water

MDPI

Article

Deep Learning-Based Salinity Forecasting in the Vietnamese Mekong Delta: A Cung Hau Estuary Case Study

Nguyen Phuoc Cong [1], Tran Van Ty [1], Dinh Van Duy [1], Dang Thi Hong Ngoc [2], Nigel K. Downes [3], Huynh Vuong Thu Minh [3], Amit Chatterjee [4], Shamik Chakraborty [5] and Pankaj Kumar [6,*]

1 Water Resources Engineering Faculty, College of Engineering, Can Tho University, Can Tho 94119, Vietnam; npcong@ctu.edu.vn (N.P.C.); tvty@ctu.edu.vn (T.V.T.); dvduy@ctu.edu.vn (D.V.D.)
2 Faculty of Natural Resource and Environment, Kien Giang University, Kien Giang 920000, Vietnam; dthngoc@vnkgu.edu.vn
3 Water Resources Department, College of Environment and Natural Resources, Can Tho University, Can Tho 94119, Vietnam; nkdownes@ctu.edu.vn (N.K.D.); hvtminh@ctu.edu.vn (H.V.T.M.)
4 Department of Geography, Visva-Bharati, Santiniketan 731235, India; amit.chatterjee@visva-bharati.ac.in
5 Graduate School of Sustainability Studies, Global Research Centre for Advanced Sustainability Science (GRASS), University of Toyama, Toyama 930-8555, Japan; shamik@eco.u-toyama.ac.jp
6 Institute for Global Environmental Strategies, Hayama 240-0115, Japan
* Correspondence: kumar@iges.or.jp

Abstract

Salinity intrusion has become a critical threat to agricultural stability and water resource management in the Vietnamese Mekong Delta (VMD), particularly in coastal regions. This study evaluates the efficacy of the Long Short-Term Memory (LSTM) neural network, a sophisticated deep learning (DL) architecture, for predicting salinity concentrations at two monitoring stations: Hung My and Tra Vinh. Using historical salinity data, the research explores the impact of varying the lookback window from 15- to 45-day and the forecast horizons (1- to 3-day) on model performance. Experimental results demonstrate that the 15-day lookback window provides the most robust temporal context, enabling the model to achieve high predictive accuracy for short-term horizons. For 1-day forecast horizon, the model achieved Nash–Sutcliffe Efficiency (NSE) values exceeding 0.85 and low Root Mean Square Error (RMSE) at both stations. However, a progressive decline in performance was observed as the lead time extended to 3-day forecast horizon, primarily due to increased prediction uncertainty and the inherent non-linearity of estuarine dynamics. A detailed analysis of the results reveals a consistent underestimation of extreme salinity peaks, a phenomenon attributed to the smoothing effect of the Mean Squared Error (MSE) loss function and the absence of real-time exogenous inputs such as wind speed and tidal pressure. These findings provide a valuable scientific foundation for developing early warning systems, offering actionable insights for farmers and supporting evidence-based decision-making for policymakers in managing salinity risks.

Keywords: deep learning; estuarine dynamics; salinity intrusion; time-series forecasting; the Vietnamese Mekong Delta; water resource management

Check for updates

Academic Editors: Chunyan Li and Biyun Guo

Received: 25 March 2026
Revised: 12 May 2026
Accepted: 16 May 2026
Published: 20 May 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

The Mekong Delta is the largest delta in Vietnam, covering an area of approximately 3.9 million hectares and accounting for nearly 12% of the country's total landmass, and it is home to about 18 million people. This region plays a pivotal role in national food security, contributing roughly 50% of Vietnam's total rice production and 70% of its seafood export

volume. However, the VMD is currently facing severe challenges stemming from climate change, among which salinity intrusion has emerged as one of the most pressing issues [1]. In the Mekong Delta, salinity intrusion occurs primarily during the dry season (from December to May annually). During this period, the freshwater flow from the upstream Mekong River decreases, while high tides from the East Sea and the West Sea push saltwater deep into the inland [2]. In recent years, the situation has become increasingly severe for several reasons. Climate change and rising sea levels have increased average sea levels, facilitating deeper saltwater penetration [3]. Additionally, the construction of hydropower dams upstream has reduced the volume of freshwater reaching the downstream areas, weakening the river's ability to repel saltwater [1]. Furthermore, the expansion of canal systems and irrigation infrastructure has created favourable pathways for saltwater to migrate further inland [4].

Salinity intrusion poses a severe threat to freshwater availability, agricultural productivity, and ecosystem health, particularly in vulnerable deltaic regions such as the Mekong Delta in Vietnam [5,6]. This phenomenon is intensifying due to climate change challenges, rising sea levels, and upstream hydrological shifts that exacerbate salinity penetration [7]. Agriculture in the VMD accounts for a significant portion of Vietnam's economy; therefore, accurate salinity forecasting has become an urgent requirement for sustainable water resource management and the maintenance of agricultural yields [6,8].

Recent statistics indicate a sharp increase in salinity intrusion, causing substantial crop losses and directly threatening water security [6,9–11]. This salinisation process, closely linked to climate change and anthropogenic activities, is directly impacting the livelihoods of millions of people and the stability of coastal ecosystems [12,13]. Over the past decade, research has shifted significantly from traditional hydrodynamic models toward the integration of advanced Machine Learning (ML) techniques, reflecting a trend of applying data-driven approaches to improve spatial and temporal resolution in salinity prediction [14].

Addressing the challenges of developing and operating effective monitoring, surveillance, and early-warning systems is paramount. Over the past decade, salinity forecasting methodologies have advanced significantly, transitioning from traditional hydrodynamic models (such as MIKE 11 and HEC-RAS) and classical statistical methods (such as ARIMA) to advanced ML techniques [14–16]. While physics-based models remain valuable for scenario interpretation, DL models, particularly LSTM, have demonstrated strong performance in capturing the non-linear characteristics and temporal dependencies of salinity data [10,17,18]. However, researchers have also noted that pure ML models occasionally face challenges when forecasting under extreme conditions that have not appeared in the training datasets [19,20].

Long Short-Term Memory networks are increasingly optimised for real-time salinity forecasting with high accuracy, frequently achieving coefficient of determination (R^2) values above 0.9 for short-term predictions [21]. The efficacy of LSTM has been confirmed across various major estuarine systems worldwide, such as the Rhine-Meuse Delta (Netherlands) and the Qiantang Estuary (China) [22,23]. A promising new direction involves integrating satellite data (such as SMAP, SMOS, and Landsat) into LSTM models to expand monitoring coverage in coastal areas lacking physical stations [24,25]. Nevertheless, challenges regarding signal noise and the sparse temporal resolution of satellite products still require complex data calibration and fusion techniques [26,27].

Machine Learning and DL are becoming indispensable tools for forecasting salinity intrusion in the VMD. These models can learn from historical datasets, identifying complex patterns and trends to provide highly accurate salinity predictions. Duc et al. (2025) [10] conducted a comparative analysis of ML and DL model performances in predicting salinity

levels across the VMD. Their study evaluated various algorithms, including Random Forest (RF), Support Vector Machines (SVM), Artificial Neural Networks (ANN), and LSTM. The findings indicated that DL models—specifically LSTM—outperformed traditional ML models, particularly in long-term time-series forecasting. Similarly, Nguyen et al. (2023) [22] developed an LSTM model specifically for salinity forecasting at the Dai Ngai monitoring station. As a specialised type of Recurrent Neural Network (RNN), LSTM is designed to learn long-term dependencies within time-series data, making it exceptionally well-suited for salinity modelling. The results demonstrated high precision, suggesting that LSTM can be effectively applied to both short-term and medium-term salinity forecasting at other regional stations. In a localised study, Dang et al. (2025) [5] employed diverse ML algorithms for short-term salinity prediction in the coastal areas of Soc Trang Province. By comparing Linear Regression, Decision Trees, Random Forest, Gradient Boosting, and Neural Networks, the research highlighted that ensemble models (such as Random Forest and Gradient Boosting) achieved the highest accuracy, successfully predicting salinity levels 1–3 days in advance with minimal error margins. Furthermore, Lam (2023) [28] utilised statistical modelling to predict salinity intrusion for coastal rice-growing regions in the VMD. This model integrated multiple variables, including upstream discharge, tidal peaks, sea-level rise, and climatic factors. The high accuracy of the results suggests that such models provide vital decision support for farmers, enabling them to align their cropping schedules with fluctuating salinity conditions.

The Cung Hau estuary is one of the most critical gateways of the Tien River, situated at the border between Tra Vinh and Ben Tre provinces, where it is directly and powerfully influenced by the East Sea's tidal regime [9]. Monitoring results at the Hung My station (23 km from the sea) identify it as a highly volatile brackish–freshwater transition zone, frequently reaching 10–15 g/L during the peak dry season. Meanwhile, the Tra Vinh station (located further inland, 35 km from the sea) was historically a stable freshwater zone but now records sudden salinity spikes of up to 10–13 g/L [9]. Analysis reveals that the growth rate of minimum salinity (Smin) at Tra Vinh (6.19%/year) and Hung My (2.95%/year) is increasing rapidly, signalling that the river system is losing its natural capacity to flush out salt due to the severe decline in upstream discharge.

Salinity intrusion poses a severe threat to freshwater availability, agricultural productivity, and ecosystem health. Currently, traditional soil and water analysis methods often rely on manual sampling and laboratory analysis, which are time-consuming, costly, and difficult to implement at scale [20]. In response to these challenges, this study was conducted to comprehensively evaluate the applications of LSTM models for salinity forecasting in the VMD. Therefore, this study aims to evaluate the performance of the LSTM model in predicting salinity concentrations at the Hung My and Tra Vinh stations in the Cung Hau estuary. Specifically, the research investigates the influence of different lookback windows and lead times to optimize prediction accuracy for short-term salinity fluctuations.

2. Methods

2.1. Study Area

The study focuses on the coastal region of the VMD, specifically the Cung Hau estuary (Figure 1). Situated between the Tien and Hau Rivers, the two principal distributaries of the Mekong River, the estuary represents a critical transition zone between fluvial and marine processes. The area is characterised by low-lying topography and a dense network of interconnected canals, making it highly susceptible to salinity intrusion.

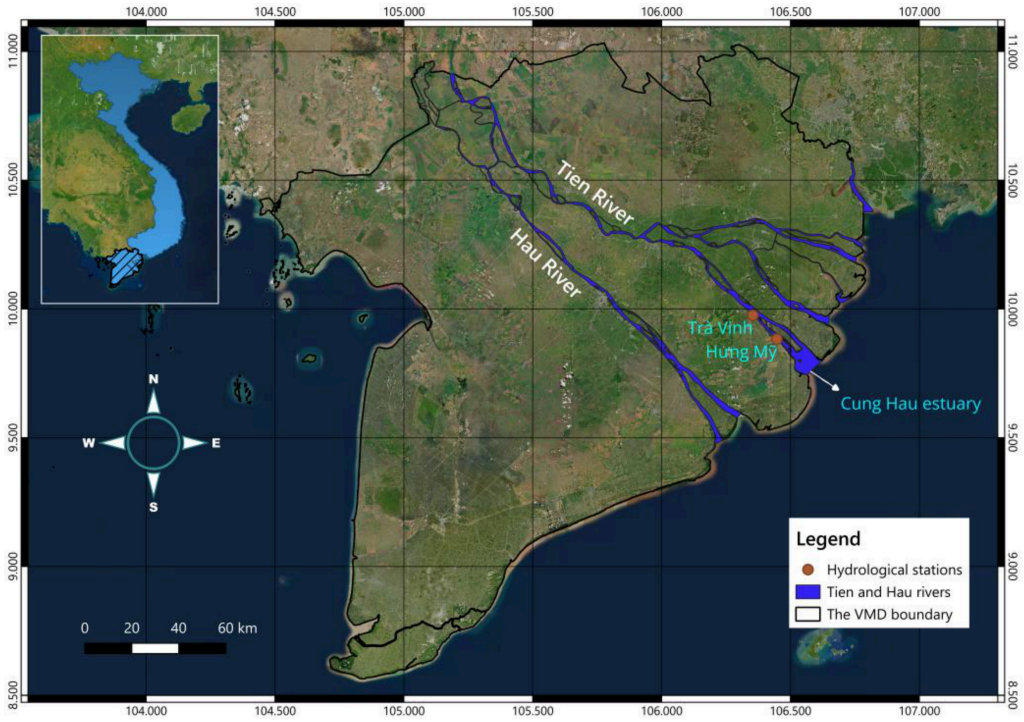


Figure 1. Location of the Cung Hau estuary within the Vietnamese Mekong Delta, showing the positions of the Hung My and Tra Vinh salinity monitoring stations along the estuarine gradient.

Salinity dynamics in the VMD are particularly important for agricultural production, especially rice cultivation, which is highly sensitive to salinity. Salinity levels exceeding 4 g/L can cause significant crop damage, resulting in reduced yields or complete crop failure. In this region, salinity intrusion occurs primarily during the dry season (December to May), when upstream freshwater discharge declines and tidal forces from the East Sea and West Sea drive saline water further inland [2].

To address this challenge, a network of salinity monitoring stations has been established across the VMD, particularly along major rivers and estuarine channels in high-risk coastal zones. According to the [29], these stations are strategically located near river mouths to monitor the inland extent of salinity intrusion, typically ranging from 30 to 52 km.

The Cung Hau estuary serves as a major conduit for tidal exchange between the South China Sea and the inland river system. It is governed by a mixed semi-diurnal tidal regime, which enhances the landward transport of saline water, particularly during periods of low upstream discharge in the dry season. Within this system, the Hung My and Tra Vinh monitoring stations capture distinct positions along the estuarine gradient, enabling analysis of spatial variability in salinity dynamics and providing a robust basis for time-series forecasting.

2.2. Long Short-Term Memory Neural Networks

Long Short-Term Memory neural networks were utilised to predict salinity based on historical data. These networks represent a specific type of recurrent neural network (RNN) architecture. LSTM networks were proposed by [30] to solve the problem of vanishing gradients in standard RNNs. In addition, an LSTM differs from feedforward neural networks because it can model the sequential nature of time-series data. Each LSTM network is composed of a memory cell and four multiplicative gates. These gates include the input, forget, candidate cell, and output gates. Each gate controls the flow of information into, within, and out of the memory cell.

The operation of an LSTM unit at the time step t can be summarised as follows.

$$f_t = \sigma\Big(W_f x_t + U_f h_{t-1} + b_f\Big), \quad (1)$$
$$i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i),$$
$$\bar{c}_t = \tanh(W_c x_t + U_c h_{t-1} + b_c),$$
$$c_t = f_t \odot c_{t-1} + i_t \odot \bar{c}_t,$$
$$o_t = \sigma(W_o x_t + U_o h_{t-1} + b_o),$$
$$h_t = o_t \odot \tanh(c_t)$$

where the input vector at time t is represented as x_t, h_t represent the hidden state, c_t represent the cell state, and W, U represent the weight matrices and b represent the bias terms, $\sigma(\cdot)$ represents the sigmoid activation function and $\odot$ represents element-wise multiplication.

The LSTM model was implemented using univariate input data, namely, historical salinity time series. The dataset was divided using a chronological split, with the periods 2000–2016, 2017–2020, and 2021–2024 used for training, validation, and testing, respectively. The salinity values were normalised using z-score standardisation prior to input into the model. The LSTM model is defined as a sequence-to-one regression problem, with a fixed-length window of input variables used to predict the residual salinity at the next time step. The input sequences were constructed using a sliding-window approach, allowing the model to learn temporal relationships across multiple preceding time intervals. In Table 1, the LSTM model comprises two LSTM layers followed by a dense layer for the final prediction. A 20% dropout regularisation penalty was applied between layers to improve robustness and generalisation. The model was trained using the Adam optimizer with a learning rate of 0.001 and Mean Squared Error (MSE) as the loss function. Early stopping was applied based on validation loss, with a patience of 10 epochs. The learning rate was reduced by a factor of 0.5 if the validation loss did not improve for 5 consecutive epochs, with a minimum learning rate of 1×10^{-6}. A batch size of 128 was used during training.

Table 1. Shapes of learnable parameters of the LSTM model.

Layer	Parameter	Shape
1st LSTM layer	$\left(W_f,\ W_i,\ W_c,\ W_o\right)$	(1, 128) each
	$\left(U_f,\ U_i,\ U_c,\ U_o\right)$	(128, 128) each
	$\left(b_f,\ b_i,\ b_c,\ b_o\right)$	(128) each
2nd LSTM layer	$\left(W_f,\ W_i,\ W_c,\ W_o\right)$	(128, 64) each
	$\left(U_f,\ U_i,\ U_c,\ U_o\right)$	(64, 64) each
	$\left(b_f,\ b_i,\ b_c,\ b_o\right)$	(64) each
Dense layer	(W_d)	(64, 1)
	(b_d)	-1

2.3. Performance Evaluation of the Model

Four complementary performance metrics were used to evaluate model skill: RMSE, MAE, R^2, and NSE. Together, these metrics quantify both the magnitude of prediction error (RMSE, MAE) and the goodness-of-fit relative to observed variability (NSE, R^2). Model performance was therefore assessed by minimizing RMSE and MAE and maximizing NSE and R^2. These measures are widely applied in hydrological and salinity forecasting studies [31–33].

RMSE captures the typical magnitude of prediction errors and is defined as shown in Equation (2).

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \quad (2)$$

where y_i and $\hat{y}_i$ are the observed and predicted values of the water salinity, respectively, and n is the number of data points.

The MAE is a measure of the mean of the errors between the predicted value and the observed value. The value of MAE is calculated as the sum of the errors divided by the sample size and is defined by Equation (3) [34].

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \quad (3)$$

Legates (1999) [34] introduced the R^2 to measure the proportion of the observed data's variance explained by the model. R^2 values range from 0 to 1, with 1 indicating a perfect correlation. It is calculated as shown in Equation (4).

$$R^2 = \left[\frac{\sum_{i=1}^{n}(y_i - \overline{y})(y_i - \overline{\hat{y}})}{\sqrt{\sum_{i=1}^{n}(y_i - \overline{y})^2}\sqrt{\sum_{i=1}^{n}(y_i - \overline{\hat{y}})^2}}\right]^2 \quad (4)$$

where $\overline{y}$ is the mean of observed values and $\overline{\hat{y}}$ is the mean of the predicted data.

The NSE, introduced by [35], is the ratio of the model prediction skill to the mean observed data. It is given by Equation (5).

$$NSE = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \overline{y})^2} \quad (5)$$

where an NSE of 1 denotes a perfect fit between observed and simulated values. An NSE value close to 0 suggests that the model's predictive power is equivalent to the observations' mean, often characterised by substantial discrepancies in process simulation despite a reliable overall mean. Values that descend significantly below zero indicate that the model is not credible.

3. Results

3.1. Descriptive Analysis of Daily Mean Salinity

The daily mean salinity data at both Hung My and Tra Vinh stations exhibit a strong seasonal signal, primarily driven by the interaction between river discharge and tidal intrusion. During the dry season (typically from January to May), both stations show a marked increase in salinity levels as freshwater discharge from upstream decreases. Figure 2 illustrates that salinity fluctuations at both stations exhibit a clear periodic pattern, reflecting the seasonal dynamics of coastal deltaic systems. Salinity levels typically peak during the dry season and decline sharply to near 0 g/L during the flood season, when upstream freshwater discharge is substantial. Notably, the fluctuation amplitude at the Hung My station (Figure 2a) is significantly higher than at the Tra Vinh station (Figure 2b), with several peak salinity events exceeding 15 g/L, particularly during major salinity intrusion periods such as 2015–2016 and 2019–2020. This indicates that Hung My is more directly influenced by salinity intrusion processes compared to Tra Vinh within the same river system. A higher frequency of extreme salinity peaks is observed in the recent decade (2014–2024) compared to the early 2000s. This trend underscores the combined effects of climate change, sea-level rise, and reduced upstream freshwater discharge. At the Tra Vinh

station, although the average salinity concentration is lower, salinity intrusion events tend to be more prolonged, with values typically ranging between 5 and 10 g/L. This poses significant challenges for freshwater abstraction for both agricultural and domestic use. These historical datasets provide a critical foundation for training deep learning models, such as LSTM, to forecast abnormal salinity episodes and support the development of timely and sustainable mitigation strategies.

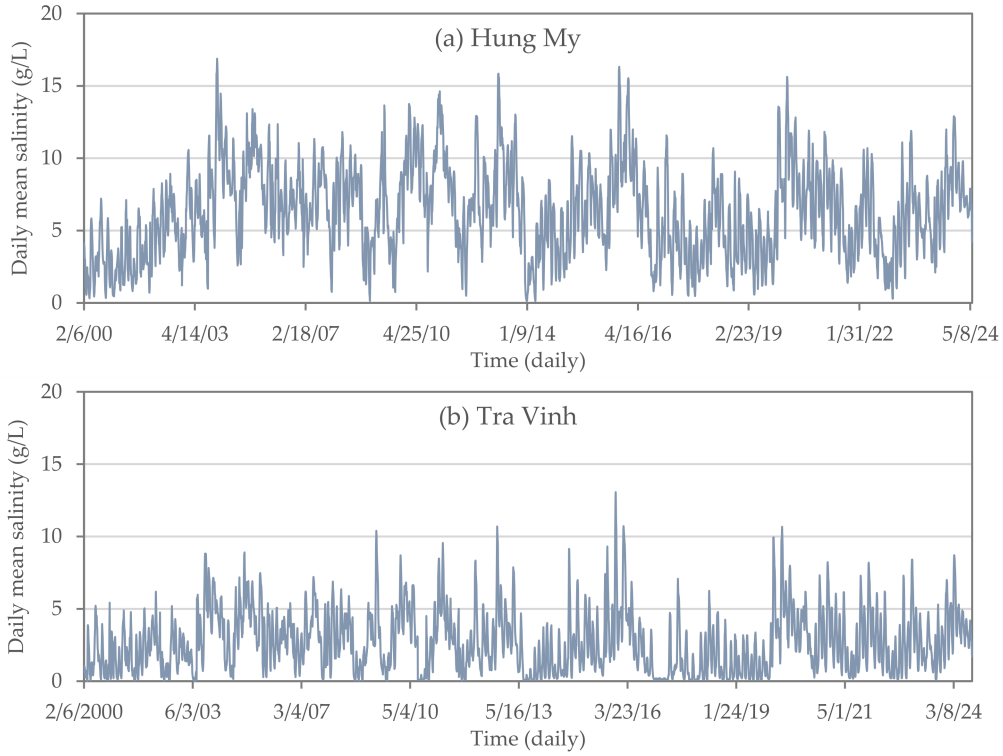


Figure 2. Time series of daily mean salinity concentrations (g/L) at (**a**) Hung My and (**b**) Tra Vinh monitoring stations from 2000 to 2024.

Figure 3 shows that the Hung My station has a significantly more severe level of salinity intrusion than the Tra Vinh station. The mean value at Hung My reached 6.311 g/L, more than 2.3 times higher than Tra Vinh (2.711 g/L). Similarly, the maximum value at Hung My also reached 16.882 g/L, while Tra Vinh only stopped at 13.067 g/L. These results are consistent with their geographical distributions, as Hung My is situated closer to the estuary mouth than the Tra Vinh station.

Observing the boxplot for Hung My station, it is evident that it is positioned higher and is larger, indicating that most daily salinity values there are concentrated at high levels (approximately 4 g/L to 8.5 g/L). Conversely, the boxplot of Tra Vinh station is lower and narrower, suggesting that salinity generally remains low (below 4 g/L) for most of the year. Median and mean salinity at Hung My are close together, reflecting a fairly balanced data distribution across the main value range, whereas at Tra Vinh, these values are concentrated near the bottom of the plot.

Although the Hung My station has higher salinity, the Tra Vinh station has a much larger coefficient of variation (CV) (Table 2). Tra Vinh's CV is 0.78, significantly higher than Hung My's 0.49. This indicates that although Tra Vinh has a lower salinity baseline, changes between days or seasons are more abrupt and unpredictable. Tra Vinh exhibits numerous outliers above the whiskers, ranging from 8 g/L to over 13 g/L, suggesting unusual salinity intrusion occurs more frequently than its average baseline.

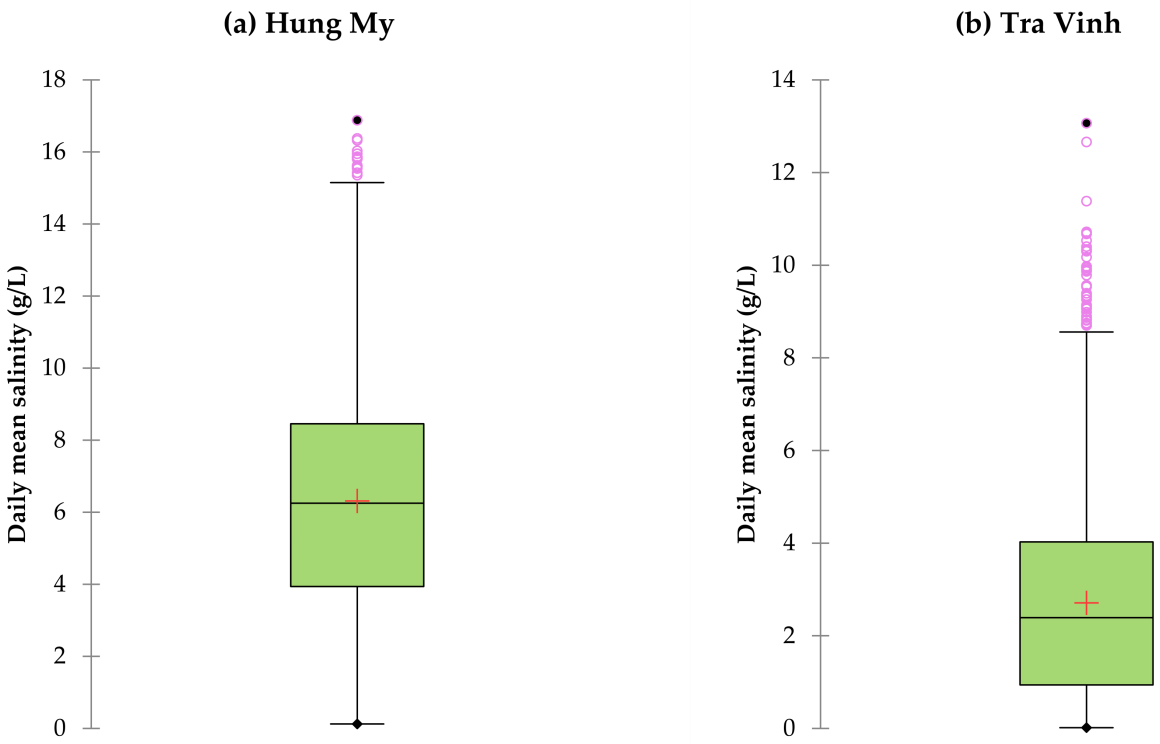


Figure 3. Boxplots of daily mean salinity for Hung My (**a**) and Tra Vinh (**b**), during 2000 and 2024. The box plot shows the average daily salinity at Hung My station. The horizontal line in the blue box denotes the median, while the red plus sign (+) and pink circles denote the mean and outliers, respectively.

Table 2. Descriptive statistics of daily mean salinity characteristics for the 2000 to 2024 periods.

Variable	Maximum	Mean	Std. Deviation	Coeff. Variation
Daily mean salinity_Hung My (g/L)	16.882	6.311	3.121	0.49
Daily mean salinity_Tra Vinh (g/L)	13.067	2.711	2.115	0.78

This comparison suggests that the Hung My station faces a persistent and prolonged risk of salinisation, putting significant pressure on the freshwater supply for agriculture. Meanwhile, although the Tra Vinh station has a lower average salinity, the risk comes from extreme salinity anomalies with extremely large variations compared to normal conditions. Modelling in Tra Vinh is therefore more challenging due to its high variability (CV), requiring the model to be able to keep up with the frequent outliers in historical data.

3.2. Long Short-Term Memory Forecasting Performance

The salinity prediction models at Hung My station shows good to excellent performance, with high correlation coefficients and NSE values typically ranging from 0.36 to 0.95 (Table 3). The RMSE and MAE exhibit a gradually increasing trend as the forecast horizon extends from 1- to 3-day. This reflects increasing uncertainty and reduced predictive signal strength at longer forecast horizons. It should be noted that the 45-day lookback window achieves the best performance with the lowest RMSE across all forecast horizons (0.63 for the 1-day forecast horizon). In contrast, the 60-day lookback window performs the poorest, suggesting that a longer historical period is insufficient for the model to capture the complex patterns of salinity fluctuations. Furthermore, increasing the lookback window to 60 days does not improve accuracy and results in a slight increase in error. Among the

tested configurations, the 15-day lookback window provides the best overall performance, particularly for the 1- and 2-day forecast horizons, with NSE values exceeding 0.83. Although prediction error increases at the 3-day forecast horizon, model performance remains within an acceptable range (NSE > 0.6), indicating that the model can support short-term salinity intrusion forecasting and early warning applications in the region.

Table 3. Testing set performance (2021–2024) for Hung My, segmented by forecast horizons (1–3 days) and lookback windows (15-, 30-, 45-, and 60-day).

Lookback Window (day)	Forecast Horizon (day)	Number of Epoch	RMSE	MAE	R^2	NSE
15	1	101	0.65	0.51	0.94	0.94
	2	57	1.11	0.86	0.83	0.83
	3	50	1.53	1.18	0.68	0.68
30	1	81	0.66	0.52	0.94	0.94
	2	43	1.26	0.98	0.78	0.78
	3	37	1.63	1.27	0.63	0.63
45	1	80	0.63	0.49	0.95	0.95
	2	40	1.33	1.03	0.68	0.68
	3	29	1.89	1.44	0.36	0.36
60	1	29	1.16	0.92	0.73	0.73
	2	31	1.41	1.13	0.61	0.61
	3	18	1.76	1.41	0.39	0.39

Figure 4 illustrates the correlation between observed and simulated daily mean salinity at Hung My station across three forecast horizons: (a) 1-day, (b) 2-day, and (c) 3-day ahead, utilising a 15-day lookback window. Similarly, Figure A1 presents the observed and predicted mean salinity at the same station during the testing period, evaluated across 1-, 2-, and 3-day forecast horizons using 30-day, 45-day, and 60-day lookback windows, respectively. Overall, the data points are closely clustered along the 1:1 diagonal, indicating that the DL model has a strong ability to accurately simulate real-world salinity dynamics. As the forecast horizon extends from 1- to 3-day, a slight increase in dispersion is observed, particularly in the lower concentration range (0–4 g/L). Nevertheless, the model captures the general magnitude of peak salinity levels, although extreme values are often underestimated (6–12 g/L). This robustness in capturing extreme values is crucial for early warning systems, despite the inherent increase in uncertainty associated with longer-term environmental fluctuations.

Figure 5 illustrates the evolution of RMSE during the training and validation processes across epochs for the 1-day ahead forecast scenario. Overall, across all four lookback windows, the loss values decrease sharply over the first 10 to 20 epochs before stabilising. This indicates that the neural network rapidly captured the core patterns of the salinity data. The smooth convergence of the learning curves demonstrates appropriate hyperparameter selection, enabling the model to reach an optimal state without significant fluctuations in error.

A noteworthy observation is that the validation error curve (Validation RMSE) remains close to and slightly below the training error curve (Training RMSE) throughout training. This behaviour can be attributed to the use of dropout regularization, which reduces model capacity during training but is not applied during validation. The similar convergence

patterns of both curves indicate stable optimisation and no strong indication of overfitting. The stable gap between the two curves suggests high model robustness, ensuring that salinity forecasts will be comparable in reliability when applied to real-world data outside the training set.

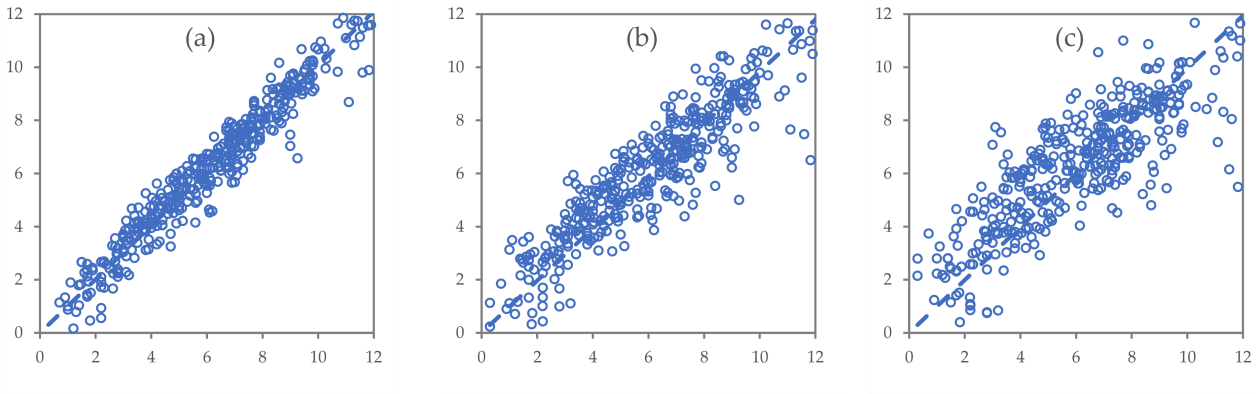


Figure 4. Scatter plots of observed vs. predicted daily mean salinity at Hung My station during the testing period. Results are shown for three forecast horizons: (**a**) 1-day, (**b**) 2-day, and (**c**) 3-day, utilizing a 15-day lookback window. Horizontal axis denotes the daily mean salinity observation while Vertical axis denotes the daily mean salinity simulation.

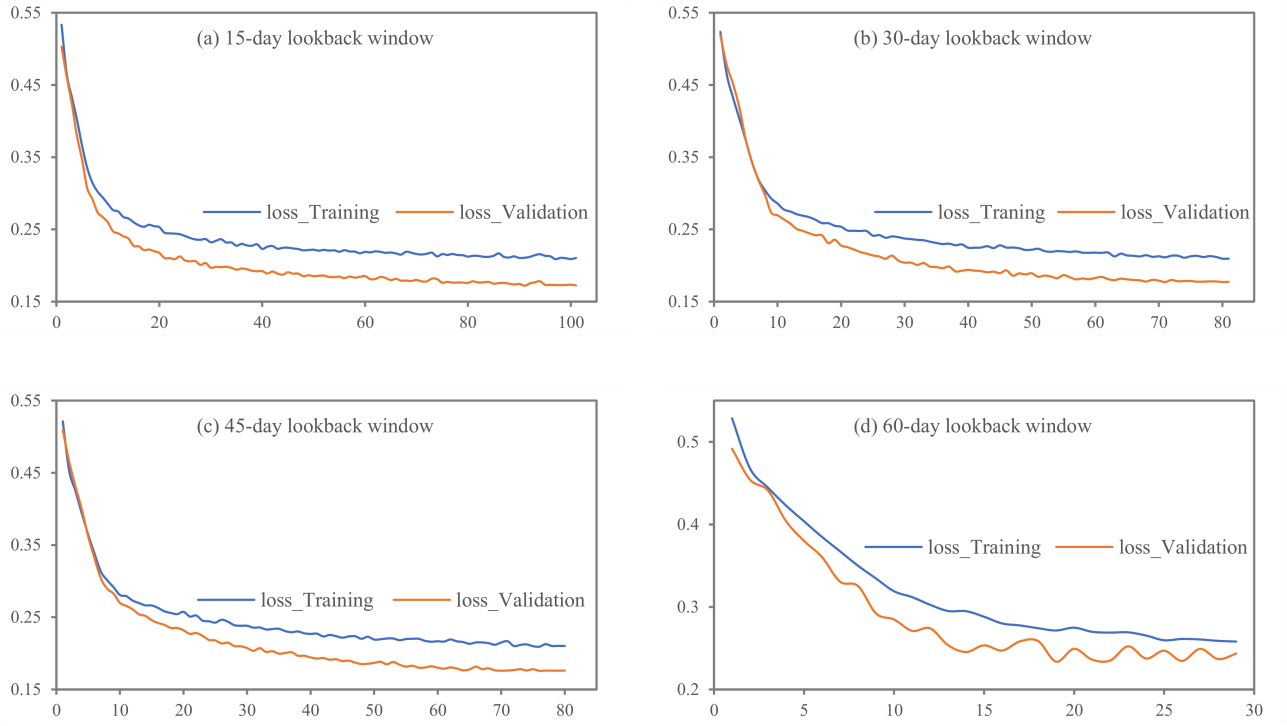


Figure 5. Training and validation RMSE for Hung My station over epochs for a 1-day forecast horizon across four lookback windows: (**a**) 15-day, (**b**) 30-day, (**c**) 45-day, and (**d**) 60-day. The horizontal axis represents Epoch and vertical axis represents Loss (RMSE).

When comparing the four forecast scenarios, the 15-day (a) and 30-day (b) lookback windows exhibit a deeper and more stable downward trend in error compared to the long-term 60-day lookback window (d). Specifically, in the 15-day, the validation RMSE curve reaches the lowest level and asymptotically stabilizes toward the end of the epoch axis. This suggests that providing sufficient historical context (15–30 days) enables the LSTM model to better capture long-term characteristics and the periodicity of salinity

intrusion. This result reinforces the hypothesis that an adequate historical data length is a prerequisite for enhancing the accuracy of time-series forecasting models in hydrology.

Figure 6 compares observed and predicted daily mean salinity at the Hung My station across three forecast horizons (1-, 2-, and 3-day). Overall, the model, using a 15-day look-back window, captures the temporal dynamics of salinity reasonably well, following the main fluctuation patterns from low-salinity periods to peak salinity intrusion events (reaching approximately 14 g/L). The close agreement between the observed and predicted time series indicates that the model can capture the dominant variability in salinity dynamics at the study site. Regarding accuracy relative to the forecast horizon, the 1-day (a) yields the best results, with close alignment between the two datasets. In this scenario, errors in both phase and amplitude are minimal, indicating that the 15-day historical sequence provides sufficient information for the model to deliver reliable short-term predictions. This high level of precision is crucial for providing immediate agricultural advisories to local farmers.

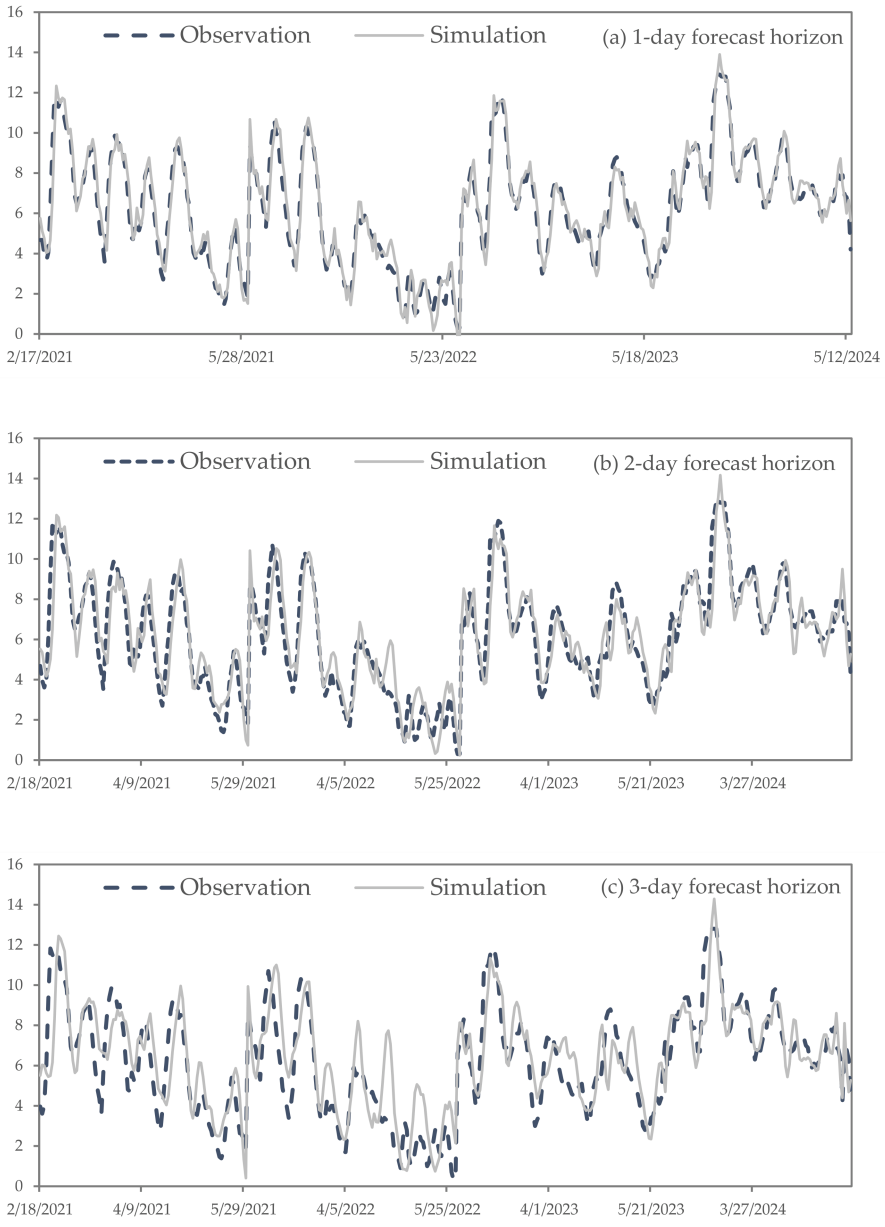


Figure 6. Comparison between observed and predicted daily mean salinity at Hung My station using a 15-day lookback window at forecast horizons of (**a**) 1-day, (**b**) 2-day, and (**c**) 3-day.

However, as the forecast horizon extends to 2 days (b) and 3 days (c), discrepancies become more pronounced, particularly at peak values. In panel (c), the model occasionally exhibits slight overestimation or a minor phase lag relative to observations. This behaviour is common in time-series forecasting, where predictive performance decreases as the forecast horizon increases due to growing uncertainty and the influence of unobserved external drivers, resulting in reduced sensitivity to abrupt fluctuations. In summary, despite a slight decline in accuracy as the forecast horizon increases, the model maintains a high level of reliability across all three scenarios. By preserving the overall variation trends, the model enables managers to grasp salinity intrusion developments up to three days in advance with an acceptable margin of error. These results affirm the effectiveness of the data-driven approach and the model's potential for early-warning systems to mitigate the impacts of drought and salinity at the Hung My station.

The model's performance at Tra Vinh station demonstrates high predictive accuracy, particularly for short-term horizons (Table 4). For a 1-day forecast, the model achieves an R^2 and NSE ranging from 0.94 to 0.95, with the lowest RMSE (0.63) observed at a 15-day lookback window. However, as the forecast horizon extends to 3 days, performance declines consistently across all lookback configurations, with RMSE values increasing to over 1.89. This trend of accuracy degradation over time is a standard characteristic of time-series forecasting due to increasing uncertainty and weakening temporal dependency at longer forecast horizons.

Table 4. Testing set performance (2021–2024) for Tra Vinh, segmented by forecast horizons (1–3 days) and lookback windows (15-, 30-, 45-, and 60-day).

Lookback Window (day)	Forecast Horizon (day)	Number of Epoch	RMSE	MAE	R^2	NSE
15	1	77	0.42	0.32	0.96	0.96
	2	62	0.83	0.63	0.82	0.82
	3	42	1.16	0.84	0.66	0.66
30	1	87	0.49	0.36	0.93	0.93
	2	33	1.03	0.79	0.70	0.70
	3	23	1.43	1.11	0.42	0.42
45	1	66	0.47	0.35	0.91	0.91
	2	39	0.91	0.68	0.68	0.68
	3	30	1.42	1.07	0.23	0.23
60	1	65	0.87	0.54	0.67	0.67
	2	34	1.29	0.93	0.29	0.29
	3	19	1.48	1.15	0.07	0.07

Compared with Hung My station, Tra Vinh generally exhibits lower error margins (lower RMSE and MAE), indicating stronger predictive performance on this dataset. While Hung My shows relatively stable performance across different lookback windows, Tra Vinh appears more sensitive to the choice of input history. Specifically, performance at Tra Vinh is strongest for lookback windows in the range of 15- to 45-day, whereas Hung My maintains good performance even with shorter windows. This suggests that salinity dynamics at Tra Vinh may be influenced by more complex seasonal or tidal processes that require a broader historical context to capture effectively Regarding the model configuration, the number of epochs required for convergence at Tra Vinh shows significant variability, peaking at 77 epochs for the 15-day lookback (1-day horizon) before dropping sharply for longer horizons. This indicates that while the model struggles more with 3-day predictions, it reaches a

steady state faster, likely due to the loss of strong linear correlations in long-term data. Compared to Hung My, the training process for Tra Vinh appears less computationally intensive for short-term windows, reflecting a less volatile data structure that necessitates more rigorous optimisation to reach peak performance.

Figure 7 illustrates the correlation between the observed and simulated salinity values at Tra Vinh station through scatter plots for three forecast horizons: 1, 2, and 3 days. In the 1-day forecast scenario (a), the data points are tightly clustered around the linear regression line, indicating a very high R^2 and minimal prediction error. The prediction error gradually increases as the forecast horizon extends. Similarly, Figure A2 presents the observed and predicted mean salinity at the same station during the testing period, evaluated across 1-, 2-, and 3-day forecast horizons using 30-day, 45-day, and 60-day lookback windows, respectively. For the 2-to-3-day forecast horizons, the model exhibits more pronounced error growth at the Tra Vinh station than at the Hung My station.

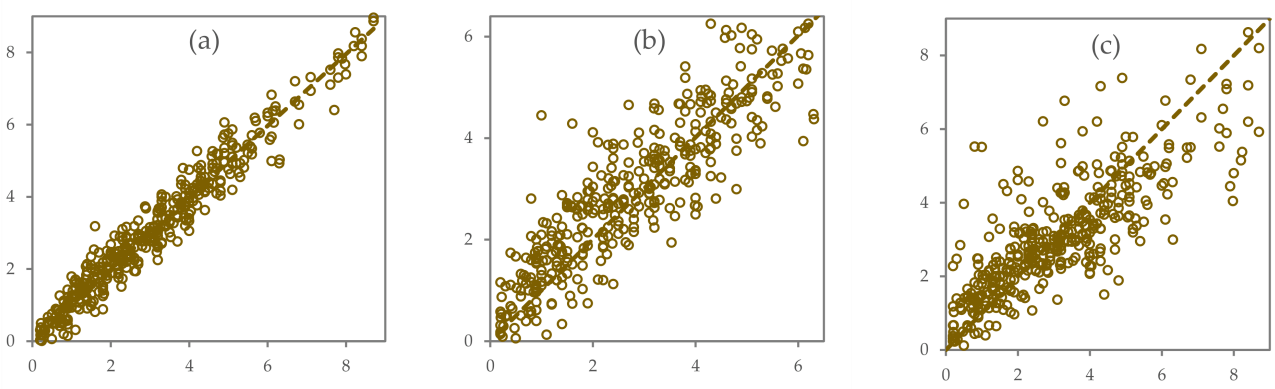


Figure 7. Scatter plots of observed vs. predicted daily mean salinity at Tra Vinh station during the testing period. Results are shown for three forecast horizons: (**a**) 1-day, (**b**) 2-day, and (**c**) 3-day, utilising a 15-day lookback window. The horizontal axis denotes the daily mean salinity observation, while the vertical axis denotes the daily mean salinity simulation.

As the forecast horizon increases to 2 days (b) and 3 days (c), the data points tend to disperse further from the regression line, reflecting an increase in prediction error over time. Particularly in the 3-day scenario, within the 6–9 high-salinity range, points more frequently fall below the symmetry line. This suggests that the model tends to underestimate actual values when encountering extreme salinity peaks, likely due to increasing uncertainty and the influence of unobserved exogenous drivers (e.g., tidal dynamics and upstream discharge) at longer forecast horizons.

Overall, despite a slight decline in accuracy as the forecast horizon is extended, the data points across all three scenarios remain closely aligned with the main regression axis and show no significant systematic bias. The slope of the trend line, which remains near 45°, confirms the model's reliable performance and its ability to accurately reflect salinity intrusion dynamics at the Tra Vinh station. These results indicate that the model is suitable for salinity monitoring and early warning in the study area. Comparing the scatter plots of the two stations, it is evident that the prediction model at Hung My station, which covers a broader salinity range (up to 12 g/L) than at Tra Vinh (below 9 g/L), remains reliable. While Tra Vinh station exhibits a very tight convergence of data points around the regression line, particularly in the 1-day forecast horizon, Hung My station shows a slightly wider scatter from the very first scenario, reflecting more complex flow dynamics at this specific location. However, a key similarity is that both stations experience an underestimation of actual values in the high-salinity ranges as the forecast horizon extends to 3-day; nonetheless, the trend line

slope at Hung My station retains a strong balance, confirming the adaptability of the 15-day lookback window to the varying hydrological characteristics of both research stations.

Figure 8 illustrates the progression of RMSE values for both training loss and validation loss across epochs for a 1-day forecast horizon, using four different lookback windows: (a) 15-day, (b) 30-day, (c) 45-day, and (d) 60-day. Generally, all four plots exhibit a rapid, sharp decline in error within the first 20 epochs, after which they stabilise and converge. This demonstrates the model's ability to quickly learn key features from the input data and indicates that the weight optimisation process is highly efficient.

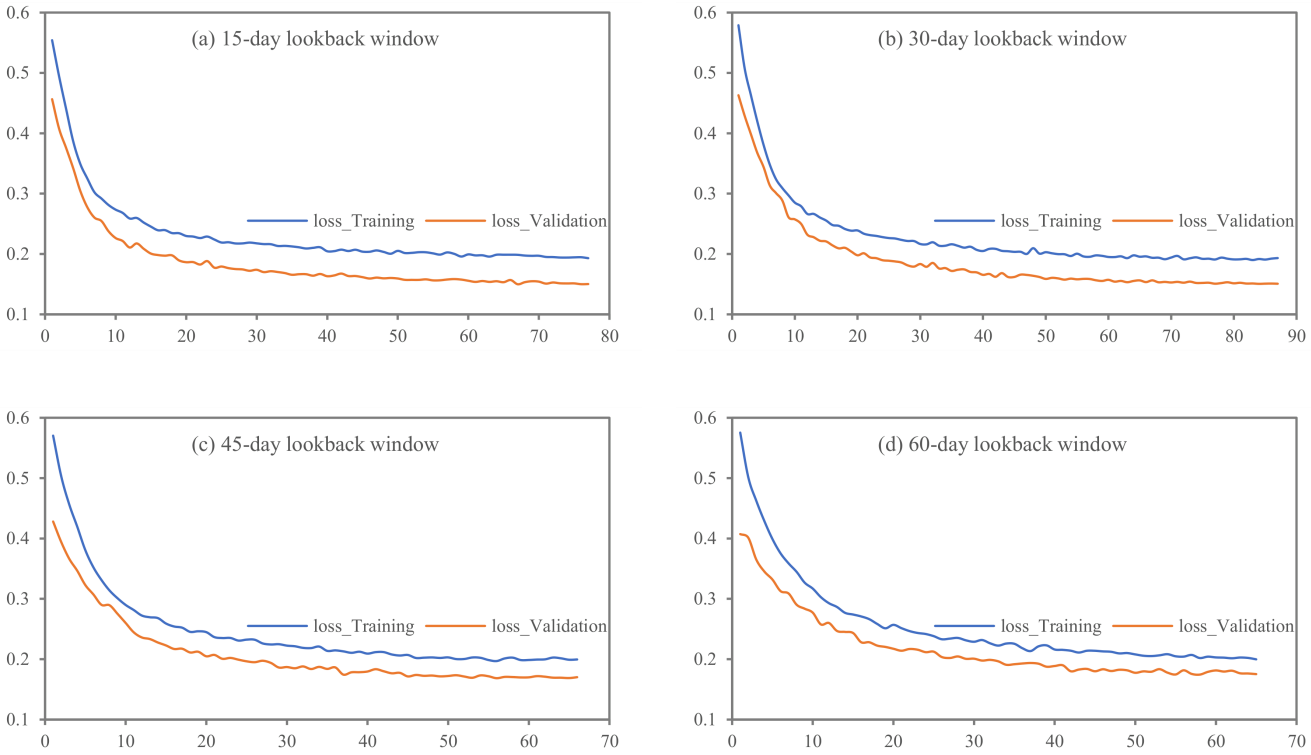


Figure 8. Training and validation RMSE for Tra Vinh station over epochs for a 1-day forecast horizon across four lookback windows: (**a**) 15-day, (**b**) 30-day, (**c**) 45-day, and (**d**) 60-day. The horizontal axis represents Epoch, and the vertical axis represents Loss (RMSE).

A notable positive aspect is that the training and validation loss curves maintain a narrow gap and run parallel toward the end of training. The fact that the validation loss remains below or closely follows the training loss indicates that the model is not suffering from overfitting; in other words, the model possesses strong generalisation capabilities on unseen data. The stability of the curves after the 40th epoch confirms that the number of epochs set is sufficient for the model to reach an optimal state without wasting computational resources.

When comparing the different lookback windows (a) 15-day and (b) 30-day, they show smoother convergence curves and achieve lower final error levels compared to the 60-day lookback window. This suggests that providing a sufficiently long historical dataset helps the model capture more stable hydrological cycles, thereby minimising instantaneous noise during learning. This result further reinforces why the 15-day lookback window was prioritised in previous analyses to achieve an ideal balance between accuracy and computational cost.

Overall, the forecasting model demonstrates strong capability in tracking actual salinity fluctuations at the Tra Vinh station across all three forecasting horizons (1-day, 2-day, and 3-day). The simulated values remain largely in phase with the observed data throughout

the 2020–2024 period (Figure 9). This close correlation confirms the model's high reliability and suggests that the algorithmic structure is well-suited to the hydrological characteristics of the study area.

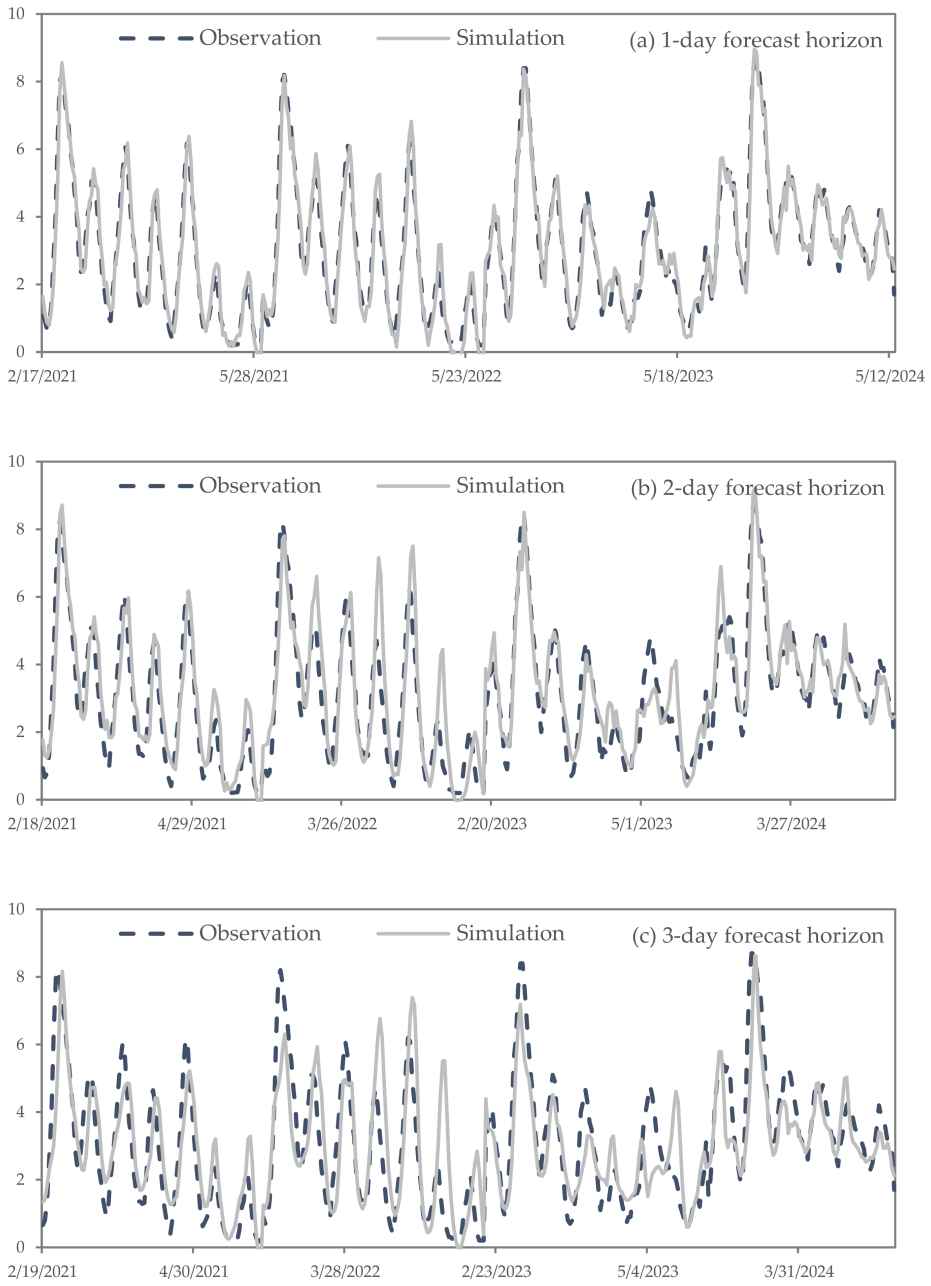


Figure 9. Comparison between observed and predicted daily mean salinity at Tra Vinh station using a 15-day lookback window at forecast horizons of (**a**) 1-day, (**b**) 2-day, and (**c**) 3-day.

The model exhibits excellent performance in simulating low-to-moderate salinity levels. However, minor discrepancies begin to emerge during extreme events, particularly at peak salinity levels exceeding 6 g/L. Specifically, in certain periods, the predicted values tend to be slightly higher or lower than the actual observations, though these deviations remain marginal. Despite the high amplitude and rapid fluctuations in salinity in Tra Vinh, the model remains stable and exhibits no significant phase lag. The RMSE values obtained in this study are consistent with those reported in previous research. For example, the model achieved an RMSE of 0.42 for the 1-day forecast horizon at the upstream Tra Vinh station, slightly higher than the upstream performance reported by Duc et al. (2025) [10]

(RMSE ≈ 0.25–0.30). At the Hung My station, located closer to the estuary mouth, the RMSE was 0.65. This value remains substantially lower than the downstream errors reported by Duc et al. (2025) [10] (RMSE ≈ 1.5–1.6), suggesting that the proposed model provides a reasonable level of predictive accuracy under comparable conditions.

A comparison between charts (a), (b), and (c) reveals that accuracy progressively declines as the forecast horizon increases. In chart (a), which represents the 1-day forecast horizon, the error is minimal because the simulation curve closely follows the observation curve. However, in charts (b) and (c), representing the 2-day and 3-day forecast horizons, respectively, there is a more pronounced divergence at the peaks and troughs. This trend aligns with time-series forecasting theory, in which uncertainty increases as the forecast horizon extends, leading to a reduced goodness-of-fit relative to short-term predictions.

Based on the experimental results, the 15-day lookback window yields reliable salinity forecasting performance at the Tra Vinh station. Although errors increase for the 3-day forecast horizon, the model retains an acceptable level of accuracy for short-term water resource management applications. These results suggest that the model can support decision-making processes, such as sluice-gate operation and water-intake management, as well as provide early warning of salinity intrusion to mitigate impacts on agricultural production.

4. Discussion

4.1. Analysis of the Underestimation of Observed Peaks at Hung My and Tra Vinh Station

A prominent common feature observed in the simulation results at both Hung My and Tra Vinh stations, especially in the medium-term forecast horizons (2- and 3-day), is the tendency for the models to produce lower peak values than the observed data. This trend becomes particularly apparent during sudden salinity surges, where the simulated output fails to match the maximum heights. Such underestimation of extreme values represents a frequent obstacle in machine learning and time-series forecasting when processing highly volatile hydrological information. The reason is that one factor is the smoothing effect generated by the optimisation process. Most forecasting models are developed to minimize the RMSE. This objective function naturally favours a cautious approach by gravitating toward mean values to prevent significant errors. Consequently, salinity at the two stations spikes under exceptional conditions, such as intense high tides or powerful winds, while the model often produces conservative results. Another factor concerns the limited availability of information and the impact of exogenous variables. Relying exclusively on a 15-day lookback period of salinity data might not adequately capture systemic fluctuations. Salinity levels at river mouths like Hung My and Tra Vinh are governed by a complex interplay between upstream freshwater flow and sea-driven tidal pressure. Without direct access to meteo-hydrological data, such as wind speed, rainfall, upstream discharge, or real-time tidal fluctuations, the model cannot effectively anticipate the rapid increase in peak amplitude during abrupt environmental shifts. This limitation is particularly relevant for early-warning applications, where underestimating extreme salinity events may reduce the system's sensitivity to high-risk conditions, potentially delaying timely management responses. The diminishing ability to forecast peak salinity values as the lead time extends from 1 to 3 days reflects increasing uncertainty and reduced predictive signal strength at longer forecast horizons. As the temporal distance between input sequences and target values increases, the dependence of historical observations on future outcomes weakens, leading to a tendency to underestimate peak magnitudes. Addressing this limitation may require model refinements such as incorporating additional exogenous variables (e.g., tidal dynamics, upstream discharge, and meteorological conditions) or adopting alternative training strategies that better capture extreme events.

4.2. Core Limitations and Challenges in Accurate Forecasting

Several studies acknowledge the critical role of environmental variables, such as upstream discharge, sea-level rise, and monsoon effects, in influencing salinity intrusion [2,36,37]. The forecasting results at Hung My station are consistent with the findings of Hoa Ho et al. (2024) [36] in the VMD, where the authors pointed out that the lack of salinity data makes it difficult for LSTM models to capture extreme values. To overcome these limitations, Hong et al. (2021) [37] suggests that integrating additional tidal and meteorological factors would enhance the model's responsiveness to sharp fluctuations at river mouths. The results at the Hung My and Tra Vinh stations indicate that the 15-day lookback window provides slightly better performance than the 30-day configuration. However, no formal statistical tests or confidence intervals were used to assess the significance of these differences. Future work should incorporate appropriate statistical testing to enable a more robust comparison of model configurations.

Salinity intrusion forecasting is a complex problem because it depends on multiple factors, including upstream freshwater discharge, high tides, sea level, and monsoons. This is consistent with recent studies highlighting increasing variability in flow and sediment dynamics in the VMD, which further complicates salinity forecasting and management [38]. Furthermore, research by Hong et al. (2021) [37] indicates that the Northeast monsoon has a significant impact, further complicating the task. Although ML and deep learning models show great potential, their accuracy remains heavily dependent on the quality and quantity of training data. Currently, the lack of key hydrometeorological variables (e.g., water level and wind data), combined with monitoring discontinuities (e.g., data collected only from January to May), poses a major challenge for integrating diverse data sources across spatial and temporal scales [4,23]. Many current studies rely on limited historical datasets, which may fail to capture emerging trends driven by climate change or anthropogenic alterations, such as dam construction and sediment starvation [1,39]. Critical data gaps, including insufficient spatial coverage and a lack of long-term datasets, have hindered model calibration and validation. Furthermore, model complexity and high data requirements can obstruct practical application by local agencies. The underestimation of peak salinity and difficulties in directly capturing short-term fluctuations diminish the operational reliability and utility of early warning systems [13,23]. This highlights the need for policy investment in integrated hydro-meteorological monitoring systems and data infrastructure to support reliable, real-time forecasting. This also aligns with broader research highlighting the ongoing transition toward more resilient and sustainable water management in the VMD under increasing climatic and anthropogenic pressures [40]. A further key limitation of this study is the absence of formal statistical testing and uncertainty quantification to assess differences between model configurations. Therefore, performance differences between lookback windows should be interpreted with caution.

4.3. Technical Evaluation and Future Research Directions

Research organisations generally agree that LSTM and other deep learning (DL) models are effective tools for short-term forecasting (ranging from hours to days) in estuarine environments, and have been shown to perform well in capturing non-linear dynamics. To enhance performance, techniques such as Gray Relational Analysis, Bayesian optimisation, and meta-heuristic algorithms (e.g., the Improved Sparrow Search Algorithm) have been employed to optimise hyperparameters [21,41]. Additionally, the use of remote sensing data, such as Landsat imagery and salinity indices, supports large-scale monitoring and helps address data sparsity in coastal regions of the Global South [42–45].

However, evidence shows that forecasting skill still declines as the lead time increases. LSTM models typically exhibit a significant drop in accuracy for forecasts exceeding 7 days,

which poses challenges for long-term seasonal or annual strategic planning [13,39,46]. Future research should focus on the comprehensive integration of hydro-meteorological, remote sensing, and even socio-economic variables to enhance model certainty. Expanding the monitoring station network and improving real-time data assimilation capabilities are essential prerequisites for supporting timely decision-making in dynamic estuarine environments [47].

5. Conclusions

This study demonstrates the LSTM model's capability to forecast salinity in the Cung Hau estuary of the VMD. The results indicate that a 15-day lookback window provides the best overall performance among the tested configurations, enabling the model to capture key temporal patterns and salinity intrusion dynamics. Performance metrics, including NSE values exceeding 0.83 for short-term forecasts (1–2 days), suggest that the model provides reliable predictions at both the Hung My and Tra Vinh stations.

However, forecasting performance declines as the lead time extends to 3 days, reflecting increasing uncertainty and the influence of unobserved exogenous drivers. In particular, the model tends to underestimate peak salinity during sudden salinity-intrusion events. This limitation is likely associated with the smoothing effect of the MSE loss function and the absence of key hydrometeorological inputs, such as tidal dynamics and upstream discharge.

In practical terms, the model can support water resource management by informing sluice-gate operations, water-intake planning, and early warning of salinity intrusion. From a policy perspective, these results demonstrate that short-term (1–2 day) salinity forecasts can provide a practical decision-support window for operational water management, particularly for sluice-gate regulation, irrigation scheduling, and freshwater abstraction planning at the provincial level. However, translating model outputs into operational practice requires institutional integration, including routine data sharing, clear agency responsibilities, and alignment with existing water management protocols. For operational uptake, forecast outputs should be linked to clear salinity thresholds for irrigation, domestic water supply, and sluice-gate operation, enabling model predictions to be translated into timely management actions. To further improve predictive performance, especially for extreme events and longer forecast horizons, future research should integrate additional data sources (e.g., wind speed, atmospheric pressure, upstream flow, and remote sensing data) and enhance real-time data assimilation capabilities. These developments will be important for advancing robust and scalable early warning systems in the Mekong Delta.

Author Contributions: Conceptualization, N.P.C., T.V.T. and H.V.T.M.; methodology, N.P.C., T.V.T., H.V.T.M. and P.K.; formal analysis, N.P.C., T.V.T. and H.V.T.M.; writing—original draft preparation, N.P.C., T.V.T., D.V.D., D.T.H.N., N.K.D., A.C., S.C., H.V.T.M. and P.K.; writing—review and editing, N.P.C., T.V.T., D.V.D., D.T.H.N., N.K.D., A.C., S.C., H.V.T.M. and P.K.; supervision, T.V.T. and H.V.T.M.; funding acquisition, P.K. All authors have read and agreed to the published version of the manuscript.

Funding: Authors gratefully acknowledge the financial support received from the Indian Council of Social Science Research (ICSSR), Government of India, and the Japan Society for the Promotion of Science (JSPS), under the ICSSR–JSPS Joint Research Programme [File No: ICSSR-JSPS (Japan)/JRP-02/2024-IC dated 2 May 2024] with project titled "Exploring the Interplay of Water and Energy Dynamics within Local Communities in Bathinda District, Punjab, India".

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A

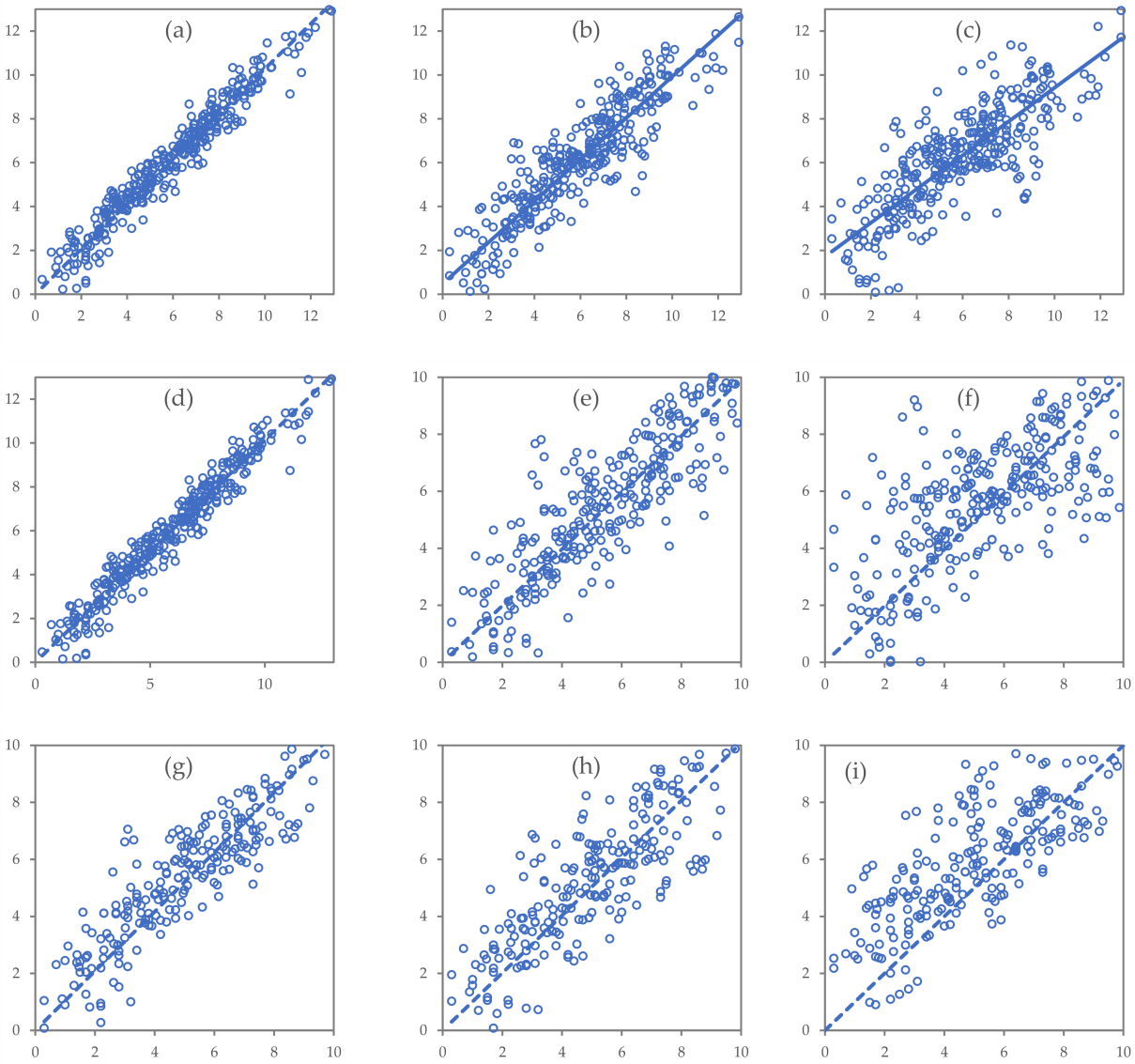


Figure A1. Comparison of observed and predicted mean salinity at Hung My station during the testing period. Subplots show forecast horizons of 1-, 2-, and 3-day across three lookback windows: 30-day (**a**–**c**), 45-day (**d**–**f**), and 60-day (**g**–**i**). Horizontal axis denotes the daily mean salinity observation while vertical axis denotes the daily mean salinity simulation.

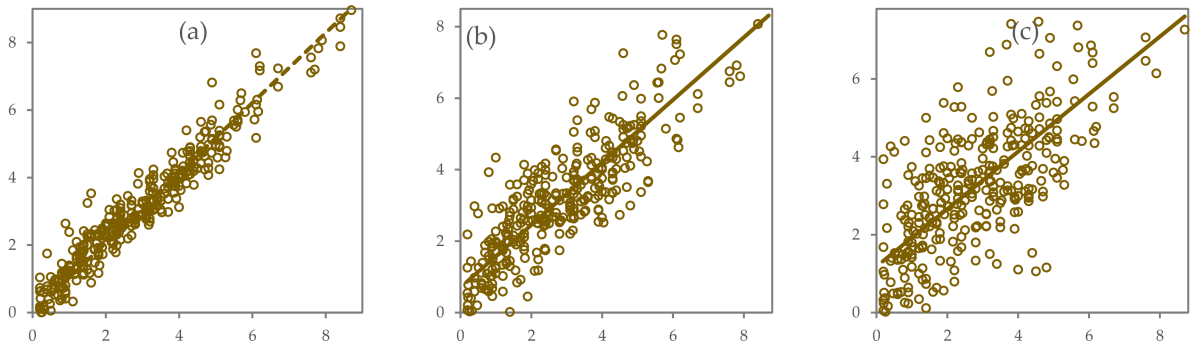


Figure A2. *Cont.*

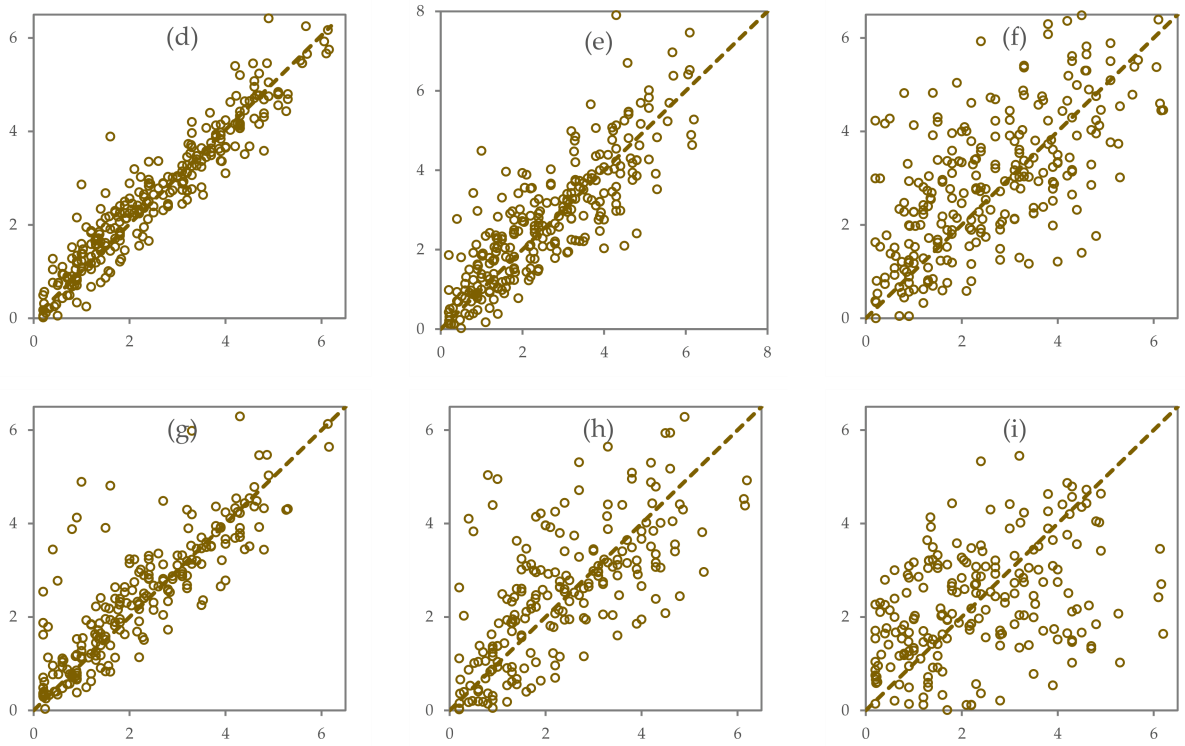


Figure A2. Comparison of observed and predicted mean salinity at Tra Vinh station during the testing period. Subplots show forecast horizons of 1-, 2-, and 3-day across three lookback windows: 30-day (**a**–**c**), 45-day (**d**–**f**), and 60-day (**g**–**i**). Horizontal axis denotes the daily mean salinity observation while vertical axis denotes the daily mean salinity simulation.

References

1. Eslami, S.; Hoekstra, P.; Minderhoud, P.S.; Trung, N.N.; Hoch, J.M.; Sutanudjaja, E.H.; Dung, D.D.; Tho, T.Q.; Voepel, H.E.; Woillez, M.-N. Projections of Salt Intrusion in a Mega-Delta under Climatic and Anthropogenic Stressors. *Commun. Earth Environ.* **2021**, *2*, 142. [CrossRef]
2. Thanh, T.N.; Huynh Van, H.; Vo Minh, H.; Tri, V.P.D. Salinity Intrusion Trends under the Impacts of Upstream Discharge and Sea Level Rise along the Co Chien River and Hau River in the Vietnamese Mekong Delta. *Climate* **2023**, *11*, 66. [CrossRef]
3. Hens, L.; Thinh, N.A.; Hanh, T.H.; Cuong, N.S.; Lan, T.D.; Van Thanh, N.; Le, D.T. Sea-Level Rise and Resilience in Vietnam and the Asia-Pacific: A Synthesis. *Vietnam J. Earth Sci.* **2018**, *40*, 126–152. [CrossRef]
4. Tran, D.D.; Pham, T.B.T.; Park, E.; Phan, T.T.H.; Duong, B.M.; Wang, J. Extent of Saltwater Intrusion and Freshwater Exploitability in the Coastal Vietnamese Mekong Delta Assessed by Gauging Records and Numerical Simulations. *J. Hydrol.* **2024**, *630*, 130655. [CrossRef]
5. Dang, L.T.T.; Ishidaira, H.; Nguyen, K.P.; Souma, K.; Magome, J. Short-Term Salinity Prediction for Coastal Areas of the Vietnamese Mekong Delta Using Various Machine Learning Algorithms: A Case Study in Soc Trang Province. *Appl. Water Sci.* **2025**, *15*, 79. [CrossRef]
6. Tran Anh, D.; Hoang, L.; Bui, M.; Rutschmann, P. Simulating Future Flows and Salinity Intrusion Using Combined One- and Two-Dimensional Hydrodynamic Modelling—The Case of Hau River, Vietnamese Mekong Delta. *Water* **2018**, *10*, 897. [CrossRef]
7. Eslami, S.; Hoekstra, P.; Kernkamp, H.W.J.; Nguyen Trung, N.; Do Duc, D.; Nguyen Nghia, H.; Tran Quang, T.; van Dam, A.; Darby, S.E.; Parsons, D.R.; et al. Dynamics of Salt Intrusion in the Mekong Delta: Results of Field Observations and Integrated Coastal–Inland Modelling. *Earth Surf. Dyn.* **2021**, *9*, 953–976. [CrossRef]
8. Nam, N.; Thuc, P.; Dao, D.; Thien, N.; Nguyen, H.A.; Tran, D. Assessing Climate-Driven Salinity Intrusion through Water Accounting: A Case Study in Ben Tre Province for More Sustainable Water Management Plans. *Sustainability* **2023**, *15*, 9110. [CrossRef]
9. Minh, H.V.T.; Ngoc, D.T.H.; Lien, B.T.B.; Diep, N.T.H.; Nguyen, P.C.; Thanh, N.T.; Lavane, K.; Downes, N.K.; Kumar, P. Freshwater–Salinity Regime Shifts in the Vietnamese Mekong Delta: Multi-Decadal Trends and Emerging Risks (2000–2024). *Environ. Geochem. Health* **2026**, *48*, 139. [CrossRef]
10. Duc, P.N.; Duc, T.T.; Van, G.P.; Van, H.N.; Minh, T.T. Predicting Salinity Levels in the Mekong Delta (Viet Nam): Analysis of Machine Learning and Deep Learning Models. *Discov. Artif. Intell.* **2025**, *5*, 79. [CrossRef]

11. Tran, D.Q.; Nguyen, N.N.; Huynh, M.V.; Bairagi, S.K.; Le, K.N.; Tran, T.V.; Durand-Morat, A. Modeling Saltwater Intrusion Risk in the Presence of Uncertainty. *Sci. Total Environ.* **2024**, *908*, 168140. [CrossRef]
12. Chong, Y.J.; Khan, A.E.; Scheelbeek, P.; Butler, A.P.; Bowers, D.; Vineis, P. Climate Change and Salinity in Drinking Water as a Global Problem: Using Remote-Sensing Methods to Monitor Surface Water Salinity. *J. Remote Sens.* **2014**, *35*, 1585–1599. [CrossRef]
13. Wullems, B.J.M.; Brauer, C.C.; Baart, F.; Weerts, A.H. Forecasting Estuarine Salt Intrusion in the Rhine–Meuse Delta Using an LSTM Model. *Hydrol. Earth Syst. Sci.* **2023**, *27*, 3823–3850. [CrossRef]
14. Qi, S.; He, M.; Bai, Z.; Ding, Z.; Sandhu, P.; Chung, F.; Namadi, P.; Zhou, Y.; Hoang, R.; Tom, B. Novel Salinity Modeling Using Deep Learning for the Sacramento–San Joaquin Delta of California. *Water* **2022**, *14*, 3628. [CrossRef]
15. Saccotelli, L.; Verri, G.; De Lorenzis, A.; Cherubini, C.; Caccioppoli, R.; Coppini, G.; Maglietta, R. Enhancing Estuary Salinity Prediction: A Machine Learning and Deep Learning Based Approach. *Appl. Comput. Geosci.* **2024**, *23*, 100173. [CrossRef]
16. Thai, T.T.; Liem, N.D.; Luu, P.T.; Yen, N.T.M.; Yen, T.; Quang, N.; Tan, L.; Hoai, P. Performance Evaluation of Auto-Regressive Integrated Moving Average Models for Forecasting Saltwater Intrusion into Mekong River Estuaries of Vietnam. *Vietnam J. Earth Sci.* **2021**, *44*, 18–32.
17. Hu, J.; Liu, B.; Peng, S. Forecasting Salinity Time Series Using RF and ELM Approaches Coupled with Decomposition Techniques. *Stoch. Environ. Res. Risk Assess.* **2019**, *33*, 1117–1135. [CrossRef]
18. Vu, M.-T.; Luu, C.; Bui, D.-Q.; Vu, Q.-H.; Pham, M.-Q. Simulation of Hydrodynamic Changes and Salinity Intrusion in the Lower Vietnamese Mekong Delta under Climate Change-Induced Sea Level Rise and Upstream River Discharge. *Reg. Stud. Mar. Sci.* **2024**, *78*, 103749. [CrossRef]
19. Gorski, G.; Cook, S.; Snyder, A.; Appling, A.P.; Thompson, T.; Warner, J.C.; Topp, S. Deep Learning of Estuary Salinity Dynamics Is Physically Accurate at a Fraction of Hydrodynamic Model Computational Cost. *Limnol. Oceanogr.* **2024**, *69*, 1070–1085. [CrossRef]
20. Wu, W.; Huo, L.; Yang, G.; Liu, X.; Li, H. Research into the Application of ResNet in Soil: A Review. *Agriculture* **2025**, *15*, 661. [CrossRef]
21. Zheng, R.; Sun, Z.; Jiao, J.; Ma, Q.; Zhao, L. Salinity Prediction Based on Improved LSTM Model in the Qiantang Estuary, China. *J. Mar. Sci. Eng.* **2024**, *12*, 1339. [CrossRef]
22. Nguyen, T.G.; Tran, T.D.; Nguyen, C.T. A Optimizing the Long Short-Term Memory (LSTM) Model by Bayesian Method for Salinity Intrusion Forecasting: A Study at Dai Ngai Station, Soc Trang Province, Vietnam. *Vietnam J. Mar. Sci. Technol.* **2023**, *23*, 223–232. [CrossRef]
23. Rummel, K.; Strauß, T.; Lauer, F.; Gräwe, U. Real-Time Prediction of Salt Intrusion in Tidal Estuaries Using Long Short-Term Memory Networks. *J. Geophys. Res. Mach. Learn. Comput.* **2025**, *2*, e2025JH000768. [CrossRef]
24. Liao, Y.-J.; Zhan, C.; Chang, Y.-S. Transformer-Based Model Sea Temperature and Salinity Prediction Using Satellite Remote Sensing and Argo Data Fusion. In Proceedings of the 2024 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Kuching, Malaysia, 6–10 October 2024; pp. 3541–3546. [CrossRef]
25. Zhang, Y.; Zhang, L.; Zhang, B.; He, Q. Results Analysis of Coastal Regions Sea Surface Salinity Retrieval from Aquarius Mission Using Deep Neural Network. In Proceedings of the IGARSS 2022—2022 IEEE International Geoscience and Remote Sensing Symposium, Kuala Lumpur, Malaysia, 17–22 July 2022; pp. 6990–6993. [CrossRef]
26. Liang, Z.; Bao, S.; Zhang, W.; Yan, H.; Duan, B.; Wang, H. Super-Resolution Reconstruction of SMOS Sea Surface Salinity from Multivariate Satellite Observations Based on Deep Learning. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2025**, *18*, 24251–24266. [CrossRef]
27. Ratheesh, S.; Jishad, M.; Dharaskar, R.V.; Sharma, N.; Agarwal, N.; Sharma, R. Implication of Errors in Space Based Sea Surface Salinity Measurement on Ocean State Estimation: An Observing System Simulation Experiment for the Tropical Indian Ocean. *IEEE Trans. Geosci. Remote Sens.* **2024**, *62*, 4206706. [CrossRef]
28. Lam, V. Predicting Salinity Intrusion for the Coastal Rice Cultivation Areas in the Mekong Delta. *Res. Crops* **2023**, *24*, 637–644. [CrossRef]
29. General Department of Meteorology and Hydrology Salinity Intrusion in the Mekong Delta. Available online: https://vnmha.mae.gov.vn (accessed on 15 February 2026).
30. Hochreiter, S.; Schmidhuber, J. Long Short-Term Memory. *Neural Comput.* **1997**, *9*, 1735–1780. [CrossRef]
31. Chen, G.; Long, T.; Xiong, J.; Bai, Y. Multiple Random Forests Modelling for Urban Water Consumption Forecasting. *Water Resour. Manag.* **2017**, *31*, 4715–4729. [CrossRef]
32. Liu, Y.; Wang, H.; Feng, W.; Huang, H. Short Term Real-Time Rolling Forecast of Urban River Water Levels Based on LSTM: A Case Study in Fuzhou City, China. *Int. J. Environ. Res. Public Health* **2021**, *18*, 9287. [CrossRef]
33. Ozan, K.; Öztürk, D. Derin Öğrenme Yöntemi Ile Kısa Süreli Rüzgar Hızı Tahmini: Bingöl Örneği (Short-Term Wind Speed Estimation with Deep Learning Method: Bingöl Example). In Proceedings of the 8th International Artificial Intelligence and Data Processing Symposium (IDAP'24), Malatya, Türkiye, 21–22 September 2024.
34. Legates, D.; Mccabe, G. Evaluating the Use Of "Goodness-of-Fit" Measures in Hydrologic and Hydroclimatic Model Validation. *Water Resour. Res.* **1999**, *35*, 233–241. [CrossRef]

35. Nash, J.E.; Sutcliffe, J.V. River Flow Forecasting through Conceptual Models Part I—A Discussion of Principles. *J. Hydrol.* **1970**, *10*, 282–290. [CrossRef]
36. Hoa Ho, V.; Xuan Quang Chau, N.; Hoang Giang Ngo, N.; Hai Pham, N.; Bay Nguyen, T. Exploring the Relationship of the Winter-Spring Crop's Rice Yield with Meteorological Drought and Water Resources in Ben Tre Province. *IOP Conf. Ser. Earth Environ. Sci.* **2024**, *1345*, 012010. [CrossRef]
37. Hong, N.V.; Hien, N.T.; Minh, N.T.T.; Toan, H.C. Forecasting Saline Intrusion under the Influence of the Northeast Monsoon in the Mekong Delta. *Vietnam J. Hydrometeorol.* **2021**, *9*, 23–36.
38. Thu Minh, H.V.; Kumar, P.; Nhat, G.M.; Diep, N.T.H.; Van Ty, T.; Van Duy, D.; Avtar, R.; Downes, N.K. Losing Ground: Flow Fragmentation and Sediment Decline in the Vietnamese Mekong Delta (2000–2023). *Earth Syst. Environ.* **2025**, 1–21. [CrossRef]
39. Cochrane, T.A.; Arias, M.E.; Piman, T. Historical Impact of Water Infrastructure on Water Levels of the Mekong River and the Tonle Sap System. *Hydrol. Earth Syst. Sci.* **2014**, *18*, 4529–4541. [CrossRef]
40. Tri, V.P.D.; Yarina, L.; Nguyen, H.Q.; Downes, N.K. Progress toward Resilient and Sustainable Water Management in the Vietnamese Mekong Delta. *Wiley Interdiscip. Rev. Water* **2023**, *10*, e1670. [CrossRef]
41. Roh, D.M.; He, M.; Bai, Z.; Sandhu, P.; Chung, F.; Ding, Z.; Qi, S.; Zhou, Y.; Hoang, R.; Namadi, P.; et al. Physics-Informed Neural Networks-Based Salinity Modeling in the Sacramento–San Joaquin Delta of California. *Water* **2023**, *15*, 2320. [CrossRef]
42. Phan, L.H.; Đặng, Đ.L.P.; Đỗ, T.N.; Nguyễn, T.T.H.; Nguyễn, T.D.M.; Hứa, H.H.; Pham, V.-M. Tích Hợp Viễn Thám và Giải Tích Số Định Lượng Xâm Nhập Mặn Tại Hạ Du Đồng Bằng Sông Cửu Long, Việt Nam. *Tạp Chí Khoa Học Và Công Nghệ Lâm Nghiệp* **2025**, *14*, 087–098. [CrossRef]
43. B, S.; Priyadarshini, R.; Bhattacharjee, S.; Kamath S, S.; U, P.; Gangadharan, K.V.; Ghosh, S. Spatio-Temporal Analysis and Modeling of Coastal Areas for Water Salinity Prediction. In Proceedings of the 2023 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS), Bhopal, India, 18–19 February 2023; p. 6.
44. Vinogradova, N.T.; Lee, T.; Boutin, J.; Drushka, K.; Fournier, S.; Sabia, R.; Stammer, D.; Bayler, E.; Reul, N.; Gordon, A.L.; et al. Satellite Salinity Observing System: Recent Discoveries and the Way Forward. *Front. Mar. Sci.* **2019**, *6*, 243. [CrossRef]
45. Xie, J.; Raj, R.P.; Bertino, L.; Martínez, J.; Gabarró, C.; Catany, R. Assimilation of Sea Surface Salinities from SMOS in an Arctic Coupled Ocean and Sea Ice Reanalysis. *Ocean Sci.* **2023**, *19*, 269–287. [CrossRef]
46. Lin, K.; Wei, S.; Li, T.; Lan, T.; Tu, X. Long-Term Daily Prediction of Saltwater Intrusion Based on the Long Memory-Double Autoregressive Model. *J. Hydroinf.* **2025**, *27*, 1431–1450. [CrossRef]
47. Lam, N.T. *Real-Time Prediction of Salinity in the Mekong River Delta*; Springer: Singapore, 2019; pp. 1461–1468.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.