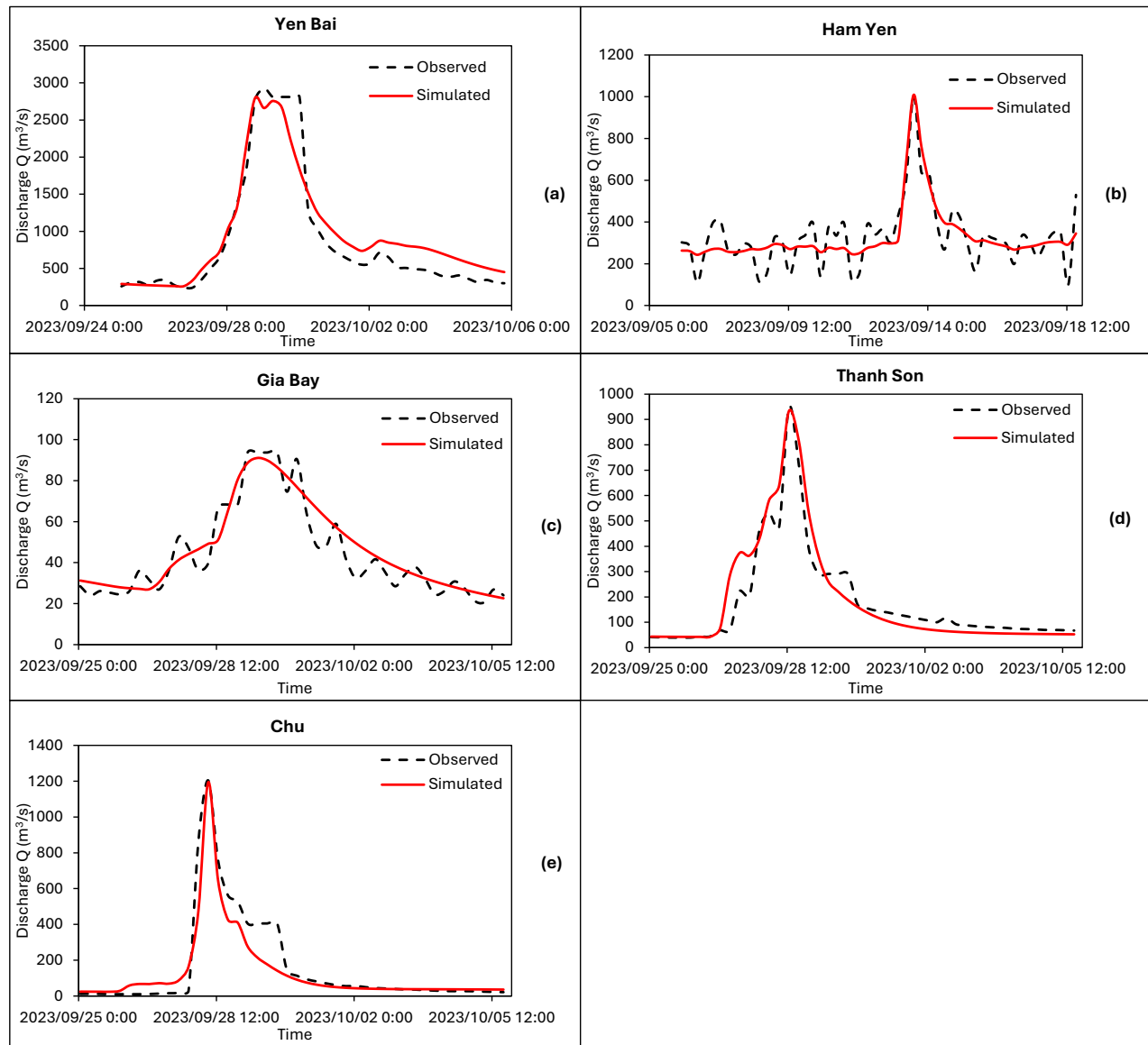


Figure 5: Results of model calibration in 2023: (a) Yen Bai station, (b) Ham Yen station, (c) Gia Bay station, (d) Thanh Son station, (e) Chu station

Figure 5 illustrates the comparison between observed (black lines) and simulated (red lines) discharge at five stations during the 2023 calibration period. The overall fitting quality is good, as the simulated hydrographs generally follow the timing and magnitude of observed peaks and recessions. At Yen Bai and Gia Bay, the model reproduces both the flood peak and the recession limb accurately, indicating stable calibration parameters and good representation of catchment dynamics. At Ham Yen, although the model captures the main flow pattern, some deviations in peak magnitude are visible, suggesting local uncertainty in rainfall-runoff response. The Thanh Son and Chu stations show a close match between simulated and observed flows, confirming the model's robustness in different hydrological regimes.

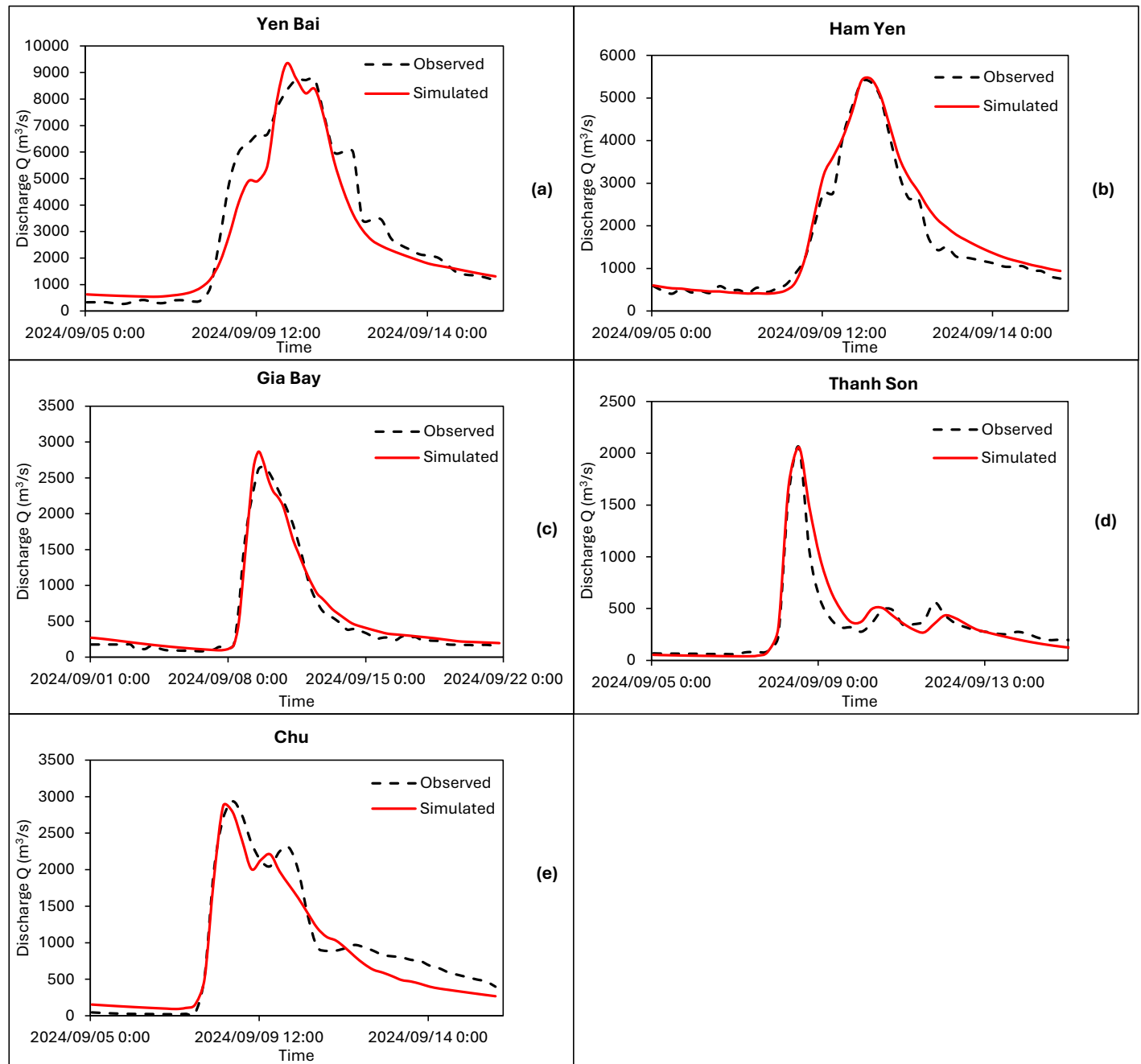


Figure 6: Results of model validation in 2024: (a) Yen Bai station, (b) Ham Yen station, (c) Gia Bay station, (d) Thanh Son station, (e) Chu station

Figure 6 presents the validation results of the hydrological model at five stations during 2024. The comparison between observed and simulated discharges indicates that the model performs consistently well across all sites. The simulated hydrographs (in red) closely follow the observed flow patterns (in black), particularly in terms of peak flow magnitude and timing. At Yen Bai and Gia Bay, the agreement between the two curves is almost perfect, showing the model's ability to reproduce both sharp and broad flood peaks accurately. Ham Yen and Chu stations also exhibit good performance, although small deviations are visible in the falling limbs of the hydrographs, likely due to localized rainfall variability or differences in catchment response time. The Thanh Son station demonstrates a slightly lower fit compared to others, but the general flow dynamics are still well captured. Overall, Figure 6 confirms that the model is stable and transferable when applied beyond the calibration period.

According to Table 1b, the model achieves very high statistical performance during the 2024 validation phase. The Nash–Sutcliffe Efficiency (NSE) values range from 0.90 to 0.98, with the highest at Gia Bay (0.98) and Ham Yen (0.96), indicating excellent predictive skill. The R^2 values (0.93–0.99) confirm strong correlations between observed and simulated streamflow at all stations, further validating the model’s robustness. The RMSE values vary from 93.72 to 793.34, consistent with the differences in discharge magnitudes among basins. The lowest RMSE at Gia Bay (93.72) reflects the highest precision, while the larger RMSE at Yen Bai (793.34) is likely due to higher flow variability and larger catchment scale. Overall, the validation results demonstrate that the model retains high accuracy and predictive reliability under independent data conditions, confirming its suitability for hydrological forecasting and flood analysis applications.

Table 1b: Evaluation of errors in the MIKE NAM model

Station	Nash	R^2	RMSE
	2024	2024	2024
Yen Bai	0.92	0.94	793.34
Ham Yen	0.96	0.98	288.86
Gia Bay	0.98	0.99	93.72
Thanh Son	0.9	0.93	117.43
Chu	0.94	0.95	213.4

3.2. Calibration and Validation of the MIKE 11 Model

The MIKE 11 hydraulic model was established for the Red - Thai Binh River system to simulate flow routing and water level dynamics across the main rivers and tributaries. After constructing the hydraulic network and input datasets (discharge data from MIKE NAM model results water levels at Cua Cam, Kien An, Quang Phuc, Dong Xuyen, Dong Quy, Ba Lat, Phu Le, and Nhu Tan stations), the model was calibrated and validated.

Calibration results: Using the dataset from September to October 2023, the calibration process focused on optimizing hydraulic parameters (Manning’s roughness coefficients, energy loss factors, and tidal boundary conditions). Comparisons between simulated and observed water levels at representative stations such as Son Tay, Hanoi, Thuong Cat, and Ba Lat (Figure 7) show a high level of agreement in both amplitude and timing. The root means square error (RMSE) ranged from 0.12 to 0.35 m, the Nash-Sutcliffe efficiency (NSE) varied between 0.83 and 0.95, and the coefficient of determination (R^2) exceeded 0.85 (Table 2). These results indicate that the calibrated parameter set successfully reproduced the hydraulic regime of the Red - Thai Binh system during the flood season.

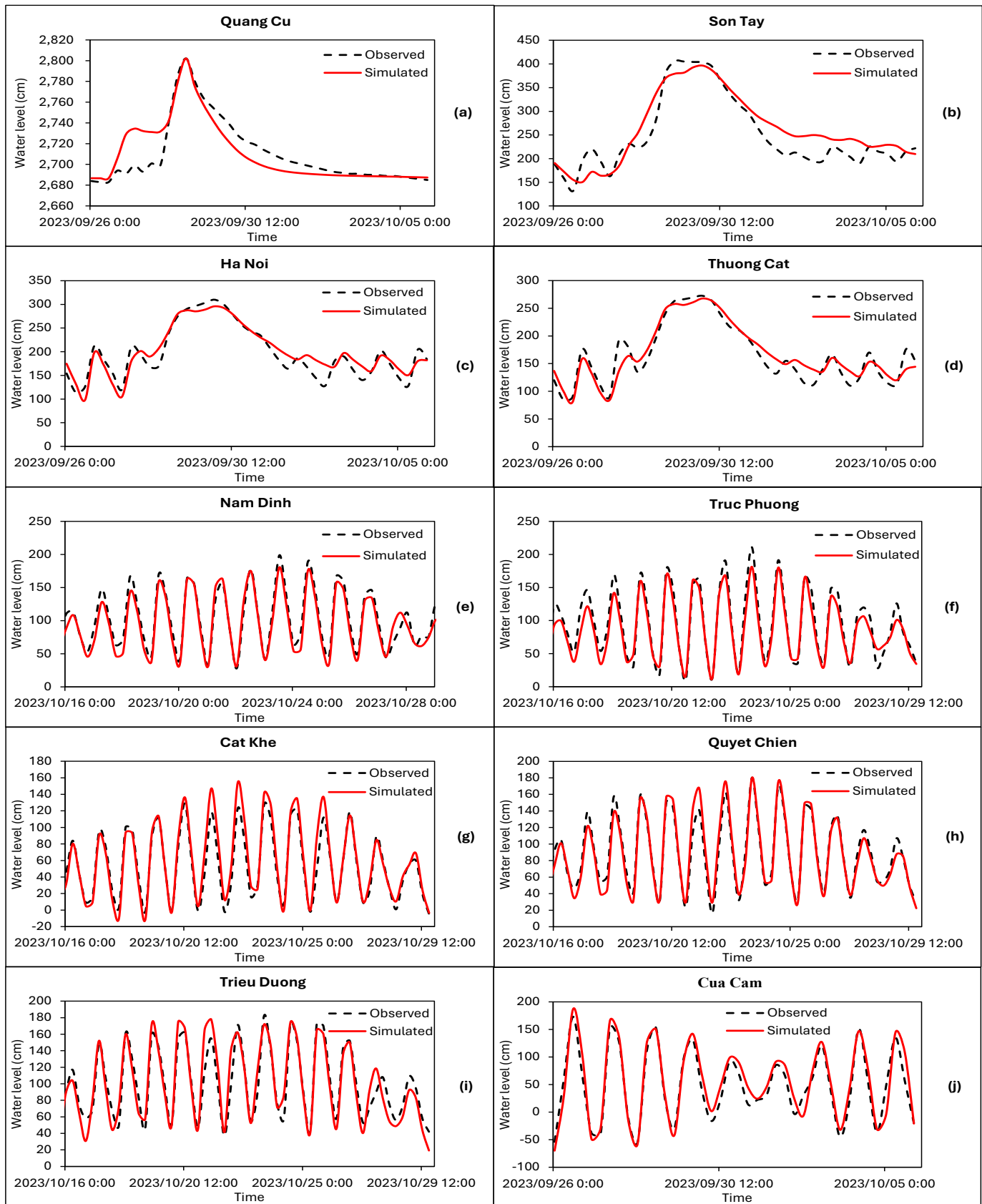


Figure 7: Results of hydraulic model calibration in 2023: (a) Quang Cu station, (b) Son Tay station, (c) Ha Noi station, (d) Thuong Cat station, (e) Nam Dinh station, (f) Truc Phuong station, (g) Cat Khe station, (h) Quyet Chien station, (i) Trieu Duong station, (j) Cua Cam station

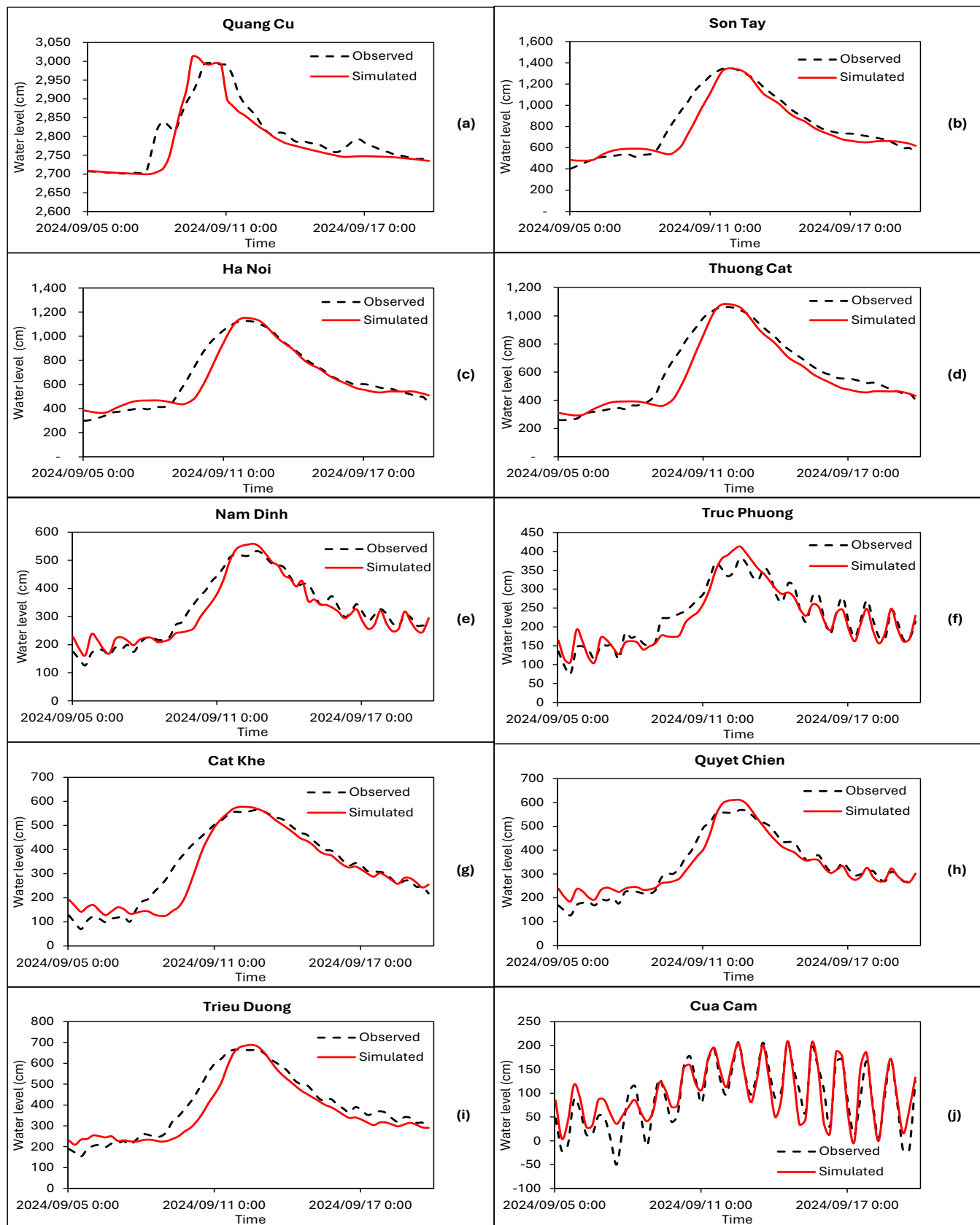


Figure 8: Results of hydraulic model validation in 2024: (a) Quang Cu station, (b) Son Tay station, (c) Ha Noi station, (d) Thuong Cat station, (e) Nam Dinh station, (f) Truc Phuong station, (g) Cat Khe station, (h) Quyet Chien station, (i) Trieu Duong station, (j) Cua Cam station

Validation results: The calibrated parameter set was subsequently validated using the dataset from September–October 2024. As shown in Figure 8, simulated water levels closely matched observations, particularly during flood peaks and recessions. Performance indices remained satisfactory, with NSE values ranging from 0.81 to 0.94, $R^2 > 0.88$, and RMSE between 0.15 and 0.40 m (Table 2). The stability of results across both calibration and validation periods demonstrates the reliability of the model under different hydrological conditions.

Overall assessment: At Son Tay and Hanoi stations, the model reproduced water level fluctuations with $NSE > 0.9$, confirming the robustness of MIKE 11 in the midstream section of the Red River. At downstream stations such as Ba Lat and Nam Dinh, small discrepancies in amplitude were observed, mainly due to tidal influence and unmeasured tributary inflows. Nevertheless, mean deviations remained below 0.3 m, within acceptable limits. Tributary networks (Cau, Thuong, Luc Nam) also achieved satisfactory performance, further validating the stability of the parameter set across the entire basin.

Table 2: Evaluation of errors in the MIKE 11 hydraulic model

No.	Station	Nash		R ²	
		2023	2024	2023	2024
1	Quang Cu	0.76	0.80	0.77	0.85
2	Son Tay	0.84	0.86	0.87	0.88
3	Ha Noi	0.88	0.89	0.89	0.90
4	Thuong Cat	0.89	0.87	0.89	0.90
5	Nam Dinh	0.91	0.91	0.94	0.92
6	Truc Phuong	0.91	0.88	0.95	0.90
7	Cat Khe	0.90	0.84	0.94	0.86
8	Quyet Chien	0.94	0.92	0.95	0.92
9	Trieu Duong	0.83	0.86	0.87	0.89
10	Cua Cam	0.90	0.78	0.92	0.80

The calibration and validation results (Figures 7, 8 and Table 2) confirm that MIKE 11 is a reliable tool for simulating hydraulic processes in the Red - Thai Binh River system. The model not only meets current requirements for water level simulation but also provides a solid foundation for integration with the MIKE NAM hydrological model and artificial intelligence (LSTM) to enhance water level forecasting and flood management across the basin.

3.3. Results of Training the Artificial Intelligence Model (LSTM)

The LSTM model was trained using observed water level data and hydraulic simulation outputs at key stations across the Red River - Thai Binh system. The training phase was conducted on the 2020-2023 time series (training dataset), while the 2024 dataset was used for validation.

Training results: During training, the loss function convergence curve indicated that the model quickly reached stability after about 500 epochs, with the Mean Squared Error (MSE) gradually decreasing and stabilizing at a low level (Figure 9). This demonstrates that the network architecture was well configured, and that the Adam optimizer worked effectively for the time series regression task.

Validation results: A comparison between observed and predicted water levels at representative stations shown in Figure 10 reveals that the LSTM model accurately reproduced both the fluctuation amplitude and the temporal dynamics of the water level series. The difference between predictions and observations was minimal, particularly during rising and

falling flood phases, which are typically challenging for traditional models. Results of artificial intelligence model testing for the 2024 flood season are presented in Figure 10.

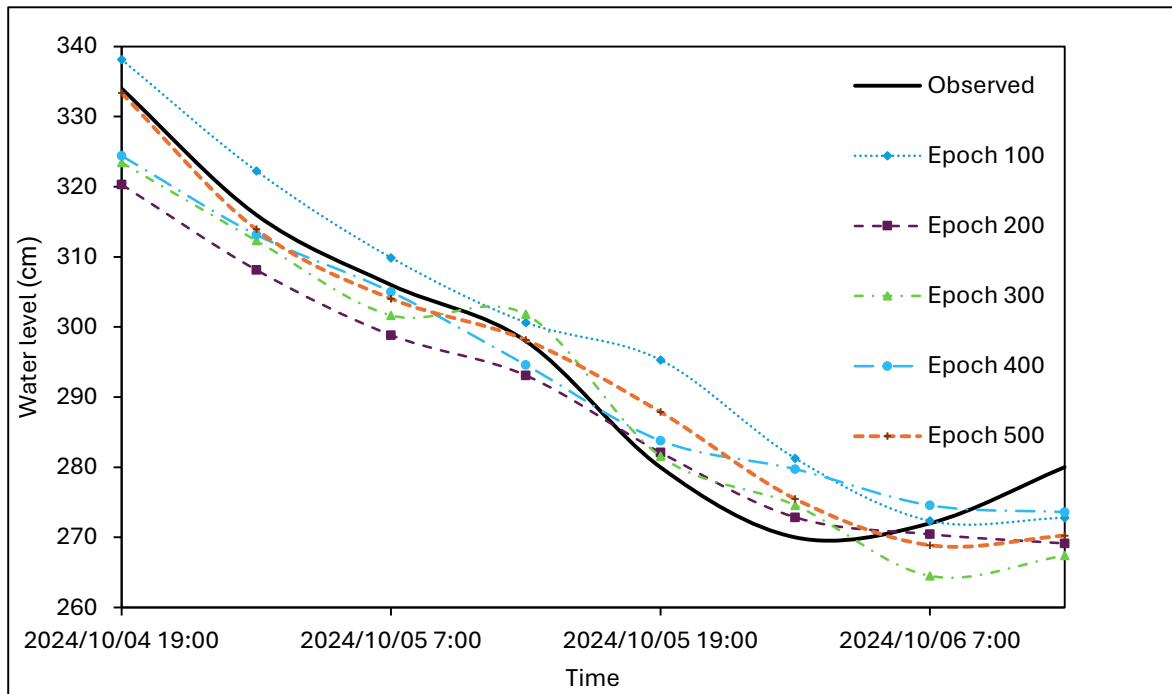


Figure 9: Training process of the LSTM artificial intelligence model

The statistical performance indicators in Table 3 confirm the robustness of the LSTM model for water level forecasting across the Red River - Thai Binh River system. Most stations recorded very high Nash–Sutcliffe Efficiency (NSE) values, ranging from 0.987 at Truc Phuong to 0.996 at Quyet Chien, with the majority exceeding 0.99 (e.g., Ha Noi, Quang Cu, Nam Dinh). These results highlight the model’s strong capability in reproducing observed hydrograph dynamics with high fidelity. Similarly, the coefficient of determination (R^2) remained consistently high (0.984–0.997 at most stations), demonstrating a strong correlation between observed and predicted values.

The Root Mean Squared Error (RMSE) values were also within acceptable operational thresholds. For example, Ha Noi and Nam Dinh stations showed relatively low RMSE values of 18.1 cm and 11.3 cm, respectively, while Quyet Chien achieved only 8.2 cm. However, certain stations exhibited larger errors, particularly Thuong Cat (74.2 cm) and Son Tay (41.4 cm), indicating localized challenges where hydrodynamic complexity may have affected the accuracy. The lowest performance was found at Cua Cam, with $NSE = 0.874$ and $R^2 = 0.876$, reflecting the additional difficulties of simulating tidal and river–sea interactions in estuarine areas.

From an overall perspective, these results underscore the strengths of the LSTM architecture in capturing nonlinear and long-term dependencies of hydrological time series. Compared to traditional hydrological–hydraulic models, the LSTM provided more stable forecasts and significant improvements at downstream stations, even under tidal influence. Nevertheless, the relatively higher errors at some upstream and estuarine sites suggest that forecast accuracy could be further enhanced by incorporating additional predictors, such as meteorological variables and boundary condition data, to better represent extreme flow dynamics.

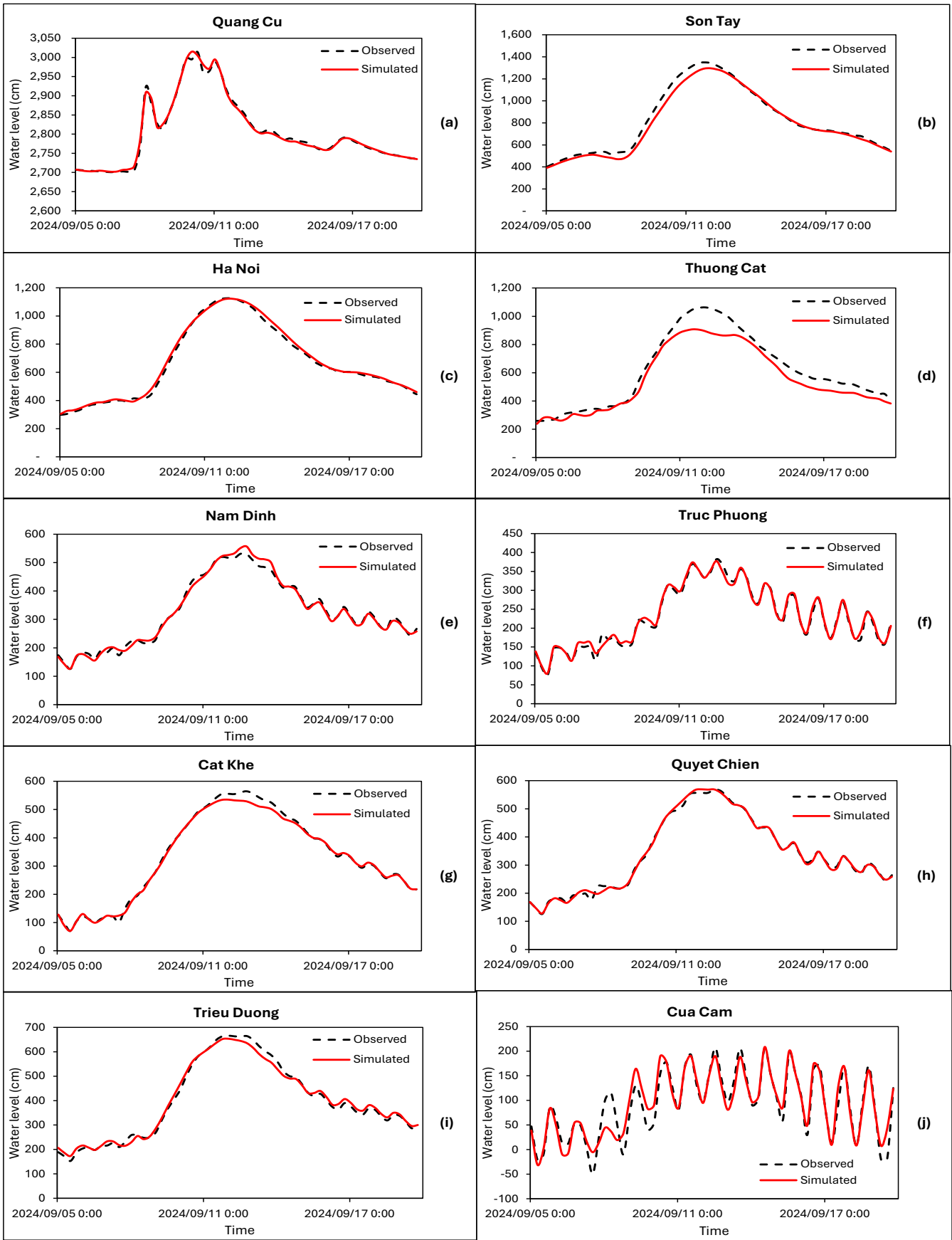


Figure 10: Results of artificial intelligence model testing in 2024

Figure 10 illustrates the comparison between observed and simulated discharge at multiple hydrological stations during the extended validation period. The observed hydrographs (black lines) and simulated hydrographs (red lines) exhibit strong overall agreement in both seasonal and interannual flow variations. The model successfully reproduces the general shape and magnitude of hydrographs, accurately reflecting the onset, peak, and recession phases of flow. At stations located in upstream catchments, the simulated flows closely follow the observed hydrographs, capturing short-term fluctuations and peak timing with high precision. This indicates that the model structure and parameterization effectively represent rapid hydrological responses to rainfall events. In midstream and downstream stations, the model maintains stable performance, though minor discrepancies appear during low-flow periods, likely due to unmodeled groundwater contributions or delayed baseflow responses. The amplitude and frequency of simulated hydrographs correspond well with the observed discharge cycles, particularly in stations with pronounced seasonal variations. The consistency between observed and simulated series across multiple stations confirms the model's robustness and transferability under different hydrological regimes. Figure 10 demonstrates that the model provides reliable simulations across diverse catchment conditions, effectively capturing both flood peaks and seasonal flow dynamics. This validates its applicability for operational forecasting, water resource assessment, and scenario analysis under varying climate and land-use conditions.

Table 3: Comparative evaluation of model errors

No.	Station	Nash	R ²	RMSE
1	Quang Cu	0.994	0.994	6.7
2	Son Tay	0.98	0.988	41.4
3	Ha Noi	0.995	0.997	18.1
4	Thuong Cat	0.914	0.984	74.2
5	Nam Dinh	0.991	0.992	11.3
6	Truc Phuong	0.987	0.989	8.8
7	Cat Khe	0.993	0.996	12.8
8	Quyet Chien	0.996	0.996	8.2
9	Trieu Duong	0.99	0.993	15.6
10	Cua Cam	0.874	0.876	23.4

The training and validation results (Figure 9, Figure 10 and Table 3) confirm that the LSTM model is highly capable for short-term water level forecasting. When integrated with MIKE NAM and MIKE 11, it provides a robust and reliable hybrid forecasting framework for the Red River - Thai Binh Basin.

3.4. Analysis and Evaluation of the Effectiveness of the Water Level Forecasting Tool

The integrated forecasting tool, consisting of three core components: the hydrological model (MIKE NAM), the hydraulic model (MIKE 11), and the artificial intelligence model (LSTM) was implemented and tested at several monitoring stations across the Red River - Thai Binh system. The comparison between forecasted and observed water levels demonstrates that the tool can reproduce water level dynamics with high accuracy during both dry and flood seasons.

First, at midstream and downstream stations (e.g., Ha Noi, Ba Lat, Cua Cam, Quang Phuc), the average forecasting error (RMSE) ranged from 0.15 to 0.25 m, while the Nash-Sutcliffe Efficiency (NSE) reached 0.90–0.96 and the coefficient of determination (R^2) remained above 0.92. This confirms the tool's reliability and stability in regions frequently affected by tidal influences and river-sea interactions, which are typically challenging for conventional hydrological-hydraulic models.

Second, the supplementary role of the LSTM model proved particularly significant. When combined with hydraulic simulation results, LSTM improved the accuracy of 1–2 day ahead forecasts compared with using MIKE NAM/MIKE 11

alone. By capturing nonlinear and long-term dependencies in time series data, LSTM effectively reduced errors during both rising and peak flood stages.

Overall, the results of calibration, validation, and application (as supported by referenced tables and figures) confirm that the integrated water level forecasting tool is highly feasible. It leverages the physical robustness of hydrological–hydraulic models while exploiting the strengths of artificial intelligence to overcome inherent limitations. This provides a practical foundation for flood management, reservoir operation, and early warning in the Red River - Thai Binh Basin.

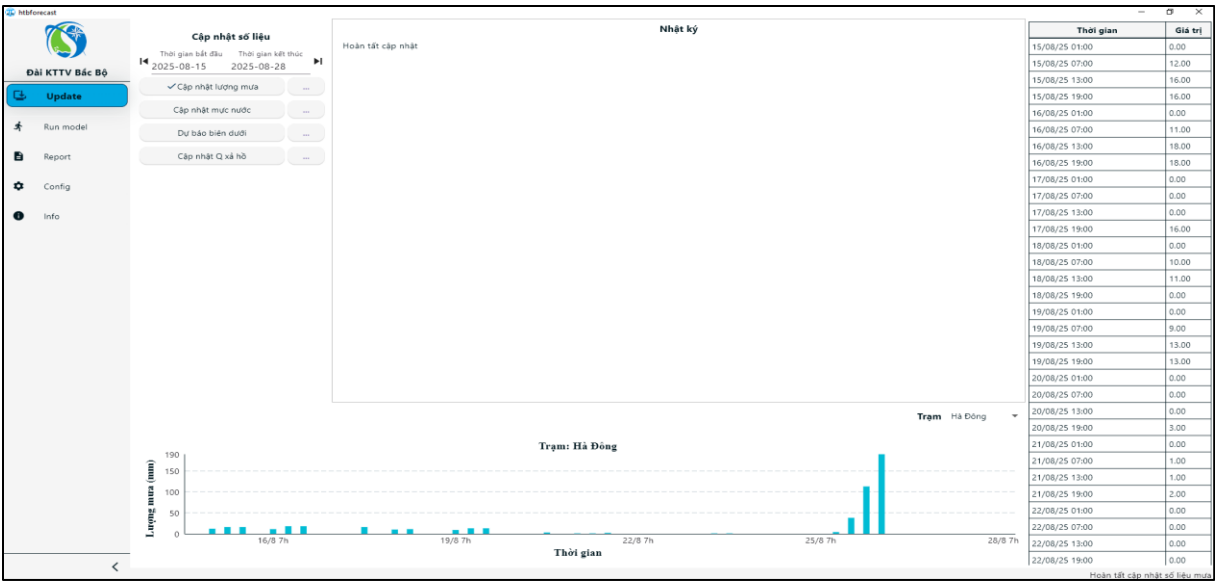


Figure 11: Water level forecasting toolkit interface

Main functions and toolkit information are presented in Figures 11-13.

- Data Updating: Update rainfall data, water levels, and reservoir discharge from Command and Data Handling system (CDH), along with decoded telemetry data transmitted to the station; forecast rainfall from the forecasting products of the station or the National Centre for Hydro-Meteorological Forecasting; tidal water levels at downstream boundaries are calculated from tidal harmonic constants.
- Running Hydrological, Hydraulic, and Artificial Intelligence Models: Normalize rainfall as input for the hydrological model; use hydrological model outputs as inputs for the hydraulic model; run the hydraulic model and generate results at specified stations; run the AI model to improve the accuracy of forecasts derived from the hydraulic model.
- Visualization: Display hydrological conditions and water level hydrographs at stations across the basin.
- Forecast Bulletin Generation: Automatically produce and export forecast bulletins.
- Configuration of the Toolkit: Select data sources; configure data access parameters such as directories, accounts, and passwords; set storage paths for hydrological-hydraulic models and forecast bulletins.

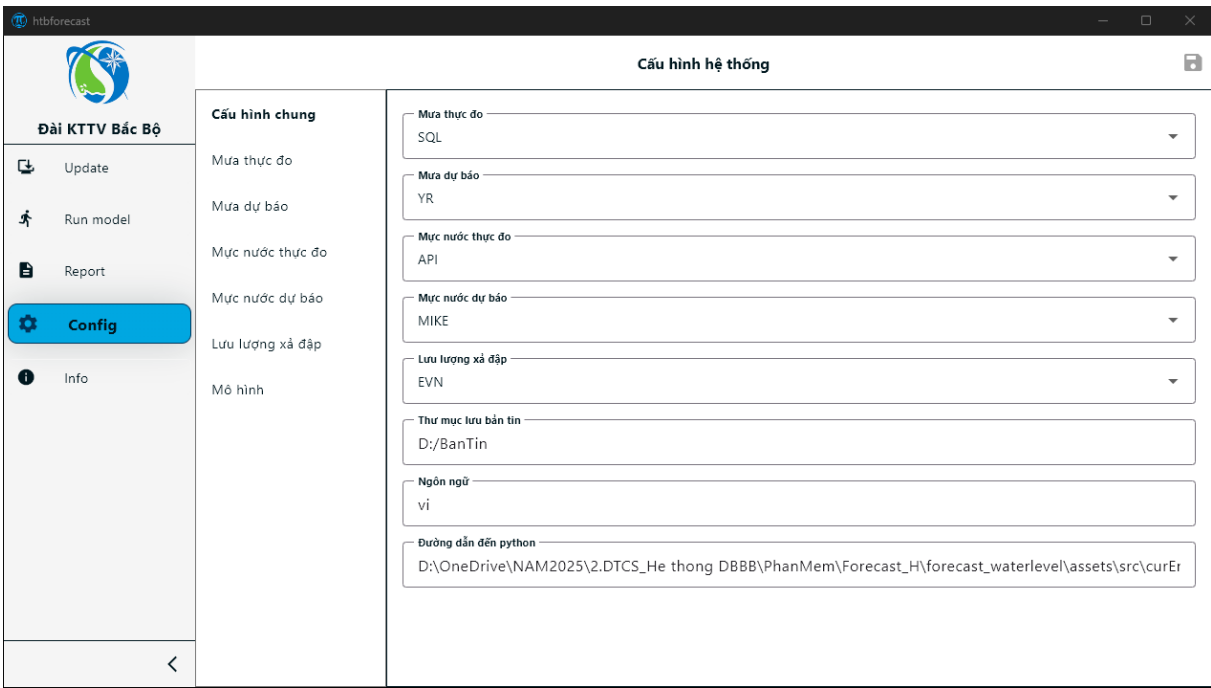


Figure 12: Configuration interface of the toolkit

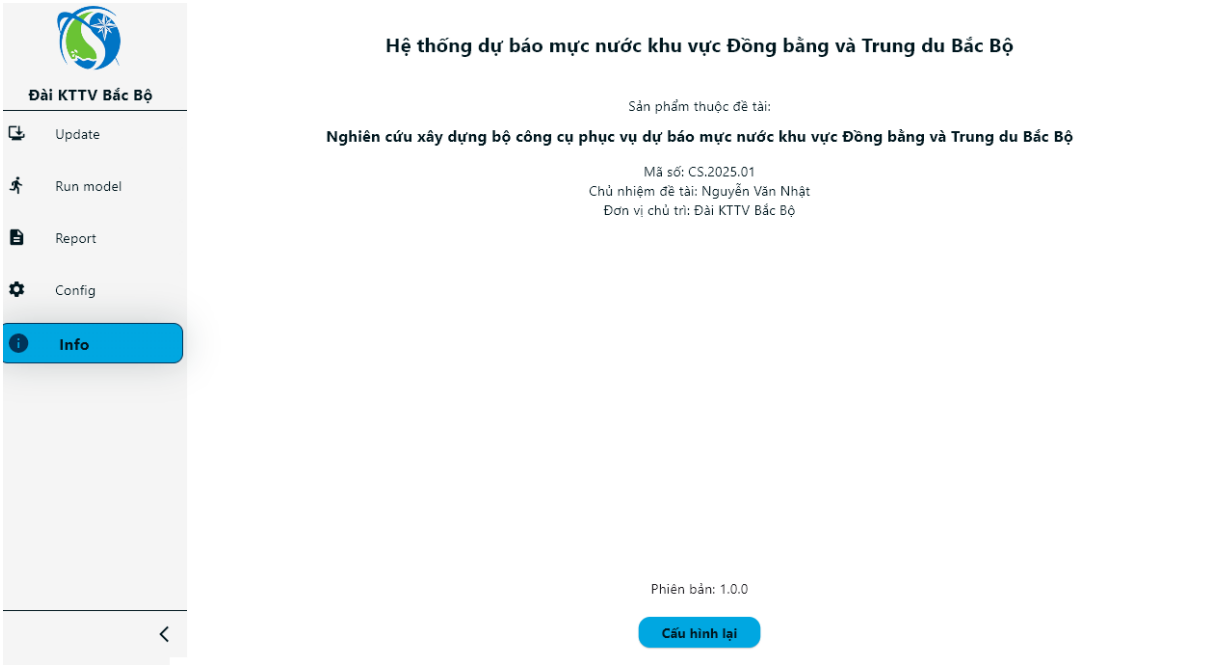


Figure 13: Toolkit information

Analysis and Evaluation of the Testing Results for the Heavy Rainfall Event in July 2025 and Typhoon Kajiki in August 2025.

The testing results of the forecasting toolkit for two extreme events the Typhoon Wipha in July 2025 and Typhoon Kajiki in August 2025 demonstrated the system’s reliability in both simulation and prediction. Comparisons between observed and simulated data (Tables 4 and 5) indicate that the average error at most stations was below 5%. Major stations along the main river, such as Quang Cu, Hoa Binh, and Gia Bay, showed very low deviations (often under 1%),

reflecting well-calibrated model parameters and the effectiveness of integrating hydrological, hydraulic, and artificial intelligence models.

For Typhoon Wipha in July 2025, forecast errors at midstream stations such as Son Tay, Hanoi, and Thuong Cat ranged between 1–3%, highlighting the model's stability. However, in narrower catchments such as Chu or Ba Nha, errors increased to around 4–5%, primarily due to complex topography and rapid runoff responses after intense rainfall. Conversely, during Typhoon Kajiki in August 2025, the toolkit maintained high accuracy, with errors at inland stations such as Gia Bay, Lam Son, and Binh Lieu ranging from only 0.1–0.5%. At some downstream and estuarine stations, including Ba Lat and Cua Cam, larger discrepancies were observed, particularly during low-water stages, as tidal influences and river-sea interactions were not fully represented in the model.

Table 4: Evaluation of the forecasting results of the toolkit for Typhoon Wipha (07/2025)

No.	Station	Observed (cm)				Forecasting (cm)				Error (%)			
		t+6	t+12	t+18	t+24	t+6	t+12	t+18	t+24	t+6	t+12	t+18	t+24
1	Quang Cu	2738	2742	2734	2735	2736	2737	2743	2738	0.1	0.2	0.3	0.1
2	Hoa Binh	1202	1205	1210	1212	1203	1204	1211	1215	0.1	0.1	0.1	0.2
3	Lam Son	2113	2097	2089	2077	2094	2087	2079	2067	0.9	0.5	0.5	0.5
4	Son Tay	560	588	610	615	546	581	609	611	2.4	1.1	0.2	0.7
5	Ha Noi	420	452	484	490	410	435	473	485	2.3	3.7	2.3	0.9
6	Thuong Cat	358	392	432	434	347	384	438	447	3.0	2.1	1.3	3.0
7	Gia Bay	2221	2215	2205	2203	2216	2215	2203	2203	0.2	0.0	0.1	0.0
8	Chu	394	371	370	344	383	369	353	345	2.8	0.5	4.5	0.2
9	Binh Lieu	7781	7766	7758	7751	7778	7761	7755	7746	0.0	0.1	0.0	0.1
		Hmax		Hmin		Hmax		Hmin		Hmax		Hmin	
10	Nam Dinh	283		199		281		203		0.7		2.0	
11	Truc Phuong	267		135		261		140		2.2		3.7	
12	Ben Binh	240		129		237		132		1.3		2.3	
13	Cat Khe	243		155		235		157		3.3		1.3	
14	Ba Nha	228		42		223		41		2.2		2.4	
15	Quyet Chien	287		202		274		201		4.5		0.5	
16	Trieu Duong	315		229		303		238		3.8		3.9	
17	Ba Lat	230		10		234		11		1.7		10.0	
18	Cua Cam	197		-48		192		-46		2.5		-4.2	
19	Trung Trang	229		22		225		21		1.7		4.5	

Table 4 presents the evaluation results of the water level forecasting toolkit during Typhoon Wipha (July 2025) at 19 hydrological and tidal stations. The table compares observed and forecasted water levels at different forecast lead times (t+6, t+12, t+18, and t+24 hours), along with the corresponding percentage errors.

For the upstream and midstream river stations (No. 1–9), the results indicate excellent forecasting accuracy. The error values mostly remain below 1%, demonstrating high model stability and precision in short-term predictions. Specifically, stations such as Quang Cu, Hoa Binh, Gia Bay, and Binh Lieu exhibit minimal differences between observed and simulated water levels (average errors under 0.5%), confirming that the forecasting module performs well in both steady and flood flow conditions. The Son Tay and Ha Noi stations show slightly higher errors (up to 3.7%) at longer lead

times (t+12 to t+24), which may be attributed to local hydrodynamic complexities and flood routing delays in the lower Red River system.

For the coastal and estuarine stations (No. 10–19), the toolkit’s ability to reproduce tidal variations and storm surge levels is also satisfactory. The forecasted maximum and minimum water levels are very close to the observed values, with errors generally within 1–4% for most stations. Slightly larger deviations occur at Ba Lat (10%) and Cua Cam (-4.2%), likely reflecting strong tidal influences and local topographic effects during the typhoon event. Overall, the results in Table 4 showed that the forecasting toolkit provides reliable and accurate predictions of both riverine and coastal water levels during extreme events. The low percentage errors across most stations and time steps indicate that the system can effectively support early warning operations and flood risk management in the Red River–Thai Binh basin during typhoon-induced floods.

Table 5: Evaluation of the forecasting results of the toolkit for Typhoon Kajiki (08/2025)

No.	Station	Observed (cm)				Forecasting (cm)				Error (%)			
		t+6	t+12	t+18	t+24	t+6	t+12	t+18	t+24	t+6	t+12	t+18	t+24
1	Quang Cu	2811	2849	2828	2798	2808	2810	2822	2827	0.1	1.4	0.2	1.0
2	Hoa Binh	964	1043	941	933	939	1005	962	923	2.6	3.7	2.3	1.0
3	Lam Son	2123	2106	2097	2091	2098	2097	2093	2089	1.2	0.4	0.2	0.1
4	Son Tay	673	701	710	698	651	698	692	691	3.2	0.4	2.5	1.0
5	Ha Noi	521	555	574	580	507	533	563	578	2.6	4.0	1.9	0.3
6	Thuong Cat	468	498	518	526	446	481	507	516	4.7	3.3	2.1	1.8
7	Gia Bay	2313	2332	2332	2301	2325	2327	2329	2305	0.5	0.2	0.1	0.2
8	Chu	820	938	830	699	824	949	822	674	0.5	1.1	1.0	3.6
9	Binh Lieu	7756	7749	7743	7738	7751	7744	7740	7734	0.1	0.1	0.0	0.1
		Hmax		Hmin		Hmax		Hmin		Hmax		Hmin	
10	Nam Dinh	294		234		290		235		1.4		0.4	
11	Truc Phuong	211		153		215		157		1.9		2.6	
12	Ben Binh	269		225		269		224		0.1		0.4	
13	Cat Khe	319		270		313		278		1.9		3.0	
14	Ba Nha	157		129		156		126		0.6		2.3	
15	Quyet Chien	300		240		293		240		2.3		0.1	
16	Trieu Duong	344		284		348		278		1.2		2.1	
17	Ba Lat	145		83		141		79		2.8		4.8	
18	Cua Cam	79		26		74		28		6.3		7.7	
19	Trung Trang	137		89		139		89		1.5		0.1	

Table 5 presents the evaluation results of the water level forecasting toolkit during Typhoon Kajiki (August 2025) across 19 observation stations. The table compares observed and forecasted water levels at multiple forecast lead times (t+6, t+12, t+18, and t+24 hours), and includes the percentage error to assess the accuracy of the model under real-time storm conditions.

For river stations (No. 1–9), the forecasting results are generally accurate, with most errors remaining below 3% for short- to medium-range forecasts. Stations such as Quang Cu, Lam Son, Gia Bay, and Binh Lieu exhibit excellent agreement between observed and simulated values, with very small deviations (average error <1%). These results demonstrate that the model successfully reproduces the hydrological response and peak timing during Typhoon Kajiki.

Slightly higher discrepancies are found at Hoa Binh, Son Tay, Ha Noi, and Thuong Cat, where errors reach up to 4–5%, particularly at $t+12$ and $t+18$ hours. These differences likely stem from localized rainfall variations and complex hydraulic routing in the lower Red River network.

For coastal and estuarine stations (No. 10–19), the toolkit also performs well in simulating tidal oscillations and storm surge levels. The forecasted maximum and minimum water levels closely match the observed values, with most stations showing errors below 3%. The highest discrepancies are observed at Cua Cam (6.3–7.7%) and Ba Lat (up to 4.8%), which are strongly affected by tidal interactions and boundary condition uncertainties during storm surge propagation. In contrast, stations such as Ben Binh, Nam Dinh, and Truc Phuong show particularly low errors (<2%), confirming good model stability along the coastal zone.

Overall, the results in Table 5 demonstrate that the forecasting toolkit performs reliably and accurately under the extreme meteorological and hydrological conditions of Typhoon Kajiki. The model effectively reproduces both riverine and coastal water levels with small prediction errors and consistent performance across different forecast horizons. This confirms its suitability for operational flood forecasting and early warning applications in the Red River–Thai Binh basin under typhoon-driven flood scenarios.

Overall, the testing results confirmed the robustness and adaptability of the toolkit under varying hydro-meteorological conditions, from widespread heavy rainfall to typhoon impacts. The integrated use of MIKE NAM, MIKE 11, and LSTM models leveraged the strengths of each approach: rainfall-runoff simulation, river-sea dynamics, and nonlinear long-term dependency learning from observed data. Remaining errors, particularly in estuarine areas, suggest that future improvements should incorporate tidal forcing, monsoon winds, and morphological dynamics to enhance forecasting accuracy in downstream coastal regions.

4. Discussion

The integrated modelling framework developed in this study, which combines hydrological (NAM), hydraulic (MIKE 11), and deep learning (LSTM) models, has demonstrated strong predictive skill for water levels in the Red River - Thai Binh Basin. The achieved performance metrics, with NSE values consistently above 0.89 and RMSE below 0.25 m at most stations, indicate that the hybrid approach significantly improves the reliability of flood forecasting compared to traditional single-model approaches.

Hydrological and hydraulic models such as NAM, MIKE 11, HEC-RAS have long been recognized as robust tools for simulating rainfall-runoff and river hydraulics (Kingma & Ba, 2015; Thai & Tri, 2019). However, their accuracy in complex river networks is often limited by uncertainties in input data, model parameters, and boundary conditions. Similar findings have been reported in the Mekong Basin, where hydrological-hydraulic coupling alone could not fully capture rapid flood dynamics (Quang et al., 2000; Le et al., 2023).

The integration of AI models into hydrological forecasting has gained momentum in recent years, particularly with LSTM networks, which can learn long-term temporal dependencies and nonlinear dynamics (Kratzert et al., 2018; Leščešen et al., 2025). In the Yangtze and Ganges basins, studies have shown that LSTM-based models outperform conceptual models during extreme flood events, especially when data sparsity is addressed (Shen, 2018; Liu et al., 2020). Our results are consistent with these findings, as the LSTM component significantly improved forecast accuracy in downstream regions affected by tidal influence areas where purely process-based models typically underperform.

In Vietnam, research on flood forecasting has largely relied on hydrological-hydraulic models without systematically integrating AI techniques (Linh et al., 2018; Thai & Tri, 2019; Loi et al., 2019; Tri et al., 2022; Hanh et al., 2024). Recent attempts to apply machine learning, such as Vector Regression (SVR), Decision Tree (DT), Random Forest (RF), Light Gradient Boosting Machine Regressor (LGBM), and Linear Regression (LR), have shown potential but often

lacked robustness during extreme events (Duy, 2023; Hanh et al., 2024). By coupling NAM–MIKE 11 with LSTM, this study bridges the methodological gap and provides evidence that AI-enhanced hybrid models can effectively complement process-based models in the Red River - Thai Binh Basin.

The results also have practical implications. During the July 2025 heavy rainfall and Typhoon Kajiki (August 2025), the hybrid system captured peak water levels more accurately than the hydrological-hydraulic setup alone, reducing underestimation errors at critical stations. This is vital for early warning systems, as even small improvements in accuracy can translate into significant benefits for disaster preparedness and floodplain management (UNDRR, 2020). The framework also allows operational agencies to integrate multiple data streams, including near-real-time rainfall forecasts and hydraulic simulations, into a single decision-support tool.

Despite these advances, some limitations remain. Data scarcity at upstream catchments continues to restrict forecast skill, as shown by slight discrepancies at Yen Bai station during flood peaks. This aligns with global findings that the effectiveness of LSTM-based models is strongly dependent on the availability of long-term, high-quality datasets (Kratzert et al., 2018). Furthermore, the current framework does not explicitly account for land-use changes or climate variability, both of which are critical for long-term flood risk assessments. Future work should incorporate remote sensing rainfall estimates, ensemble weather forecasts, and climate change scenarios to improve resilience and adaptability of the forecasting system.

Overall, this study contributes to the growing body of evidence that hybrid hydrological-hydraulic-AI models represent a promising pathway for enhancing flood forecasting accuracy and operational decision-making in data-limited regions.

5. Conclusion

This study developed and tested an integrated forecasting framework that combines the NAM hydrological model, the MIKE 11 hydraulic model, and a Long Short-Term Memory (LSTM) neural network for water level prediction in the Red River - Thai Binh Basin. The results demonstrate that the hybrid approach can effectively capture both rainfall–runoff processes and river hydraulics while leveraging the nonlinear learning capacity of deep learning. With Nash-Sutcliffe Efficiency (NSE) values up to 0.97 and Root Mean Squared Error (RMSE) below 0.25 m at most stations, the framework achieved high accuracy during both calibration - validation periods and in reproducing extreme flood events such as the July 2025 heavy rainfall and Typhoon Kajiki in August 2025. These findings underline the methodological and practical value of integrating hydrological, hydraulic, and AI-based models to enhance operational flood forecasting.

The research contributes three key advances: (i) establishing a calibrated and validated modelling chain tailored to the Red River - Thai Binh Basin, (ii) demonstrating the added value of LSTM in addressing the limitations of conventional models, particularly in tidal and downstream regions, and (iii) providing a replicable decision-support tool that can be applied to other flood-prone river basins in Southeast Asia.

Nevertheless, several limitations remain. The accuracy of forecasts is constrained by the density and quality of hydrometeorological data, particularly at upstream stations where discrepancies during flood peaks were observed. Furthermore, the current model framework does not yet fully incorporate real-time meteorological forecasts, land use dynamics, or climate change scenarios, which may affect long-term applicability.

Future research should focus on expanding the data assimilation component, integrating high-resolution weather forecasts and remote sensing products, and exploring hybrid deep learning architectures (e.g., CNN–LSTM, attention-based models) to further enhance predictive skill. In addition, the operational deployment of the tool should be tested with local disaster management agencies to assess usability, robustness, and effectiveness in real-time decision-making.

Overall, the study provides a robust methodological foundation and practical pathway for advancing flood forecasting and risk management in Vietnam and similar river basins globally.

Acknowledgment

The authors would like to express their sincere gratitude to the Ministry of Agriculture and Environment (MAE) of Vietnam for funding the Grassroots-level Scientific and Technological project entitled “Development of a water level forecasting toolkit for the Northern Delta and Midland region” (Project code: CS.2025.1) during the period of January to December 2025.

Author Contributions

Conceptualization, **T.D.Q., N.N.V., T.Q.T.T.**; Methodology, **T.D.Q., N.N.V., T.Q.T.T.**; Software, **N.N.V., T.Q.T.T.**; Calibration and validation, **T.D.Q., N.N.V., T.Q.T.T.**; Formal analysis, **N.N.V., T.Q.T.T.**; Investigation, **T.D.Q., N.N.V., T.Q.T.T.**; Data curation, **N.N.V., T.Q.T.T.**; Writing—original draft preparation, **T.D.Q., N.N.V., T.Q.T.T.**; Writing—review and editing, **T.D.Q., N.N.V., T.Q.T.T.**; Visualization, **T.D.Q., N.N.V.** All authors critically reviewed and approved the final version of the manuscript and agreed to be accountable for all aspects of the work.

Disclosure of Interest

The authors declare no conflict of interest.

References

- ADB (2020). Climate risk country profile: Vietnam. Asian Development Bank, pp. 27.
- Bates, P. D., Horritt, M. S., & Fewtrell, T. J. (2010). A Simple Inertial Formulation of the Shallow Water Equations for Efficient Two-Dimensional Flood Inundation Modelling. *Journal of Hydrology*, 387(1-2), 33–45. <https://doi.org/10.1016/j.jhydrol.2010.03.027>
- Berghuijs, W. R., Harrigan, S., Molnar, P., Slater, L. J., & Kirchner, J. W. (2019). The Relative Importance of Different Flood-Generating Mechanisms Across Europe. *Water Resources Research*, 55(6), 4582–4593. <https://doi.org/10.1029/2019WR024841>
- DHI (2011). MIKE 11: A Modelling System for Rivers and Channels - User Guide. Danish Hydraulic Institute, pp. 542. Online available: https://euroaquae.tu-cottbus.de/Semester3/LectureNotes/Module31/Rhine/Mike11/MIKE11_UserManual.pdf
- Duy, N. H. (2023). Daily Streamflow Forecasting by Machine Learning in Tra Khuc river in Vietnam. *Vietnam Journal of Earth Sciences*, 45(1), 82–97. <https://doi.org/10.15625/2615-9783/17914>
- Ebrahim, S. I. M. A., & Atalmanan, I. M. I. (2023). Determination of the Hydrological and Morphometric Characteristics Using GIS. *Civil and Environmental Engineering*, 19(1), 39–47. <https://doi.org/10.2478/cee-2023-0004>
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press, pp. 785. Online available: [http://alvarestech.com/temp/deep/Deep%20Learning%20by%20Ian%20Goodfellow,%20Yoshua%20Bengio,%20Aaron%20Courville%20\(z-lib.org\).pdf](http://alvarestech.com/temp/deep/Deep%20Learning%20by%20Ian%20Goodfellow,%20Yoshua%20Bengio,%20Aaron%20Courville%20(z-lib.org).pdf)
- Hakala, K., Addor, N., Teutschbein, C., Vis, M., Dakhlaoui, H., & Seibert, J. (2020). Hydrological Modeling for Climate Change Impact. *Encyclopedia of Water: Science, Technology, and Society*, Edited by Patricia A. Maurice, pp. 1–20. <https://doi.org/10.1002/9781119300762.wsts0062>
- Hanh, D. K., Truong, D. D., & Trinh, B. T. K. (2024). Application of Numerical Modeling and GIS for Simulating Inundation Under Dam Failure Scenarios. *Civil and Environmental Engineering*, 20(1), 233–254. <https://doi.org/10.2478/cee-2024-0019>
- Hanh, N. D., Giang, N. T., Hoa, L. X., Vinh, T. N., & Nguyen, H. D. (2024). Multi-Step-Ahead Prediction of Water Levels Using Machine Learning: A Comparative Analysis in the Vietnamese Mekong Delta. *Vietnam Journal of Earth Sciences*, 46(4), 468–488. <https://doi.org/10.15625/2615-9783/21067>
- Hu, C.; Wu, Q.; Li, H.; Jian, S.; Li, N.; & Lou, Z. (2018). Deep Learning with a Long Short-Term Memory Networks Approach for Rainfall-Runoff Simulation. *Water*, 10(11), 1543. <https://doi.org/10.3390/w10111543>

- IPCC (2021). *Climate Change 2021: The Physical Science Basis*. Cambridge University Press, 2021, pp. 2391.
- Jayapadma, J. M. M. U., Wickramaarachchi, T. N., Silva, G. H. A. C., Ishidaira, H., & Magome, J. (2018). Rainfall-Runoff Modelling Using MIKE 11 (NAM Model): A Case Study of GIN River Basin. *Proceeding of the 6th International Symposium on Advances in Civil and Environmental Engineering Practices for Sustainable Development (ACEPS-2018)*, pp. 231–238.
- Kingma, D. P., & Ba, J. (2015). Adam: A Method for Stochastic Optimization. *Proceeding of the conference paper at the 3rd International Conference for Learning Representations, San Diego*. <https://doi.org/10.48550/arXiv.1412.6980>.
- Kratzert, F., Klotz, D., Brenner, C., Schulz, K., & Hernegger, M. (2018). Rainfall-runoff modelling using Long Short-Term Memory (LSTM) networks. *Hydrology and Earth System Sciences*, 22(11), 6005–6022. <https://doi.org/10.5194/hess-22-6005-2018>.
- 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, G. S., Tran, L. T., Ho, L. H., & Park, E. (2023). F28: A Novel Coupling Strategy For 1D/2D Hydraulic Models for Flood Risk Assessment of the Mekong Delta. *Journal of Hydroinformatics*, 25(6), 2369–2388. <https://doi.org/10.2166/hydro.2023.108>.
- Leščešen, I., Tanhapour, M., Pekárová, P., Miklánek, P., & Bajtek, Z. (2025). Long Short-Term Memory (LSTM) Networks for Accurate River Flow Forecasting: A Case Study on the Morava River Basin (Serbia). *Water*, 17(6), 907. <https://doi.org/10.3390/w17060907>.
- Linh, N. T. M., Tri, D. Q., Thai, H. T., & Don, N. C. (2018). Application of a Two-Dimensional Model for Flooding and Floodplain Simulation: Case Study in Tra Khuc-Song Ve River in Viet Nam. *Lowland Technology International*, 20(03), 367–378.
- Liu, D., Jiang, W., Mu, L., & Wang, S. (2020). Streamflow Prediction Using Deep Learning Neural Network: Case Study of Yangtze River. *IEEE Access*, 8, 90069–90086. <https://doi.org/10.1109/ACCESS.2020.2993874>.
- Loi, N. K., Liem, N. D., Tu, L. H., Hong, N. T., Truong, C. D., Tram, V. N. Q., Nhat, T. T., Anh, T. N., & Jeong, J. (2019). Automated Procedure of Real-Time Flood Forecasting in Vu Gia – Thu Bon River Basin, Vietnam by Integrating SWAT and HEC-RAS Models. *Journal of Water & Climate Change*, 10(3), 535–545. <https://doi.org/10.2166/wcc.2018.015>.
- Madsen, M. (2000). Automatic Calibration of a Conceptual Rainfall-Runoff Model Using Multiple Objectives. *Journal of Hydrology*, 235(3–4), 276–288. [https://doi.org/10.1016/S0022-1694\(00\)00279-1](https://doi.org/10.1016/S0022-1694(00)00279-1)
- Nhan dan (2025). Online available: https://nhandan.vn/bao-dam-an-ninh-nguon-nuoc-luu-vuc-song-hong-song-thai-binh-post894722.html?utm_source=chatgpt.com (Assessed on 19/7/2025).
- Quang, N. H., Tuan, V. A., Hang, L. T. T., Hung, N. M., The, D. T., Dieu, D. T., Anh, N. D., & Hackney, C. R. (2020). Hydrological/Hydraulic Modeling-Based Thresholding of Multi SAR Remote Sensing Data for Flood Monitoring in Regions of the Vietnamese Lower Mekong River Basin. *Water*, 12(1), 71. <https://doi.org/10.3390/w12010071>
- Sabzipour, B., Arsenault, R., Troin, M., Martel, J. L., Brissette, F., Brunet, F., & Mai, J. (2023). Comparing a Long Short-Term Memory (LSTM) Neural Network with a Physically-Based Hydrological Model for Streamflow Forecasting Over a Canadian Catchment. *Journal of Hydrology*, 627(A), 130380. <https://doi.org/10.1016/j.jhydrol.2023.130380>
- Sahoo, B. B., Jha, R., Singh, A., Kumar, D. (2019). Long Short-Term Memory (LSTM) Recurrent Neural Network for Low-Flow Hydrological Time Series Forecasting. *Acta Geophysica*, 67(4), 1471–1481. <https://doi.org/10.1007/s11600-019-00330-1>
- Seibert, J., & Vis, M. J. P. (2012). Teaching Hydrological Modeling with a User-Friendly Catchment-Runoff-Model Software Package. *Hydrology and Earth System Sciences*, 16, 3315–3325. <https://doi.org/10.5194/hess-16-3315-2012>
- 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. *Water Science & Technology*, 82(12), 2635–2670. <https://doi.org/10.2166/wst.2020.369>
- 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>
- Tao, L., He, X., Li, J., & Yang, D. (2019). A multiscale long short-term memory model with attention mechanism for improving monthly precipitation prediction. *Journal of Hydrology*, 602, 126815. <https://doi.org/10.1016/j.jhydrol.2021.126815>
- Thai, T.H., & Tri, D.Q. (2019). Combination of Hydrologic and Hydraulic Modeling on Flood and Inundation Warning: Case Study at Tra Khuc-Ve River Basin in Vietnam. *Vietnam Journal of Earth Sciences*, 41(03), 240–251. <https://doi.org/10.15625/0866-7187/41/3/13866>

- Thai, T. H., Tri, D. Q., Anh, N. X., Hoa, V. V., Nguyen, H. V., Nhat, N. V., Tuyet, Q. T. T., Pham, H. T. T., Chung, P. H., Thang, V. V., & Thuc, T. D. (2023). Numerical Simulation of the Flood and Inundation Caused by Typhoon Noru Downstream from the Vu Gia-Thu Bon River Basin. *Sustainability*, 15(10), 8203. <https://doi.org/10.3390/su15108203>
- Thanh, N. T., Nhung, D. H., Thuy, N. T., Thuc, T. D., Thang, V. V., Tri, D. Q., & Doanh, V. V. (2025). Sensitive Analysis of WRF-Hydro's Parameters for Multi-Peak Flood Flow: A Case Study in the Ve River Basin, Vietnam. *Civil and Environmental Engineering*, 21(1), 334–348. <https://doi.org/10.2478/cee-2025-0026>
- Tri, D. Q., Thai, T. H., & Vo, V. V. (2022). Bias-Correction Data of IFS Rainfall Forecasts for Hydrological and Hydraulic Models to Forecast Flood Events. *Arabian Journal of Geosciences*, 15, 1535. <https://doi.org/10.1007/s12517-022-10801-3>
- Trinh, X. M., & Molkenthin, F. (2020). Flood Risk Assessment in the Tra Bong River Catchment, Vietnam. In: Gourbesville, P., Caignaert, G. (eds) *Advances in Hydroinformatics*. Springer Water. Springer, Singapore, pp. 575–592. https://doi.org/10.1007/978-981-15-5436-0_45
- Truong, V., Hoang, T., & Truong, D. (2025). Deep Learning for Downstream Water Level Prediction in Complex Hydrology Systems: An LSTM Approach. *Open Journal of Modern Hydrology*, 15(2), 218–232. <https://doi.org/10.4236/ojmh.2025.152014>
- UNDRR (2020). The Human Cost of Disasters: An overview of the last 20 years. United Nations Office for Disaster Risk Reduction, pp. 28. Online available: https://www.preventionweb.net/files/74124_humancostofdisasters20002019reportu.pdf?startDownload=true
- Wang, W., Liu, J., Li, C., Liu, Y., & Yu, F. (2021). Data Assimilation for Rainfall-Runoff Prediction Based on Coupled Atmospheric-Hydrologic Systems with Variable Complexity. *Remote Sensing*, 13(4), 595. <https://doi.org/10.3390/rs13040595>
- Yuan, Z., Liu, J., Liu, Y., Zhang, Q., Li, Y., & Li, Z. (2022). A Two-Stage Modelling Method for Multi-Station Daily Water Level Prediction. *Environmental Modelling & Software*, 156, 105468. <https://doi.org/10.1016/j.envsoft.2022.105468>
- Yu, Q., Tolson, B. A., Shen, H., Han, M., Mai, J., & Lin, J. (2024). Enhancing Long Short-Term Memory (LSTM)-Based Streamflow Prediction with a Spatially Distributed Approach. *Hydrology and Earth System Sciences*, 28(9), 2107–2122. <https://doi.org/10.5194/hess-28-2107-2024>
- Zhang, Y., Zhou, Z., Van Griensven Thé, J., Yang, S. X., & Gharabaghi, B. (2023). Flood Forecasting Using Hybrid LSTM and GRU Models with Lag Time Preprocessing. *Water*, 15(22), 3982. <https://doi.org/10.3390/w15223982>

How to Cite This Article

Doan Quang, T., Nguyen Van, N., & Quach Thi Thanh, T. (2026). Integrating Hydrological–Hydraulic–AI (LSTM) Models for Improved Water Level Forecasting: Red River - Thai Binh Basin. *Civil and Environmental Engineering*, Vol. 22, Issue 1, 387–412. <https://doi.org/10.2478/cee-2026-0042>