JEE Journal of Ecological Engineering

Journal of Ecological Engineering, 2025, 26(11), 114–134
https://doi.org/10.12911/22998993/207704
ISSN 2299–8993, License CC-BY 4.0

Received: 2025.06.18
Accepted: 2025.08.01
Published: 2025.08.15

Estimating salinity using long short-term memory in the Vietnamese Mekong Delta and analyzing its dynamics

Le Van Chin[1], Chien Pham Van[1*]

[1] Faculty of Water Resources Engineering, Thuyloi University, 175 Tay Son, Dong Da, Hanoi, Vietnam
* Corresponding author's e-mail: Pchientvct_tv@tlu.edu.vn

ABSTRACT

Accurate estimation of salinity is critical for water resource management in deltas and coasts. Traditional methods such as numerical models often rely on physical transport processes, leading to significant uncertainty in modeling parameter estimation. Data-driven models like long short-term memory (LSTM) offer an effective alternative. This study implements the LSTM model to estimate salinity at multiple locations in the Vietnamese Mekong Delta. Hourly tidal data from Vung Tau and discharge data from Chau Doc and Tan Chau were applied as inputs, with salinity data from six locations as outputs. The model is trained and tested using data collected from 01/01/2014 to 30/06/2017, before being applied. The model's accuracy was evaluated using several statisitcal indicators, including Nash-Sutcliffe efficiency (NSE), Pearson's correlation coefficient (r), mean error (ME), mean absolute error (MAE), and using root-mean-square error (RMSE). The findings indicated that the LSTM model accurately reproduced salinity, with dimensionless errors between 0.84 and 0.99, and dimensional errors from -0.31 to 0.38 PSU. These results demonstrate the reliability and generalizability of LSTM models for salinity estimation in the Vietnamese Mekong Delta. Moreover, the integration of wavelet and wavelet coherence analyses provided novel insights into the temporal structure and multiscale interactions of salinity with key hydrodynamic drivers, such as river discharge and tidal forcing. This study contributes to the growing body of literature advocating for hybrid and machine learning-based modeling approaches in hydro-environmental science, offering scalable, interpretable, and efficient tools for forecasting and decision support in data-scarce coastal regions worldwide.

Keywords: salinity, Vietnamese Mekong Delta, long short-term memory, wavelet analysis, wavelet coherence, wavelet power spectrum.

INTRODUCTION

Salinity estimation in deltaic and coastal regions, characterized by the mixing of freshwater from rivers and marine water from the sea, has received significant attention due to its biodegradable propertiess in shaping environmental dynamics (Tran et al., 2023), water supply and agricultural activities (Nguyen et al., 2019), water resources management (Nguyen et al., 2023a), and ecosystem development (Pereira et al., 2019). Salinity levels shape the physical properties of water, influencing processes like sediment transport and erosion (Anthony et al., 2015). Indeed, salinity gradients, where freshwater from rivers meets saltwater from the sea, create diverse habitats that support a wide range of species, including specialized plants like mangroves and salt marsh grasses, and animals adapted to varying salinity levels (Vu et al., 2024). These gradients also impact primary production by affecting nutrient availability, driving high productivity in estuaries and deltas that supports complex food webs. In terms of water supply and humain activities in deltaic regions, agriculture is heavily influenced by salinity levels in both soil and water (Nguyen et al., 2019). High salinity can reduce crop yields and soil fertility, posing challenges for sustainable farming and requiring to manage salinity for maintaining agricultural productivity (Nguyen et al., 2021). Futhermore, controlling salinity is vital for ensuring the availability of fresh water for human consumption and irrigation (Nguyen et al., 2019). Deltas and coasts are also particularly vulnerable to climate

change, where rising sea levels increase the risk of salt intrusion, altering salinity patterns and further stressing ecosystems and water resources (Smajgl et al., 2015; Pereira et al., 2019). Thus, accurately estimating salinity in deltaic and coastal regions still remains challenging, not only for determining precise values but also for supporting health and functionality of these environments.

Different methods have been applied to measure and calculate salinity in coastal and deltaic regions such as field campaign measurement, numerical models, remote sensing technologies, data-driven models, etc. Field campaign measurement normally determine salinity using electrical conductivity sensors in natural settings, providing data for models' calibration and validation as well as related analysis (Vu et al., 2024). Numerical models such as one-dimensional models (Tran et al., 2023), combination one-and two-dimensional models (Tran et al., 2018), two-dimensional models (Vu et al., 2024), three-dimensional models (Tran et al., 2024) are another powerful tool, employing mathematical equations to simulate salinity variations and intrusion in estuaries and coastal regions, based on input data like river discharge, tide, and meteorological conditions. In addition to field campaign measurement and numerical models, remote sensing technologies using satellite imagery are employed to monitor salinity changes in coastal and deltaic areas, providing insights into broader spatial patterns (e.g., Nguyen et al., 2021; Tran et al., 2023). However, it must be noted that these approaches (i.e., filed campaign measurement, numerical models, and remote sensing techniques) have some limitations even though they provide valuable information for understanding and managing salinity's impact on ecosystems and human activities. For instance, filed campaign measurement cannot provide salinity with a high spatio-temporal resolution over a large region since it is time-consuming and labor-intensive (Tran et al., 2023), while numerical models require a lot of data including bathymetry, hydrological data, infrastructure, and constructions, etc which are often lacking in coastal and deltaic regions. Applications of remote sensing technologies are also constrained by high costs and the need for substantial memory. All of this has restricted the use of these approaches in many coastal and deltaic regions worldwide. Therefore, there remains a critical need to refine current methodologies or to develop more robust and automated alternatives capable of overcoming these constraints

Among various data-driven models, the long short-term memory (LSTM) have recently increased attention from scientific communities worldwide for its highly accurate salinity prediction capabilities (Nguyen et al., 2021; Tran et al., 2022; Saccotelli et al., 2024). Note that LSTM has been widely applied to tackle the nonlinear and stochastic transport processes of salinity, with less emphasis on the physical properties of transport processes. This study emphasizes the use of data-driven, black-box approaches to establish optimal mathematical relationships between input and output variables. The LSTM model is particularly advantageous due to its ability to handle large and multi-scale datasets effectively, while maintaining robustness against missing data (Saccotelli et al., 2024). Nguyen et al. (2021) used the LSTM model combined with remote sensing to predict salt intrusion in the Vietnamese Mekong Delta (VMD). Tran et al. (2022) compared the performances of LSTM with four machine learning algorithms (Simple Linear, K-Nearest Neighbors, Random Forest, Support Vector Machine) for predicting salt intrusion in the Ham Luong river in the VMD using limited salinity monitoring data, revealing that LSTM was the most accurate and efficient model when input data was limited. Saccotelli et al. (2024) employed the LSTM model to improve the accuracy of salinity predictions in the Po River estuary (Italy), aiming to better understand the complex interplay of factors influencing estuarine salinity dynamics. These examples demonstrate that the LSTM model can be effectively applied to estimate salinity in the region examined in this study.

The study is designed to address the following objectives: (i) implement the LSTM model to estimate hourly salinity at six locations in the VMD based on the available discharge at Chau Doc and Tan Chau as well as tidal data at Vung Tau, (ii) identify suitable hyper-parameters for this estimation, and (iii) examine salinity dynamics, periodicity, and the relationship between salinity and discharge/tide using wavelet analysis. Besides these objectives, it is worth noting that this study is the first to evaluate the performance of the LSTM model in estimating salinity at various locations within the VMD. To the best of our knowledge, this is the first study apply LSTM for multi-site salinity estimation in the VMD using available hydrological data from specific

upstream and coastal boundary conditions. The research uniquely integrates statistical and wavelet-based approaches to evaluate both the predictive skill and temporal structure of the modeled salinity signals. Hydrological data, including discharge at Tan Chau and Chau Doc, tidal data at Vung Tau, and salinity at six given locations within the VMD were collected in the period from 01/01/2014 to 31/12/2018. To quantitatively compare the estimated results with the observed salinity data, five statistical performance indicators are applied, including Nash–Sutcliffe efficiency (NSE), Pearson's correlation coefficient (r), mean error (ME), mean absolute error (MAE), and using root-mean-square error (RMSE).

The novel contribution of the present study to existing researches lies not only in its methodological implementation but also in its potential to provide a transferable, scalable framework for salinity monitoring in other vulnerable coastal and deltaic regions. The findings from the study are also expected to inform water resource managers, agricultural planners, and climate adaptation strategies, offering timely support in delatic and coastal regions where traditional modeling approaches are limited by data scarcity or complexity.

THE VIETNAMESE MEKONG DELTA AND DATA COLLECTION

The Vietnamese Mekong Delta

The Mekong River has known as one of the ten most crucial deltas globally due to its hydrodynamic properties, transport of tracers like salinity and sediment, and its significant vulnerability to rising sea levels and climate change (Anthony et al., 2015; Smajgl et al., 2015; Tran et al., 2022; Tran et al., 2023; Vu et al., 2024; Tran et al., 2024). The river originates from the Tibetan Plateau in China and traverses through six countries (i.e., China, Laos, Cambodia, Thailand, Myanmar, and Vietnam), spanning a total length of 4,350 km (see Figure 1). It discharges into the East Sea through the Vietnamese Mekong Delta (VMD), which spans an area of about 39,400 km² (Tran et al., 2023). About 19,000 km² of this delta is dedicated to rice cultivation, contributing around 50% of Vietnam's total rice production. The delta, home to nearly 20 million people, is vital for socio-economic growth, particularly in agriculture and food security (Nguyen et al., 2019).

The VMD features a predominantly flat topography, with an average elevation ranging from 0 to 2 meters above sea level. It also indicates a complex network of distributaries and rivers, including Tien and Hau Rivers, along with a dense system of channels containing thousands of culverts and structures used to support water distribution for various activities (Tran et al., 2024). The region experiences a tropical climate, with a distinct rainy season from May to October and a dry season from November to April. Annual rainfall in the delta varies between 1.400 and 2.200 mm, with 90–95% of it occurring during the rainy season.

In terms of tide and flow, the tidal patterns in the delta are complex, with semi-diurnal tides in the East Sea and diurnal tides in the West Sea. In the semi-diurnal regime, high tide lasts around 6 hours, while low tide lasts approximately 7 hours, with tidal heights varying between 3 and 4 meters, occasionally reaching a maximum of 4.1 meters. The diurnal regime features two high and two low tides daily, with tidal ranges of 0.8 to 1.2 meters (Tran et al., 2023; Vu et al., 2024). Tides in the delta are expected to be particularly vulnerable to the effects of climate change, worsening both drought conditions and salt intrusion in dry season. On the other hand, the delta receives an annual flow volume of approximately 500 km³, with 23 km³ (4.6%) directly from precipitation and the remaining 477 km³ from the upstream Mekong River. Hourly flow measurements at Tan Chau station, collected from 2014 to 2017, ranged from -5.000 to 26.000 m³/s. Negative values indicate flow from the East Sea back toward the upstream Mekong, while positive values indicate flow in the opposite direction. Most of the annual flow (around 70–80%) occurs during the rainy season, while only 20–30% takes place in the dry season (Nguyen et al., 2023b).

In the VMD, salinity is a critical factor for water quality, influencing various aspects and purposes of water use, such as drinking water supply and irrigation for rice cultivation (Nguyen et al., 2019; Nguyen et al., 2021). The temporal and spatial distribution of salinity, as well as salt intrusion and salinity dynamics in the region, are heavily influenced by both tidal patterns and the flow of the upstream Mekong River. Additionally, the VMD is significantly impacted by climate change and rising sea levels. A number of efforts have been made to investigate salinity and salt intrusion in the VMD as consequences of global warming (Smajgl et al., 2015) and intensive

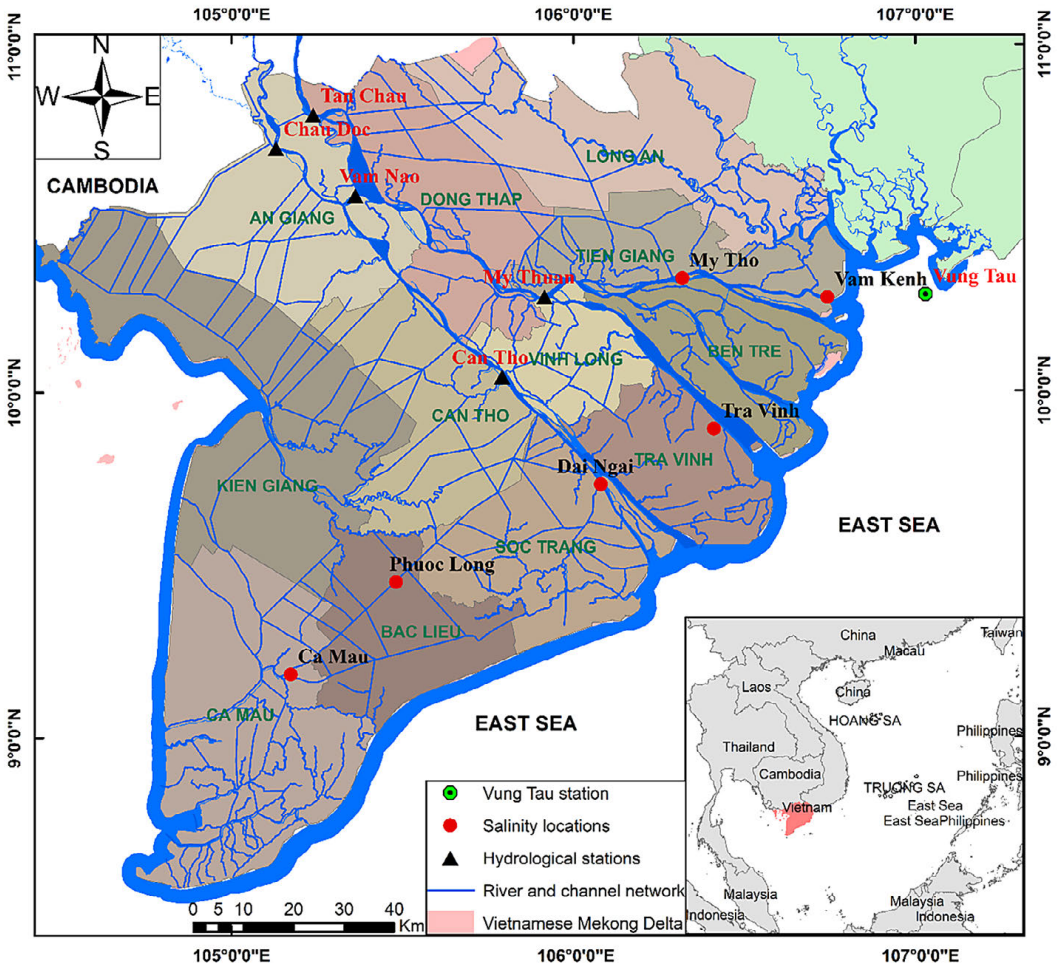

Figure 1. Map of the Vietnamese Mekong Delta, together with the hydrological stations and estimated salinity locations

human activities (Anthony et al., 2015). However, accurate prediction of salinity in the VMD is still critical not only for understanding salinity dynamics but also for developing suitable adaptation measures in agricultural and aquaculture production as well as freshwater supplies within the VMD while safeguarding its environment.

Data collection

Regarding data collection, hourly tidal data at Vung Tau station (downstream of the delta, see Figure 1) and water discharge at Tan Chau and Chau Doc stations (upstream of the delta, see Figure 1) were collected in the period from 01/01/2014 to 31/12/2018. Salinity data at six monitoring locations (i.e., Ca Mau, Phuoc Long, Dai Ngai, Tra Vinh, My Tho, and Vam Kenh, see Figure 1) were collected from 01/01/2014 to 30/06/2017. Flow, tidal, and salinity data were collected during the aforementioned periods because they were the only available datasets. Figure 2 shows the time-series of these data, while various statistical characteristics (e.g., range, mean, standard deviation, skewness, and kurtosis) of tide, discharge as well as salinity in the considered period are summarized in Table 1. Discharge (at Tan Chau and Chau Doc) and tide (at Vung Tau) versus salinity at six collected locations are shown in Figure 3 to Figure 5. Within each figure panel, the solid line represents the linear trend of salinity variation, which is a first-order polynomial. The correlation coefficients between discharge, tide and salinity are summarized in Table 2. Most salinity locations show a negative correlation, with values ranging from -0.455 to -0.023, except for the Vam Kenh location.

Salinity at collected locations was mostly measured from January to June of each year on an hourly timescale. However, it must be noted that these time-series data of salinity are still discontinuous throughout the entire period and across different years due to various factors, such as limitations

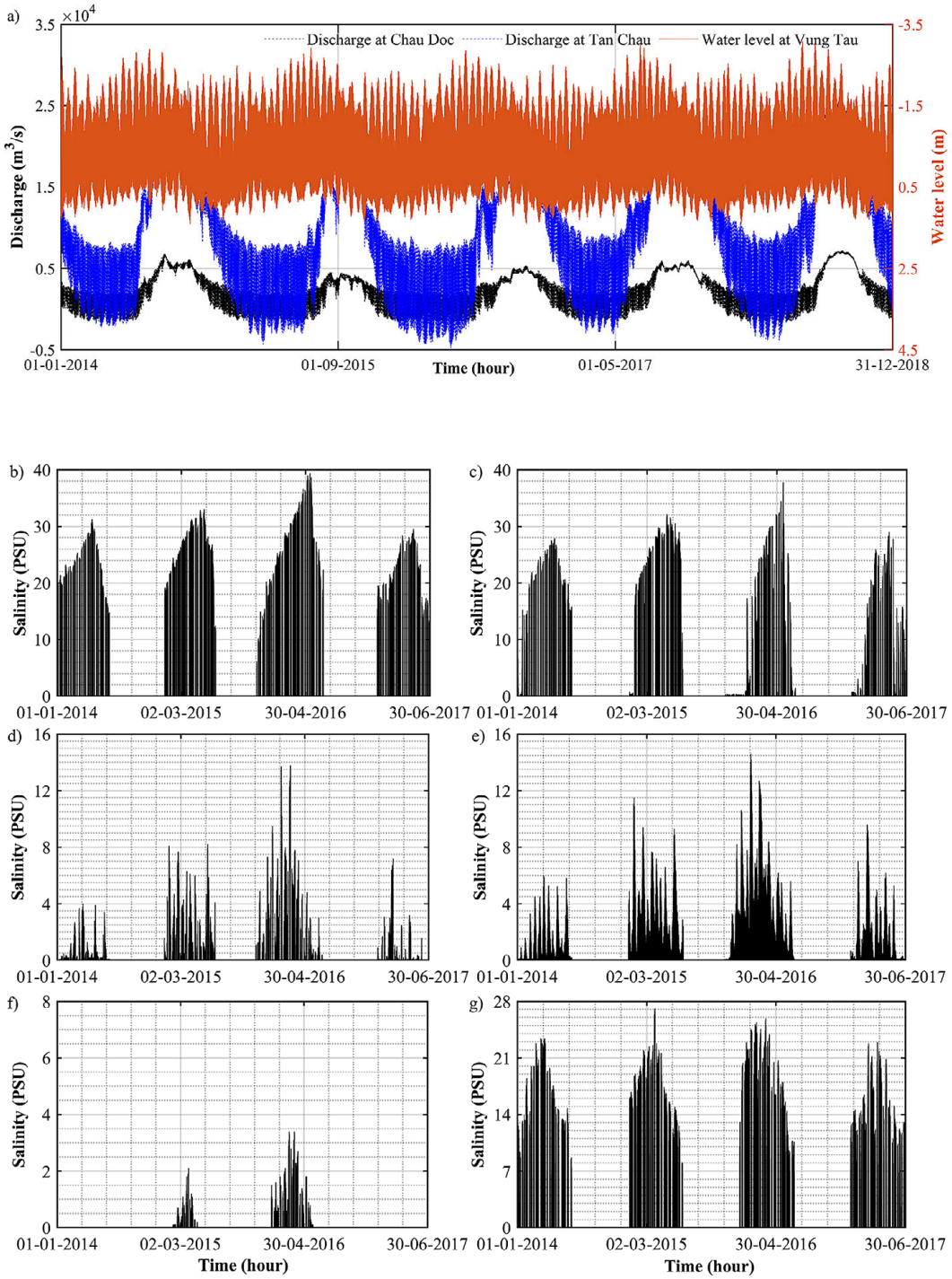

Figure 2. Hourly time series of: a) discharge and water level at collected stations and salinity at b) Ca Mau, c) Phuoc Long, d) Dai Ngai, e) Tra Vinh, f) My Tho, and g) Vam Kenh locations

in sampling equipments and the extreme large studied region. In addition, due to the fragmented information-sharing policies and the absence of real-time salinity monitoring systems, users in different provinces within the VMD are unable to monitor and collect salinity data in real-time. This has negatively impacted the region's socio-economic development, particularly in agriculture and water supply. In this study, tidal data from Vung Tau, discharge data from the Tan Chau and Chau Doc stations, and salinity data from six locations between 01/01/2014 and 30/06/2017 were used to train and test the LSTM model.

In terms of data processing, missing values of salinity in the collected period from 2014–2017 were interpolated linearly to create hourly time

Table 1. Discharge, water level and salinity collection at monitoring stations throughout the VMD

Name	Locations		Statistical characteristics					Quantity	Data collected period
	Longitude	Latitude	Range*	Mean	Standard deviation	Skewness	Kurtosis		
Chau Doc	105°07'00"	10°42'00"	-1850 ÷ 7210	2244.0	2040.6	0.2	2.5	Discharge (m^3/s)	2014–2018
Tan Chau	105°51'00"	10°50'00"	-4780 ÷ 26200	10225.6	6887.7	0.1	2.2		
Vung Tau	107°06'00"	10°18'00"	-3.08 ÷ 1.41	-0.2	0.8	-0.6	2.7	Water level (m)	
Ca Mau	105°09'00"	9°10'00"	4.9 ÷ 39.4	24.05	5.83	-0.12	2.85	Salinity (PSU)	2014–2017
Phuoc Long	105°27'40"	9°26'00"	0.1 ÷ 37.8	16.10	10.46	-0.35	1.70		
Dai Ngai	106°02'40"	9°47'30"	0 ÷ 13.8	1.44	2.01	2.35	9.83		
Tra Vinh	106°21'00"	9°58'40"	0.1 ÷ 14.6	2.05	2.32	1.78	6.64		
My Tho	106°22'00"	10°21'00"	0 ÷ 3.4	0.72	0.70	1.30	4.08		
Vam Kenh	106°45'00"	10°16'00"	0.1 ÷ 27.1	11.27	5.30	0.37	2.43		

Note: * A negative discharge value indicates flow direction from the East Sea to river.

Table 2. Correlation coefficients between salinity at different locations and discharge at Tan Chau and Chau Doc or water level at Vung Tau

Salinity at	Total number of hourly salinity data	Discharge at		Water level at	Note
		Chau Doc	Tan Chau	Vung Tau	
Ca Mau	4740	-0.288	-0.455	-0.066	Negative correlation
Phuoc Long	4692	-0.204	-0.323	-0.126	
Dai Ngai	4884	-0.221	-0.297	-0.027	
Tra Vinh	9168	-0.296	-0.387	-0.023	
My Tho	1523	-0.235	-0.268	-0.105	
Vam Kenh	4440	0.106	-0.001	0.509	Negative and positive correlations

series, in order to render the data processing as simple as possible. The hydrological and salinity data exhibit strong non-linearity, leading to numerical challenges during the training process of the LSTM model and significantly affecting estimation accuracy (Niu et al., 2020). Therefore, normalizing the data is essential to mitigate these issues. In this study, the hydrological and salinity datasets (water level, discharge, and salinity) at each station were normalized to a range between zero and unity.

METHOD

Long short–term memory model

The LSTM model, a prominent variant of recurrent neural networks (RNN) designed to retain information over extended time periods, has been employed for predicting salinity levels (Tran et al., 2022; Saccotelli et al., 2024). The model was first presented by Hochreiter and Schmidhuber (1997) to address a common issue such as the vanishing gradient problem in training traditional RNN, where gradients shrink exponentially during backpropagation. The fundamental element of the LSTM is the memory cell, which plays a role analogous to that of a hidden layer in conventional neural networks. Each memory cell is equipped with a recurrent connection characterized by a specific set of weights, purposefully designed to mitigate issues related to vanishing and exploding gradients. This recurrent pathway maintains a value known as the cell state, which serves as a conduit for information flow across time steps. The cell state from the previous time step (denoted as C_{t-1}) is transmitted to the current time step, allowing the network to retain long-term dependencies. Importantly, this state is updated to form the current cell state (denoted as C_t) through controlled mechanisms, rather than being directly scaled by weight multipliers. An LSTM cell contains three specialized gating mechanisms (namely the input gate, forget gate, and output gate) which collectively control the flow and retention of information within the network.

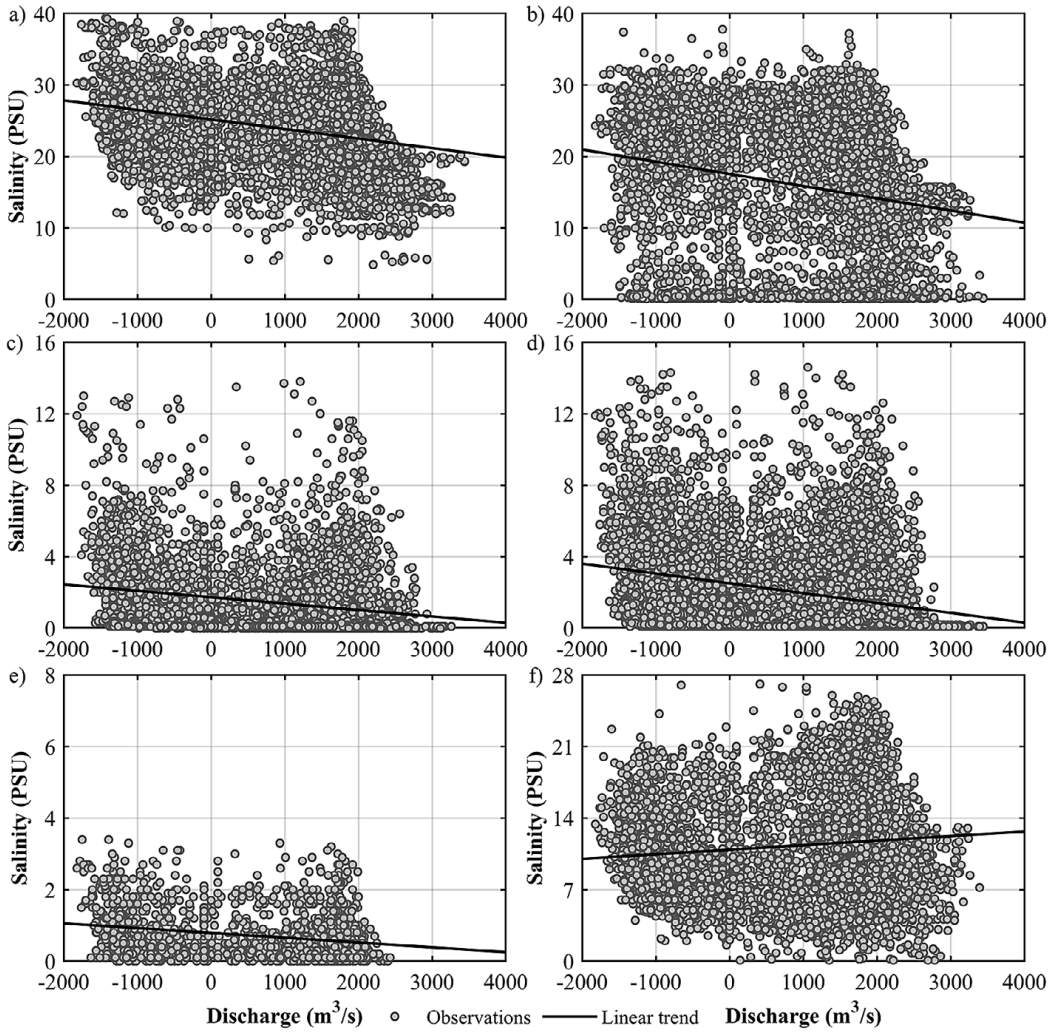

Figure 3. Discharge at Chau Doc versus salinity, at: a) Ca Mau, b) Phuoc Long, c) Dai Ngai, d) Tra Vinh, e) My Tho, and f) Vam Kenh locations

The input gate (i_t) and the input node (g_t) play a key role in updating the cell state. Their values are calculated using the following expressions (Pham Van and Nguyen-Ngoc, 2022).

$$i_t = \sigma\left(W_{xi}x_t + W_{hi}h_{t-1} + b_i\right) \quad (1)$$

$$g_t = \tanh\left(W_{xg}x_t + W_{hg}h_{t-1} + b_g\right) \quad (2)$$

where: W_{xi}, W_{hi}, W_{xg}, W_{hg} represent the weight matrices within the LSTM architechture, where x_t denotes the input at current time step t, and h_{t-1} corresponds to the hidden state from the previous time step t-1. The function σ(×) refers the sigmoid activation function, , while tanh refers to the hyperbolic tangent function. Additionally, b_i and b_g are bias terms associated with the respective gates.

The update of the cell state at time t is governed by the following equation:

$$C_t = (C_{t-1} \odot f_t) \oplus (i_t \odot g_t) \quad (3)$$

where: $\oplus$ denotes element-wise addition, $\odot$ is represents element-wise multiplication, and f_t is the forget gate.

The forget gate (f_t) plays a crucial role in controlling the retention and removal of information in the memory cell, thereby preventing the unboundaed accumulation of the cell state. The computation of the forget gate is expressed as follows:

$$f_t = \sigma\left(W_{xf}x_t + W_{hf}h_{t-1} + b_f\right) \quad (4)$$

The output gate, denoted by o_t, governs the updating process of the hidden state. It integrates the current input x_t, the hidden state from the

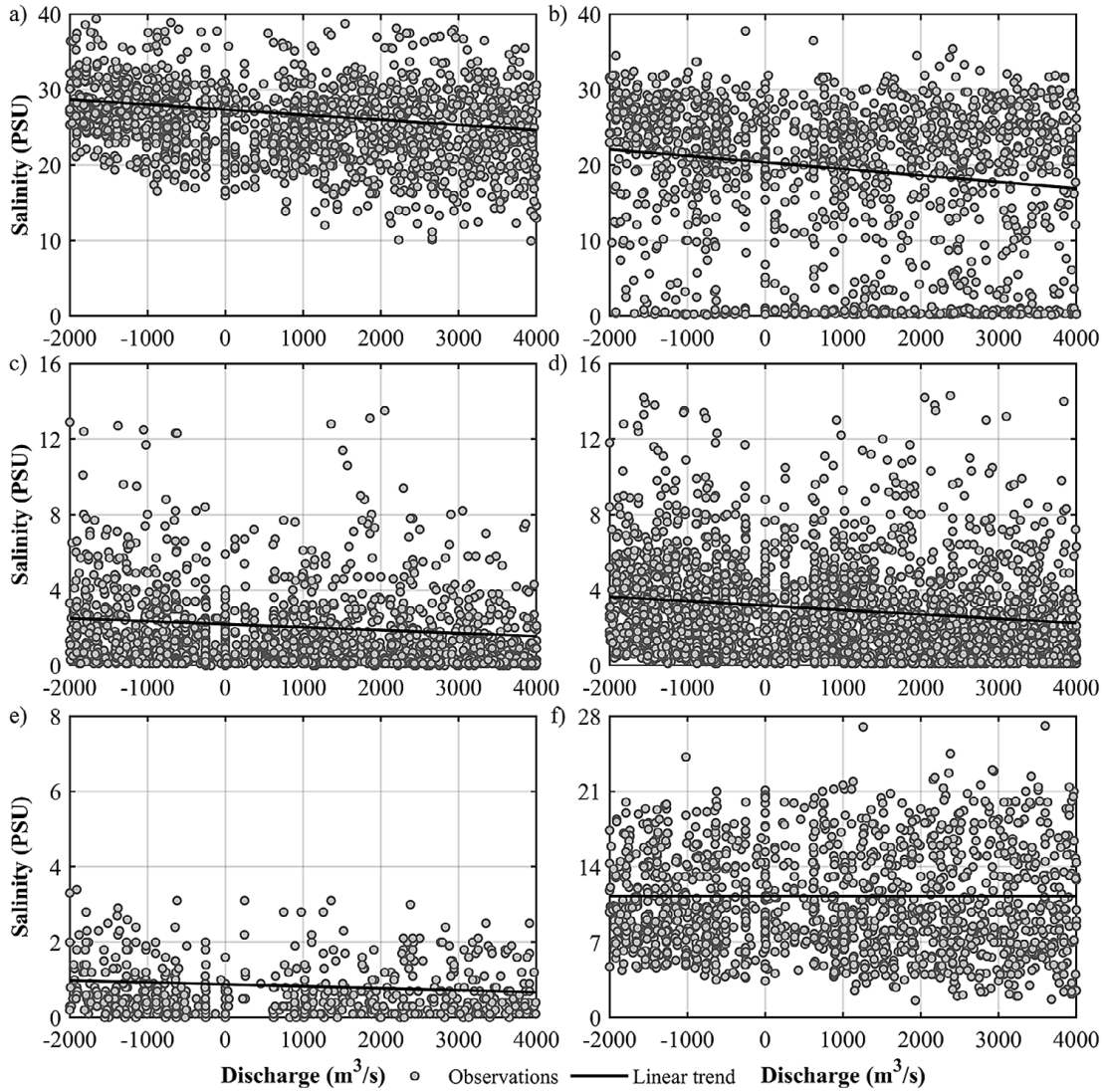

Figure 4. Discharge at Tan Chau versus salinity, at: a) Ca Mau, b) Phuoc Long, c) Dai Ngai, d) Tra Vinh, e) My Tho, and f) Vam Kenh locations

previous time step h_{t-1}, and the corresponding bias term b_o to compute the output as follows:

$$o_t = \sigma\left(W_{xo}x_t + W_{ho}h_{t-1} + b_o\right) \quad (5)$$

It is noteworthy that the hidden state at the current time step is determined using the following expression:

$$h_t = o_t \odot tanh(C_t) \quad (6)$$

The number of epochs, the number of hidden units, and the learning rate are three key hyperparameters that are optimized through a trial and error approach during the training phase. For model development at each location, 70% of the salinity dataset is used for training, while the remaining 30% is reserved for testing.

Performance metrics

Performance evaluation plays a critical role in the development and validation of LSTM model. In this study, five commonly used statistical indicators, i.e., Nash-Sutcliffe efficiency (NSE), Pearson's correlation coefficient (r), mean error (ME), mean absolute error (MAE), and using root-mean-square error (RMSE) are employed to quantitatively assess the correspondence between simulated and observed salinity values across various locations in the VMD.. The mathematic expression of these statistical indicators are reported in the following:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(S_{obs,i} - S_{sim,i}\right)^2} \quad (7)$$

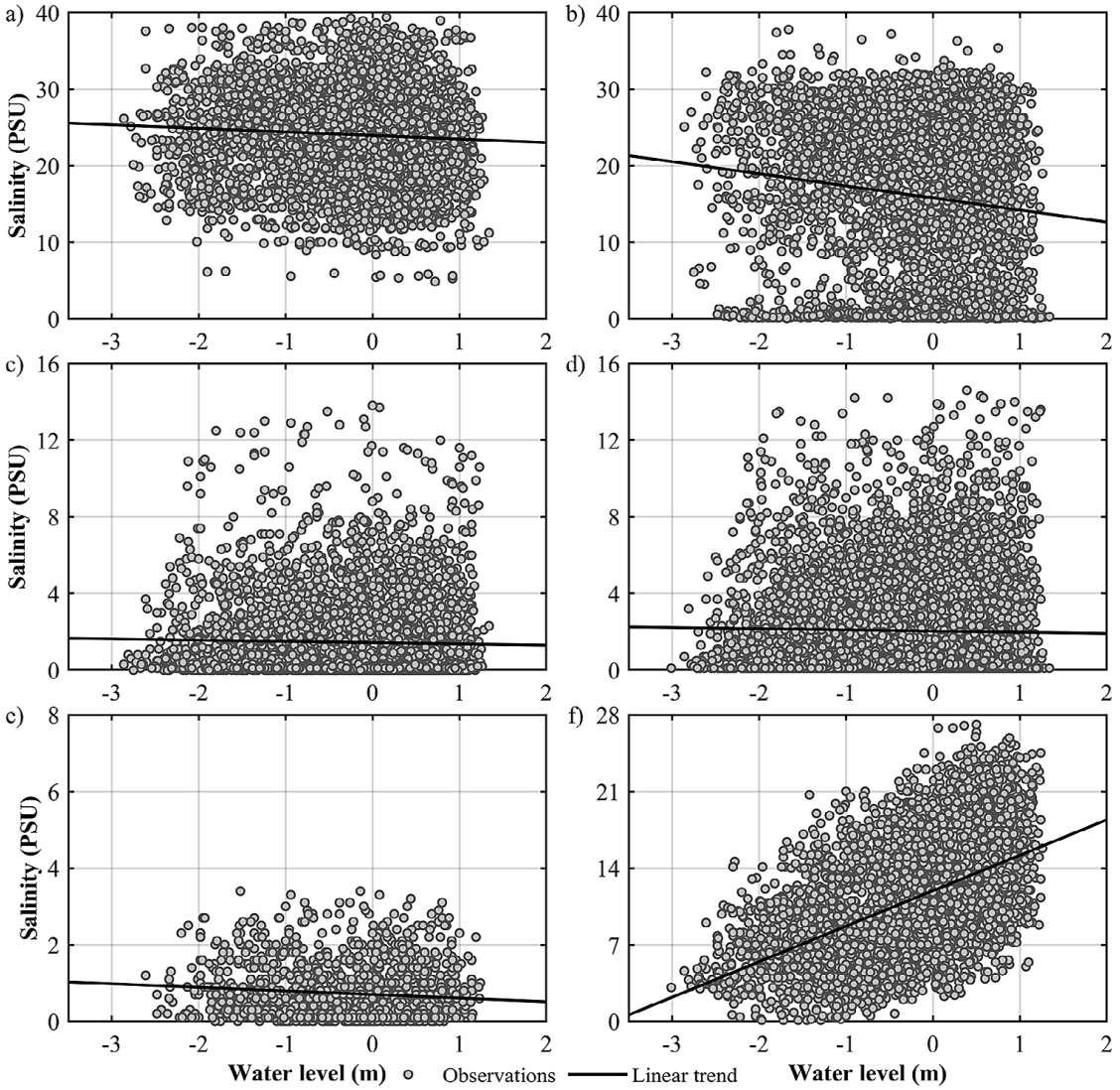

Figure 5. Water level at Vung Tau versus salinity, at: a) Ca Mau, b) Phuoc Long, c) Dai Ngai, d) Tra Vinh, e) My Tho, and f) Vam Kenh locations

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|S_{obs,i} - S_{sim,i}\right| \quad (8)$$

$$ME = \frac{1}{N}\sum_{i=1}^{N}\left(S_{sim,i} - S_{obs,i}\right) \quad (9)$$

$$r = \frac{\sum_{i=1}^{N}\left(S_{obs,i} - S_{obs,m}\right)\left(S_{sim,i} - S_{sim,m}\right)}{\sqrt{\sum_{i=1}^{N}\left(S_{obs,i} - S_{obs,m}\right)^2}\sqrt{\sum_{i=1}^{N}\left(S_{sim,i} - S_{sim,m}\right)^2}} \quad (10)$$

$$NSE = 1 - \frac{\sum_{i=1}^{N}\left(S_{sim,i} - S_{obs,i}\right)^2}{\sum_{i=1}^{N}\left(S_{obs,i} - S_{obs,m}\right)^2} \quad (11)$$

where: S_{obs} and S_{sim} present observed and estimated salinity, respectively; i and N indicate the time step and total length of the data, respectively; and $S_{obs,m}$ and $S_{sim,m}$ stand for mean observed and simulated salinity, respectively.

Notably, RMSE, MAE, and ME report errors in the same units as the target variable, thereby enabling direct and intuitive interpretation of model accuracy (Pham Van et al., 2023). In contrast, the Pearson correlation coefficient (r) and Nash-Sutcliffe efficiency (NSE) assess the proportion of variance in the residuals relative to that in the observations, offering a comprehensive measure of agreement between simulated and measured values. By combining these

complementary metrics, a more nuanced evaluation of model performance is achieved, guiding further refinement of the predictive framework.

Wavelet analysis

Wavelet analysis is a versatile technique for examining signals and has found extensive use in fields such as environmental data interpretation (Torrence and Compo, 1998; Addison, 2017; Grinsted et al., 2004). It is commonly employed to explore the wavelet power spectrum of time series at specific sites or locations. By decomposing time series data, this approach reveals key patterns of variability and tracks their evolution across multiple time scales (Addison, 2017). A central component of this method is the continuous wavelet transform, which involves projecting the data series (e.g., river discharge) into a wavelet domain through convolution with a scaled and shifted Morlet wavelet.

Let $x(\tau)$ represent the time series of the variable. Its continuous wavelet transform is expressed in Equation 12:

$$w_x(\nu,t)=\frac{1}{\sqrt{\nu}}\int_{-\infty}^{+\infty}x(\tau)\psi\left(\frac{\tau-t}{\nu}\right)d\tau \qquad (12)$$

where: ν and $\psi(t)$ are the scale parameter and complex conjugate of the Morlet wavelet, respectively. The $\psi(t)$ is defined as:

$$\psi(t)=\pi^{-1/4}\left(e^{i2\pi f_0 t}-e^{-(2\pi f_0)^2/2}\right)e^{-t^2/2} \qquad (13)$$

In this study, the central frequency f_o of the Morlet wavelet is set to 0.85, a value commonly adopted in practical applications (Addison, 2017).

To evaluate the relationship between rainfall and water discharge over time and frequency, wavelet coherence analysis is applied. Given two time series $y(t)$ and $x(t)$, with their respective wavelet transforms $w_y(\nu,t)$ and $w_x(\nu,t)$, the wavelet coherence between them is calculated as shown in Equation 14:

$$WCH^2_{x,y}(\nu,t)=\frac{\left|\left\langle w_x(\nu,t)\,w_y(\nu,t)\right\rangle\right|}{\left\langle\left|w_x(\nu,t)\right|^2\right\rangle\left\langle\left|w_y(\nu,t)\right|^2\right\rangle} \qquad (14)$$

where: $\langle\cdot\rangle$ denotes a localized smoothing function applied in both time and scale domains to the wavelet transform components.

Wavelet analysis involves two components: the wavelet power spectrum and wavelet coherence. The wavelet power spectrum identifies areas of high energy in the time-frequency domain, thereby exposing the primary modes of variability in a time series. In contrast, wavelet coherence measures how strongly two time series co-vary across both time and frequency, offering insights into their dynamic relationships (Torrence and Compo, 1998). To assess the statistical significance of these findings, the wavelet power spectrum is typically compared to a red noise background, which helps distinguish genuine signals from stochastic variability. Significance thresholds for wavelet coherence are typically determined through Monte Carlo simulations, in which surrogate datasets are generated to establish confidence intervals for the observed coherence values.. These analytical tools are particularly effective in detecting complex temporal structures and interrelations in time series data – for example, in salinity studies, where understanding the temporal patterns and shifts in dominant frequencies is crucial for analyzing transport dynamics (Grinsted et al., 2004).

RESULTS AND DISCUSSION

Training results of LSTM

At each location, the input data (i.e., discharge at Chau Doc and Tan Chau, and tide at Vung Tau) and output data (i.e., salinity at collected locations) were normalized firstly, and then normalized databases are split into 70% and 30% for the training and testing steps, respectively. In the training step, the appropriate values of hyper-parameters in the LSTM model are determined by trial-and-error method, resulting a total of 350 epochs, a dimensionality of 128 for the hidden units, and a learning rate set at 0.001. The trial-and-error approach was adopted due to its flexibility and practical effectiveness. Other methods liked grid search or Bayesian optimization offer comprehensive parameter tuning, however, they are computationally expensive (Nguyen et al., 2021). The chosen configuration demonstrated stable convergence and high predictive accuracy based on both dimensionless and dimensional evaluation metrics, supporting the adequacy chosing the trial-and-error approach in the context of the present study.

Figure 6 presents the comparison between observed and LSTM-estimated salinity at six selected locations during the training phase. The results indicate that the LSTM model effectively captures

the observed salinity patterns. The dimensional error metrics (including RMSE, MAE, and ME) range from -0.05 to 2.76 PSU (as reported in Table 3), corresponding to less than 8.01% of the observed salinity values at the monitoring stations. The correlation coefficient (*r*) between observed and predicted salinity exceeds 0.96 across all locations, demonstrating the model's strong ability to replicate temporal variations in salinity. Furthermore, the NSE values range from 0.93 to 0.98, underscoring the model's high predictive performance at all sites considered (Figure 7).

Regarding the estimated salinity in detail, among six locations of interest in the VMD, the plot of observations versus estimations of salinity at Phuoc Long shows apparent scatters, revealing complex transport processes of salinity at this location. The latter is result from the complex flow dynamics caused by multiple rivers and channels networks, interaction between river flow and tide, human activities, and local construction operations like culverts. The values of dimensional errors (i.e., RMSE, MAE, and ME) of estimated salinity range from 0.26 to 2.76 PSU (see Table 3 and Figure 7). The dimensionless errors (i.e., r and NSE) vary from 0.93 to 0.96, indicating a slight smaller value in comparison with other locations.

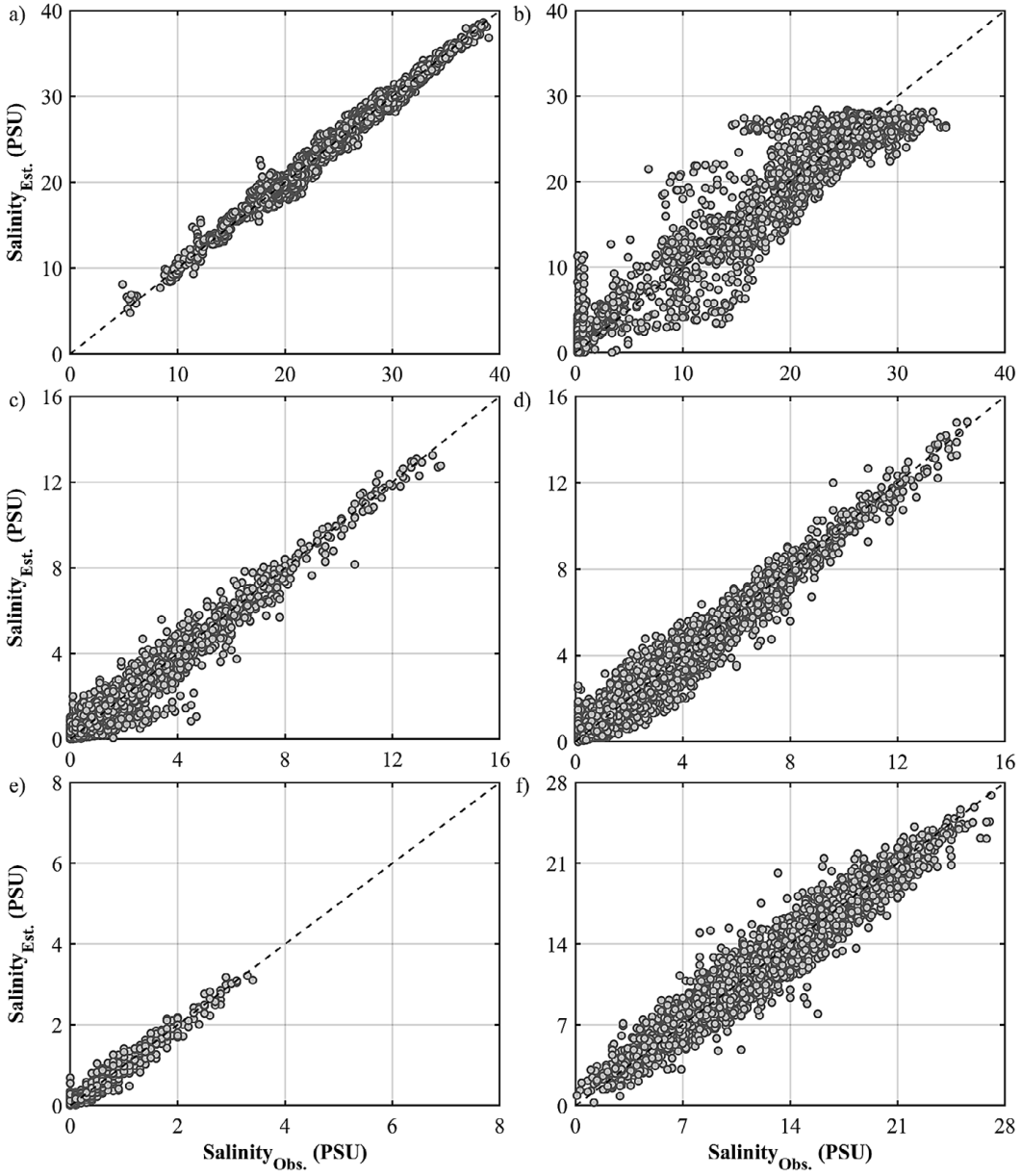

Figure 6. Observed versus estimated salinity, at: a) Ca Mau, b) Phuoc Long, c) Dai Ngai, d) Tra Vinh, e) My Tho, and f) Vam Kenh locations for training step of the LSTM

Table 3. Statistical indicator values of hourly salinity in the training phase

Locations	*RMSE*		*MAE*		*ME*		*r*	*NSE*
	PSU	%	PSU	%	PSU	%		
Ca Mau	0.75	1.92	0.57	1.46	0.042	0.11	0.992	0.984
Phuoc Long	2.76	8.01	2.00	5.79	0.261	0.76	0.965	0.930
Dai Ngai	0.49	3.56	0.33	2.39	0.015	0.11	0.975	0.950
Tra Vinh	0.49	3.35	0.35	2.38	-0.002	-0.01	0.981	0.962
My Tho	0.13	3.77	0.09	2.73	-0.002	-0.05	0.977	0.954
Vam Kenh	1.33	4.92	1.02	3.78	0.059	0.22	0.968	0.937

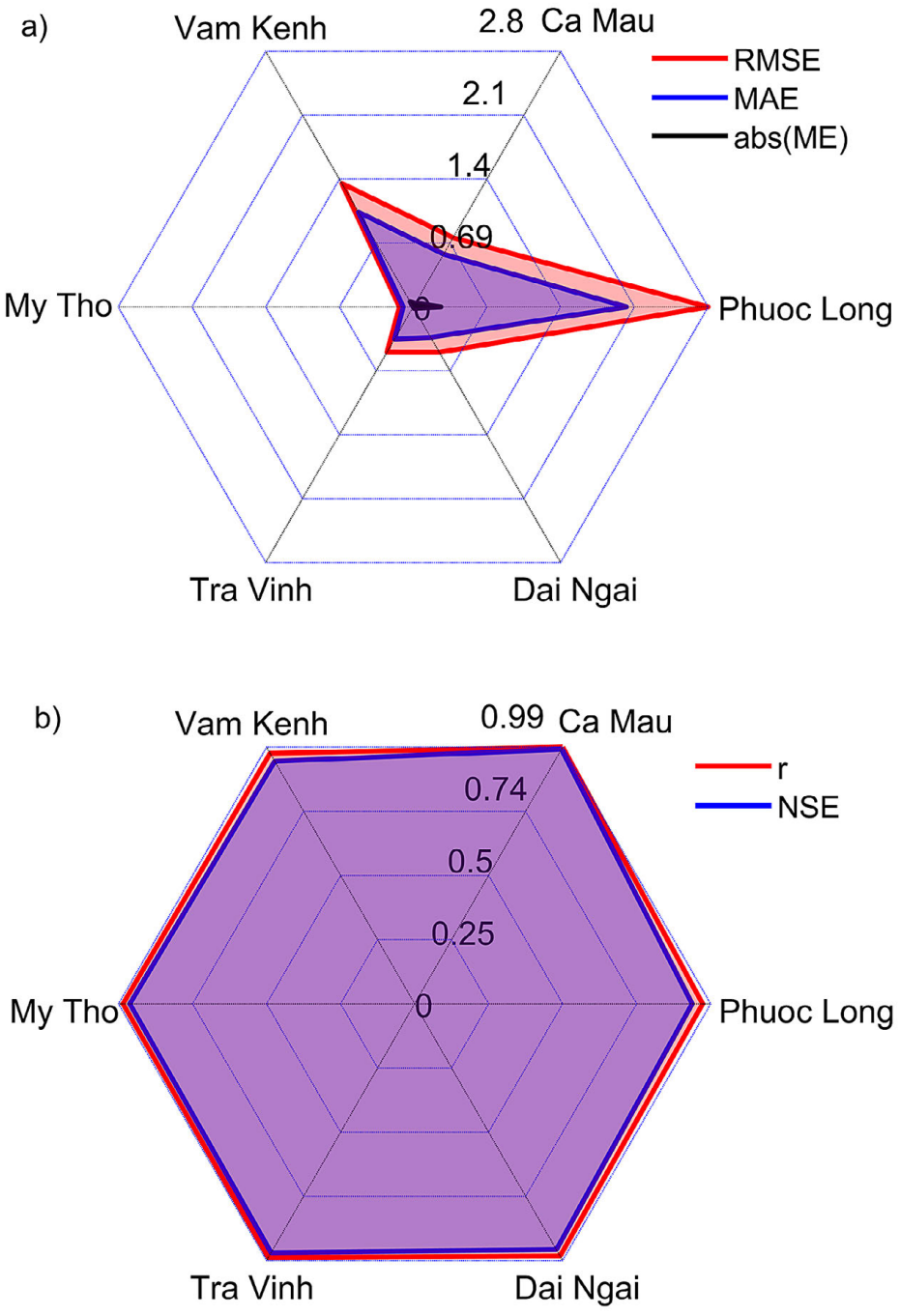

Figure 7. Radar plots for: a) dimensional and b) dimensionless performance metrics of hourly salinity for training step of the LSTM

Table 4. Statistical indicator values of hourly salinity in the testing phase

Locations	*RMSE*		*MAE*		*ME*		*r*	*NSE*
	PSU	%	PSU	%	PSU	%		
Ca Mau	1.09	2.77	0.85	2.16	-0.069	-0.18	0.980	0.961
Phuoc Long	3.80	10.04	2.40	6.36	-0.305	-0.81	0.920	0.839
Dai Ngai	0.40	5.57	0.27	3.80	-0.031	-0.43	0.956	0.913
Tra Vinh	0.41	4.28	0.27	2.77	0.006	0.07	0.965	0.931
My Tho	0.14	4.06	0.10	2.90	0.000	-0.01	0.986	0.973
Vam Kenh	1.26	5.47	0.98	4.25	-0.094	-0.41	0.961	0.924

Testing results of LSTM

In the testing step, the observed versus estimated salinity is shown Figure 8, while Figure 9 displays radar plots representing RMSE, MAE, ME, r and NSE indicators. A good agreement between observations and estimations of salinity is obtained for most locations. Overall, a strong agreement between observed and estimated salinity is observed across most locations. The RMSE values range from 0.14 to 3.80 PSU, whereas MAE varies between 0.10 and 2.40 PSU (Table 4). These error magnitudes represent approximately 10% of the observed salinity values at the measurement sites. Notably, the *ME* remains small, accounting for less than 0.5% of the observed values. The correlation coefficient *r* ranges from 0.92 to 0.98, indicating a very strong linear relationship between observed and estimated salinity. Additionally, the *NSE* values span from 0.84 to 0.97, reflecting the model's high capability in capturing the temporal variability of salinity. These findings demonstrate that the LSTM model is effective for estimating salinity at various locations within the VMD, using tidal data from Vung Tau and river discharge data from Tan Chau and Chau Doc as inputs.

Applications of LSTM model for estimating hourly salinity

Figure 10 shows the hourly time-series salinity at six locations in the period from 01/01/2024 to 30/06/2017. It is clearly observed that the LSTM model represents well the temporal variation of salinity in different hydrological conditions as well as in different seasons. In each year, large values of salinity are obtained in the period from January to May, while small values of salinity are archived in the period from June to December. In addition, the estimated salinity shows clearly impact of the tide, especially at Vam Kenh location, where salinity discrepancies can reach up to over 10 PSU during neap-spring tidal cycles. Tide also has a strong influence at four locations named Dai Ngai, Tra Vinh, My Tho, Vam Kenh since all of which are close to the East Sea. High salinity is observed during spring tides, with the values ranging from 0 to 25 PSU at these four locations. In neap tides, small values of salinity are obtained. At Ca Mau and Phuoc Long locations lied in land areas, river flow (or discharge) and local conditions mainly drive the variation of salinity in time.

Salinity periodicity

At six selected locations in the VMD, the combined observed and modeled hourly salinity records were analyzed via wavelet transform to identify temporal periodicities, producing the wavelet power spectrum depicted.in Figure 11. Within each figure panel, the solid black line delineates the cone of influence, highlighting edge-affected regions where variance or spectral power is diminished. The solid black contours enclose areas of the wavelet power spectrum that exceed the 95% confidence level under a yellow-noise null hypothesis. The amplitude of the wavelet power directly indicates the intensity of periodic fluctuations within the salinity data. These results illuminate the principal variability modes and reveal significant temporal patterns in the salinity time series..

Figure 11a-b reveal the hourly salinity records at Ca Mau and Phuoc Long, respectively that exhibit significant diurnal oscillations at the 95% confidence level. By contrast, the wavelet power associated with the annual cycle is markedly weaker. The inter-daily oscillations range from 4 to 32 hours in the period from January to May, while they shorten to approximately 2 to 16 hours during in the period

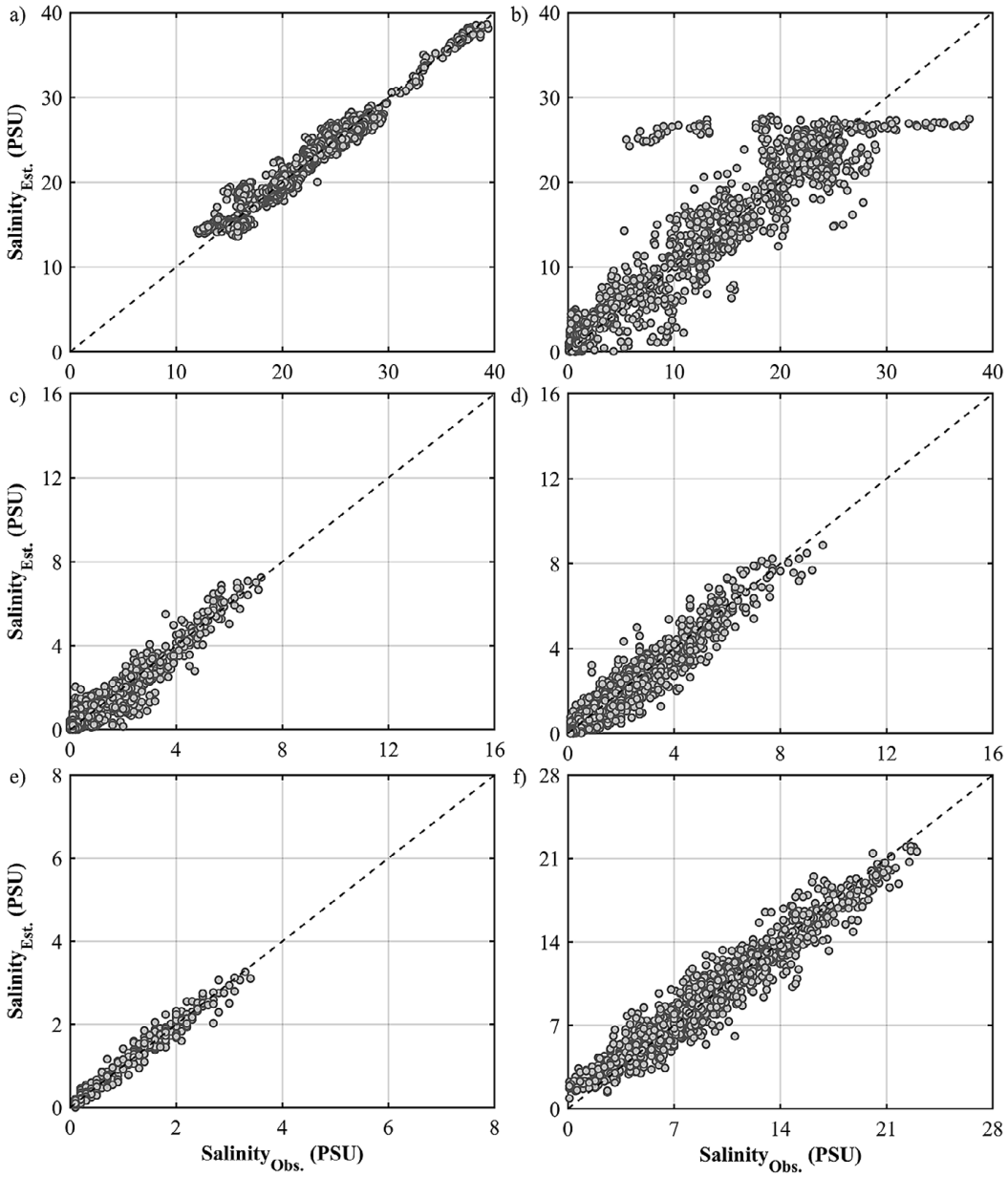

Figure 8. Observed versus estimated salinity, at: a) Ca Mau, b) Phuoc Long, c) Dai Ngai, d) Tra Vinh, e) My Tho, and f) Vam Kenh locations for testing step of the LSTM

from June to December. At Dai Ngai, Tra Vinh, and My Tho locations (Figure 11c-e), oscillations with a four-months period are observed, with significant peaks around 2016, also at the 95% confidence level. At Vam Kenh location (Figure 11f), inter-daily oscillations of 2 days are evident throughout the entire year. The wavelet power spectrum of hourly salinity also clearly illustrates the impact of tidal forces on salinity variations in time. Indeed, strong tidal impacts are observed from January to May, while weaker impacts occur from June to December. This pattern corresponds to the low and high discharges at Tan Chau and Chau Doc, respectively.

Wavelet coherence between salinity and discharge

Salinity at the selected VMD locations is governed by a combination of upstream discharge, East Sea tidal forcing, local anthropogenic activities, and intrinsic salinity transport mechanisms. To elucidate the time–frequency characteristics of the discharge – salinity relationship, a wavelet coherence analysis was performed, with results presented in Figure 12 and Figure 13. This approach enables the identification of nonstationary associations between river discharge and salinity at Tan Chau and Chau

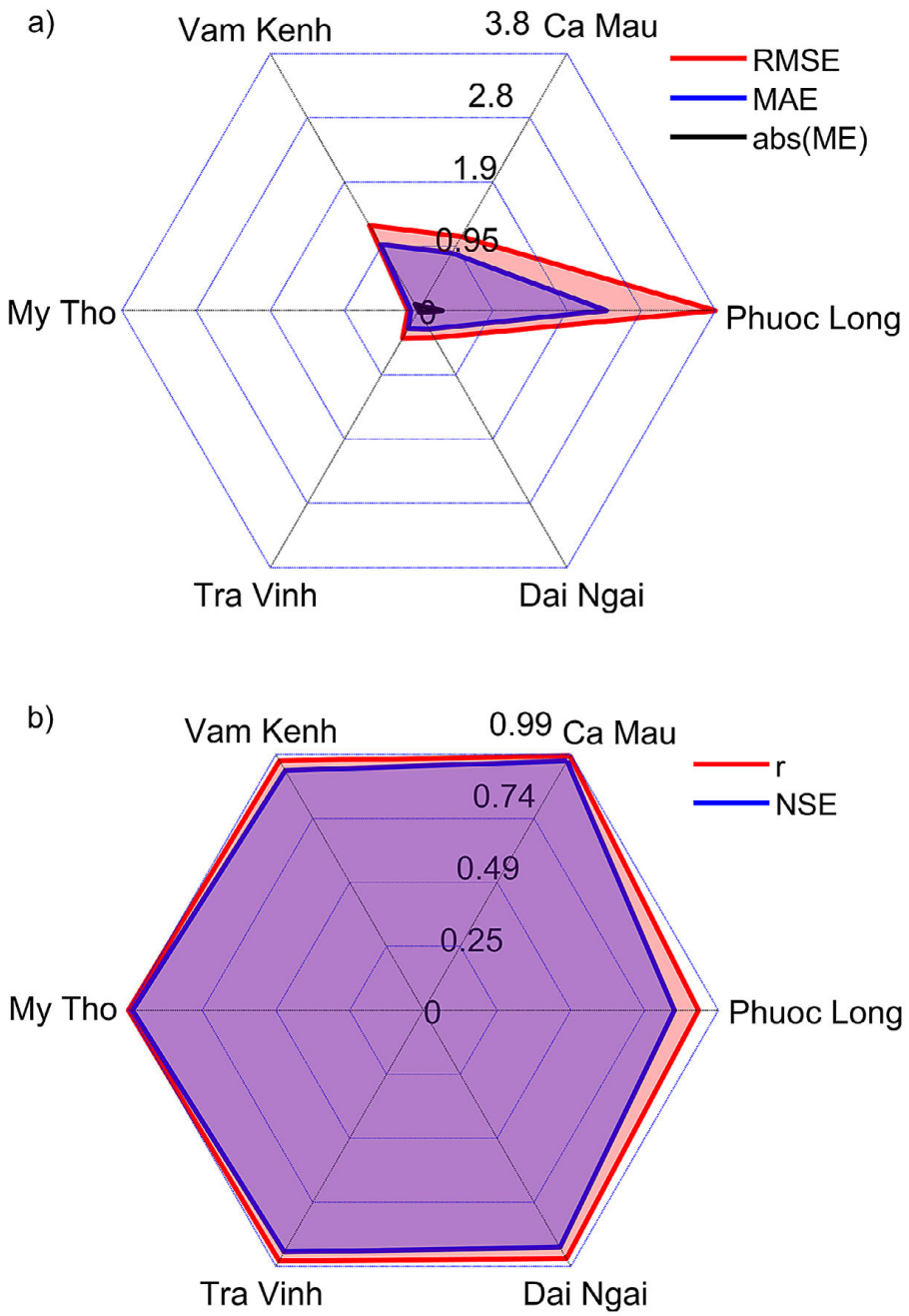

Figure 9. Radar plots for: a) dimensional and b) dimensionless performance metrics of hourly salinity for testing step of the LSTM

Doc, offering detailed insights into how variations in flow regime propagate through to salinity fluctuations. The findings thereby advance our understanding of salinity dynamics at these critical monitoring locations.

Within each panel, the horizontal axis corresponds to the temporal span of the observations, while the vertical axis represents the oscillation period. To address edge effects, a lightly shaded cone of influence delineates regions where spectral estimates may be unreliable. Solid black contours mark zones of significant coherence between discharge and salinity ($p < 0.05$). Directional arrows depict the phase relationship (θ) between the two signals, with arrows pointing right ($\theta = 0$) indicating in-phase behavior and arrows pointing left ($\theta = \pi$) indicating anti-phase. Wavelet coherence serves as a localized correlation metric: values near unity denote strong synchronization of discharge and salinity, whereas values approaching zero reflect a lack of association. This wavelet-based analysis therefore provides detailed insights

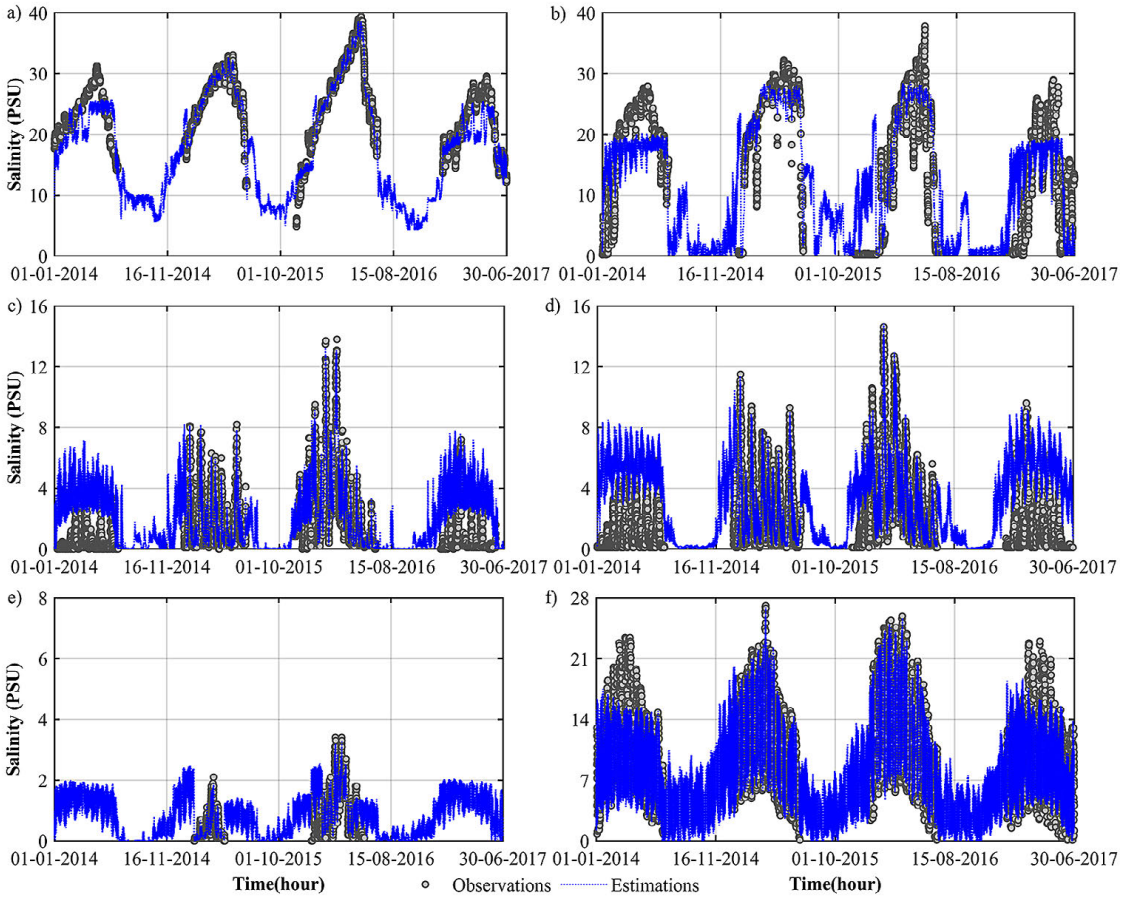

Figure 10. Time series of observed versus estimated salinity, at: a) Ca Mau, b) Phuoc Long, c) Dai Ngai, d) Tra Vinh, e) My Tho, and f) Vam Kenh locations in the period from 2014–2017

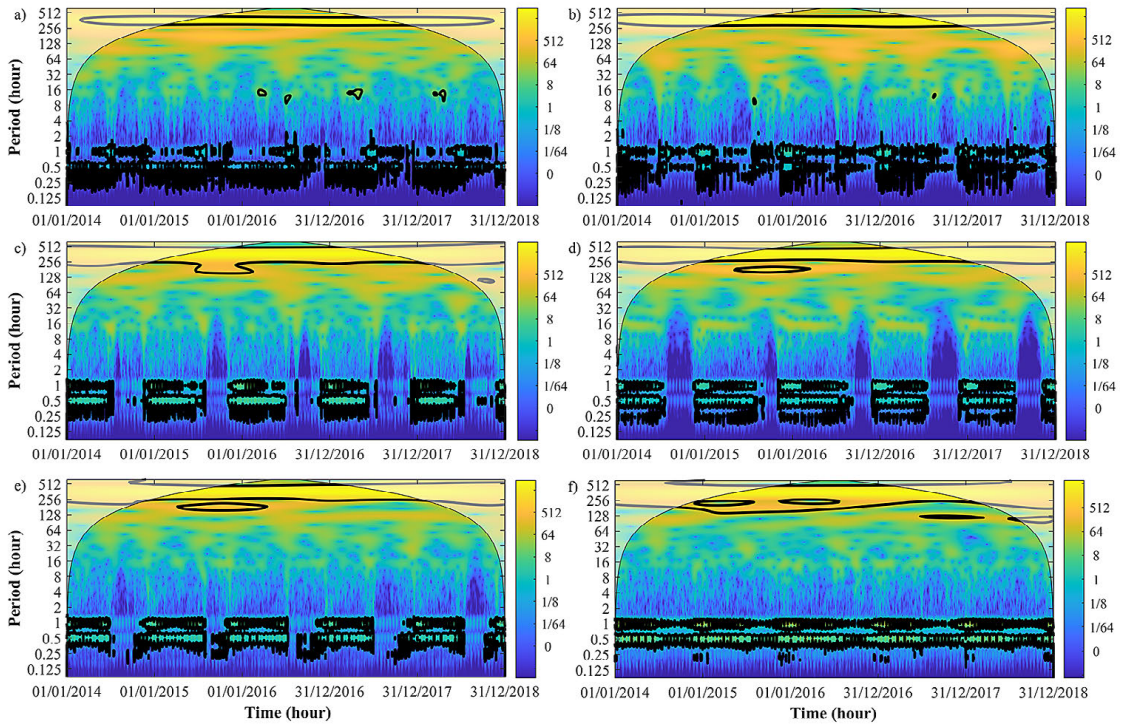

Figure 11. Wavelet power spectrum of hourly salinity, at: a) Ca Mau, b) Phuoc Long, c) Dai Ngai, d) Tra Vinh, e) My Tho, and f) Vam Kenh locations

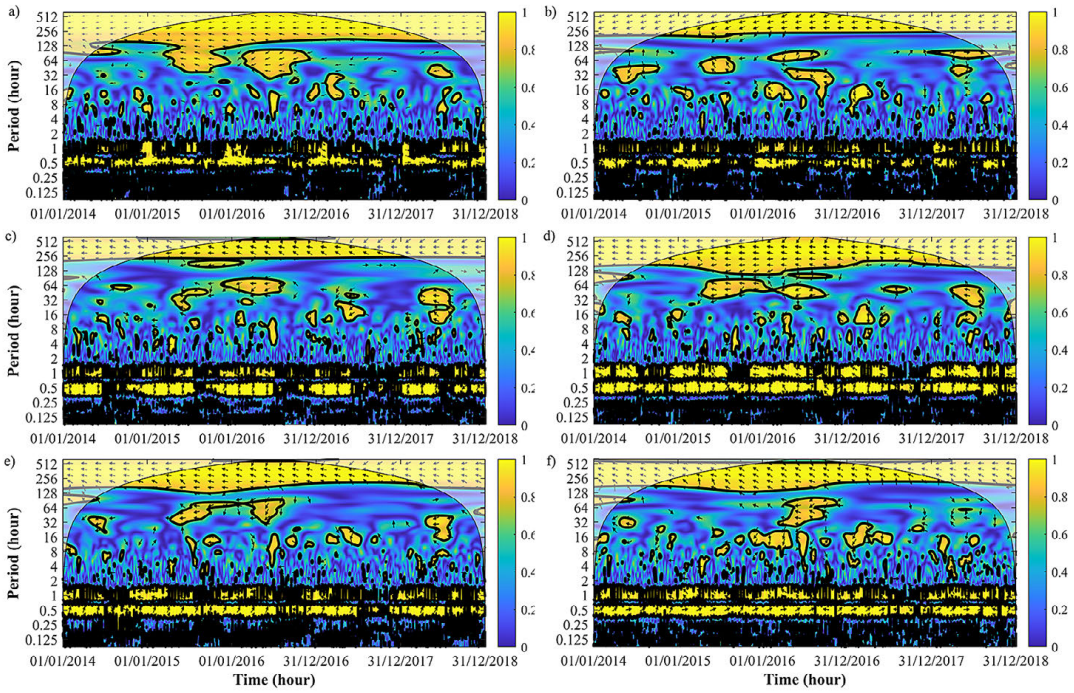

Figure 12. Wavelet coherence between hourly discharge at Chau Doc and salinity, at a) Ca Mau, b) Phuoc Long, c) Dai Ngai, d) Tra Vinh, e) My Tho, and f) Vam Kenh locations

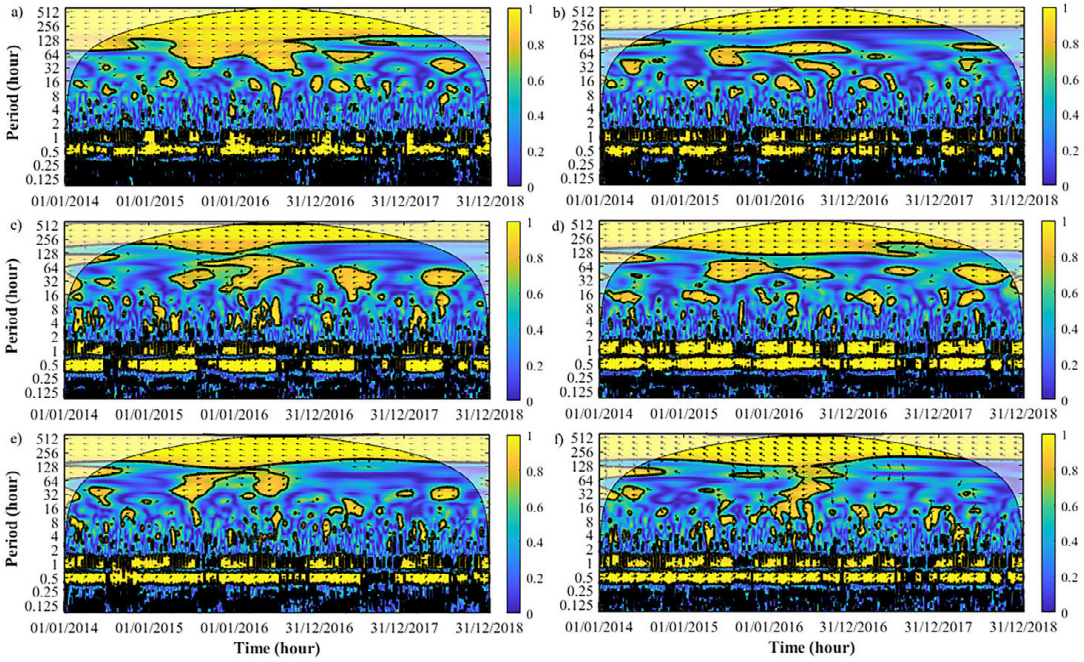

Figure 13. Wavelet coherence between hourly discharge at Tan Chau and salinity, at a) Ca Mau, b) Phuoc Long, c) Dai Ngai, d) Tra Vinh, e) My Tho, and f) Vam Kenh locations

into the time – frequency coupling of flow and salinity dynamics at the study locations.

Figure 12 and Figure 13 demonstrate a statistically significant coherence between river discharge and salinity at all examined sites. The coherence analysis reveals scale-dependent behavior: it remains significant over longer periods (approximately one to four months), with the most stable and robust multi-scale coherence occurring at seasonal time scales, thereby underscoring the strong linkage between seasonal discharge patterns and salinity variations. At shorter

time scales, the coherence exhibits more complex and variable structures, suggesting that daily discharge contributions to salinity fluctuations are less consistent. These results highlight the dynamic, multi-scale coupling between discharge and salinity and emphasize the need to consider diverse temporal scales when characterizing salinity dynamics in the VMD.

Wavelet coherence between salinity and tide

Figure 14 shows the wavelet coherence between tide at Vung Tau and salinity at six studied locations. The coherence analysis reveals that the relationship between salinity and tidal signals exhibits complex scale-dependent behavior, suggesting considerable variability in tidal contributions to salinity fluctuations. These results emphasize the dynamic and multi-scale nature of tide–salinity interactions, highlighting the need to account for such variability when assessing salinity dynamics at specific locations within the VMD. As shown in Figure 14a-b for Ca Mau and Phuoc Long locations, salinity changes in response to tidal changes with a few hours to a day. In addition, coherence is high in the period from June to December of years, revealing that seasonal flow also influences both the magnitude and temporal patterns of salinity. At Dai Ngai, Tra Vinh, My Tho, and Vam Kenh locations (in Figure 14c-f), salinity vary in response to tidal variations with a couple of days, depending on the locations of interest. The value and variability of salinity in time also clearly depict significant tidal impacts. In addition, the wavelet coherence results also show spatial variability of salinity caused by local conditions and factors.

A comprehensive evaluation of salinity fluctuations, short-term variability, and the wavelet coherence between salinity and both river discharge and tidal signals at multiple VMD locations is expected to yield a detailed characterization of salinity dynamics. These insights will support: (i) the development of salinity early-warning systems, (ii) investigation of salinity alterations driven by upstream discharge fluctuations and downstream tidal influences, and (iii) adaptation of agricultural management practices to accommodate salinity variability. Moreover, wavelet cohenrence analysis enables the detection and characterization of salinity intrusion responses to changes in external forcings (e.g., tide, river discharge), capturing both the magnitude and time lag of these responses (Gong et al., 2022). This tool is particularly valuable for analyzing highly dynamic processes in transient systems far from equilibrium, providing an unified time frequency framework to investigate salinity processes across different timescales in deltas and coastal regions. Salinity variations across the VMD exhibit particularly strong correlations with river discharge at Tan Chau and Chau Doc, as well as with tidal fluctuations at Vung Tau. This spatial heterogeneity underscores the role of seasonal forcing and location-specific factors in modulating the interplay between fluvial and marine influences. A clear understanding of these linkages is essential for optimizing water supply strategies and managing water resources throughout the delta when using numerical or data-driven models for predicting salinity associated with different riverine and marine forcing.

Traditional methods, such as numerical models, have been used to simulate salinity in the VMD. Tran Anh et al. (2018) applied one- and two-dimensional hydrodynamic models to predict future salinity intrusion in the Hau river, a main branch in the VMD. Tran et al. (2023) combined the one-dimensional model with remote sensing images to quantify saltwater intrusion in the region. Vu et al. (2024) implemented two-dimensional model to simulate hydrodynamic changes and salinity intrusion in the lower VMD under climate change-induced sea level rise and upstream river discharge. However, compared to the model presented in this study, traditional methods require extensive data, including bathymetry, river flow, tides, and salinity to determinine appropriate values of parameters in the both hydrodynamic and salinity transport modules. Moreover, they are often suffered from errors due to uncertainties in inputs, model structures, and parameterization (Tran et al., 2023). In contrast, the proposed approach is highly effective and practical as it relies only on available datasets of flow and tide as inputs. This makes the approach a valuable tool that can be easily applied to other deltas and coastal regions.

Several factors such as temperature, rainfall, and upstream human activities can influence salinity. Land-use and land-cover changes (e.g., urban expansion, deforestation) alter hydrological processes including runoff generation, evapotranspiration, and infiltration. These changes, in turn, affect discharge patterns and salinity regimes. Human-induced modifications within or upstream of the delta (e.g., dam operations,

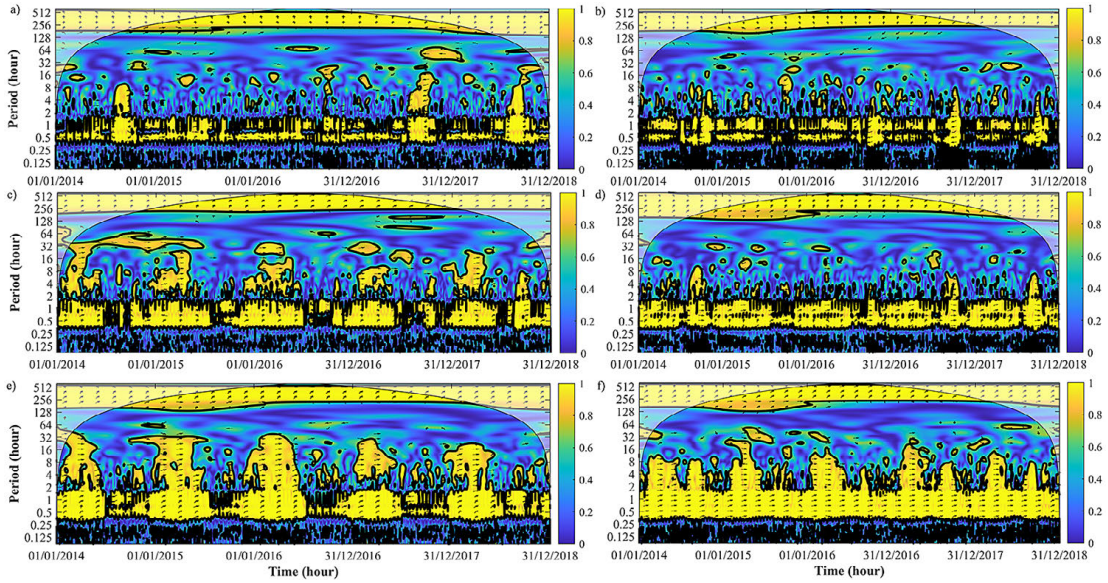

Figure 14. Wavelet coherence between hourly water level at Vung Tau and salinity, at a) Ca Mau, b) Phuoc Long, c) Dai Ngai, d) Tra Vinh, e) My Tho, and f) Vam Kenh locations

canal construction or the implementation of rice-shrimp farming systems) can significantly influence flow pathways and the extent of saltwater intrusion(Anthony et al., 2015). Incorporating land-use change data also enables the simulation of alternative development scenarios, thereby allowing for the evaluation of the potential salinity impacts associated with future land management policies or conservation efforts. This enhances the model's capacity to represent deltaic hydrology more accurately, identify anthropogenic drivers, and support scenario-based planning.

Rainfall variability (both spatial and temporal) can influence freshwater inputs and deltaic salinity dynamics. Satellite-derived precipitation products (e.g., CHIRPS, IMERG, CMORPH) can improve the model's performance by providing high-resolution input data and enabling a more dynamic representation of freshwater inflows (Vu et al., 2024). These datasets offer extensive spatiotemporal coverage, which is particularly valuable in extensive regions liked the VMD. Integrating rainfall data into the model allows for capturing both direct and indirect freshwater contributions to salinity variations, including overland flow and river discharge. Indeed, real-time or near-real-time satellite precipitation data can be assimilated to improve short-term salinity forecasting, supporting operational decision-making in water resources management.

Outputs from climate models (e.g., CMIP6) enables the salinity model to simulate future scenarios under various climate change trajectories. This capability is essential for developing effective adaptation strategies in deltaic regions that are highly susceptible to sea-level rise and changes in hydrological regimes (Smajgl et al., 2015; Nguyen et al., 2023a). Climate change introduces nonstationary trends in key variables such as temperature, precipitation, and runoff. By utilizing climate projections, the model can account for these evolving conditions, thereby overcoming the limitations of assuming a static climate. Furthermore, such projections facilitate the assessment of salinity dynamics under extreme events, including prolonged droughts and intensified wet seasons, thereby enhancing the model's robustness and resilience.

In summary, integrating climate model projections, satellite-derived precipitation, and land-use change datasets can substantially improve the salinity model by (i) providing a more comprehensive representation of both environmental and anthropogenic drivers, (ii) enhancing its transferability and generalizability within delta and temporal scales, and (iii) increasing its utility for long-term planning, climate adaptation, and integrated water resources management. The exclusion of these factors may introduce uncertainty or constrain the model's ability to capture certain hydrological events. These factors were not incorporated in this study due to limited data availability. However, they can be integrated into the calculations once the necessary datasets become available.

CONCLUSIONS

This study proposed a method for estimating salinity in the VMD based on the LSTM model and combines wave transformation analysis to examine periodic changes. Salinity at six locations in the period 01/01/2014 to 30/06/2017 and available hydrological data (i.e., discharge at Tan Chau and Chau Doc and water level at Vung Tau) in the period from 01/01/2024 to 31/12/2018 were collected and used for different calculations and analyses. The principal findings of this study can be summarized as follows:

- The LSTM model, a widely used variant of RNN, was first implemented to simulate salinity at six selected locations before being applied to the given period. In the training and testing steps, the model demonstrated a high degree of fidelity in replicating the observed salinity values at six locations in the VMD during both training and testing steps. The correlation coefficient r and NSE exceeded 0.92 and 0.84, respectively in both steps. Dimensional errors including RMSE, MAE, and ME remained below 10% of the observed salinity magnitude at the investigated locations. In the application phase, the model also effectively captured the temporal variation of salinity under different hydrological conditions and seasons.
- Wavelet analysis was applied to hourly time series of salinity data from both observations and estimations from 01/01/2014 to 31/12/2018 for investigating its periodic characteristics. The study revealed that dominant oscillatory modes in salinity were concentrated at time scales of approximately two hours to two days.
- At all six examined locations, salinity and discharge exhibit statistically significant coherence with complex, scale-dependent behavior. Notably, this coherence persists over extended periods, particularly within the one-to-four-month band. In addition, seasonal flow influenced both the magnitude and temporal patterns of salinity in a coupled manner.
- Using the wavelet analysis, wavelet coherence between water level at Vung Tau and salinity at six locations were also examined, revealing that salinity varies in response to water level variations with a couple of hours to a couple of days. Analyzing salinity fluctuations, long-term trends, and short-term variability across multiple sites in the VMD is anticipated to yield comprehensive insights into the spatial and temporal dynamics that influence salinity distribution throughout the delta.
- The proposed approach used only available flow and tide datasets to estimate salinity, making it a valuable tool for application in other deltas and coastal regions. The resuls from the case study of the VMD clearly demonstrated good predictive performance and the periodic characteristics of salinity. Thus, it is strongly believed that the proposed approach is highly effective and practical compared to traditional methods like numerical models.

Acknowledgments

The authors gratefully acknowledge the Southern Regional Hydro-Meteorological Center for providing the hydrological data used in this study.. This work was funded by project "Research and propose solutions for freshwater storage to support socio-economic development in the Vietnamese Mekong Delta", supported by the Vietnam Ministry of Science and Technology under grant number ĐTĐL.CN-37/22.

REFERENCES

1. Addison, P. S. (2017). *The Illustrated Wavelet Transform Handbook: Introductory Theory and Applications in Science, Engineering, Medicine and Finance*. Institute of Physics Publishing, Bristol, UK, https://doi.org/10.1201/9781315372556
2. Anthony, E.J., B. Guillaume, M. Besset, G. Marc, P. Dussouillez, V.L Nguyen. (2015). Linking rapid erosion of the Mekong River delta to human activities. *Scientific Reports*, 5, 14745, https://doi.org/10.1038/srep14745
3. Gong, W., Lin, Z., Zhang, H., Lin H. (2022). The response of salt intrusion to changes in river discharge, tidal range, and winds, based on wavelet analysis in the Modaomen estuary, China. *Ocean and Coastal Management*, 219, 106060, https://doi.org/10.1016/j.ocecoaman.2022.106060
4. Grinsted, A., Moore, J.C., Jevrejeva S. (2004). Application of the cross wavelet transform and wavelet coherence to geophysical time series. *Nonlinear Processes in Geophysics*, 11, 561–566, https://doi.org/10.5194/npg-11-561-2004
5. Hochreiter, S., Schmidhuber, J (1997). Long short–term memory. *Neural Computation* 9, 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735
6. Nguyen, M.T., Renaud, F.G., Sebesvari, Z. (2019). Drivers of change and adaptation pathways of

agricultural systems facing increased salinity intrusion in coastal areas of the Mekong and Red River deltas in Vietnam. *Environmental Science & Policy*, 92, 331–348, https://doi.org/10.1016/j.envsci.2018.10.016

7. Nguyen, T.H., Nguyen, H.Y., Balistrocchi, M., Peli, M., Vu, M.C., Ranzi, R. (2023a). Salinity dynamics under different water management plans coupled with sea level rise scenarios in the Red River Delta, Vietnam. *Journal of Hydro-environment Research*, 51, 67–81, https://doi.org/10.1016/j.jher.2023.10.003
8. Nguyen, T.G., Tran, N.A., Vu, P.L., Nguyen, Q.H., Nguyen, H.D., Bui, Q.T. (2021). Salinity intrusion prediction using remote sensing and machine learning in data-limited regions: A case study in Vietnam's Mekong Delta. *Geoderma Regional*, 27, e00424, https://doi.org/10.1016/j.geodrs.2021.e00424
9. Nguyen, H.D., Pham Van, C., Nguyen, Q-H., Bui, Q-T. (2023b). Daily streamflow prediction based on the long short-term memory algorithm: a case study in the Vietnamese Mekong Delta. *Journal of Water and Climate Change,* 14 (4), 1247–1267, https://doi.org/10.2166/wcc.2023.419
10. Niu, W.-J., Z.-K. Feng, W.-F. Yang, J. Zhang (2020). Short-term streamflow time series prediction model by machine learning tool based on data preprocessing technique and swarm intelligence algorithm. *Hydrological Sciences Journal,* 65(15)**,** 2590–2603, https://doi.org/10.1080/02626667.2020.1828889
11. Pereira C. S., Lopes I., Abrantes I., Sousa J. P., Chelinho S. (2019). Salinization effects on coastal ecosystems: a terrestrial model ecosystem approach. *Phil. Trans. R. Soc. B*37420180251, https://doi.org/10.1098/rstb.2018.0251
12. Pham Van, C., Nguyen-Ngoc, D. (2022). Multiple linear regression and long short-term memory for evaluating water levels in irrigation and drainage systems: an application in the Bac Hung Hai irrigation and drainage system, Vietnam. *Water Supply,* 22(12), 8587–8602, https://doi.org/10.2166/ws.2022.386
13. Pham Van, C., Le, H., Van Chin, L. (2023). Estimation of daily suspended sediment concentration in the Ca River Basin using a sediment rating curve, multiple regression, and long short-term memory model. *Journal of Water and Climate Change,* 14(12), 4356–4375, https://doi.org/10.2166/wcc.2023.229
14. Tran Anh, D., Hoang, L. P., Bui, M. D., Rutschmann, P. (2018). Simulating future flows and salinity intrusion using combined one-and two-dimensional hydrodynamic modelling – The case of Hau River, Vietnamese Mekong delta. *Water*, 10(7), 897, https://doi.org/10.3390/w10070897
15. Tran, A.P., Son, D.H., Duc, N.A., Pham Van, C., Nguyen, T.T., Trang, M.C., Nguyen, N.A., Le, P.V.V., Pham, H.V. (2023). Bayesian merging of numerical modeling and remote sensing for saltwater intrusion quantification in the Vietnamese Mekong Delta. *Environmental Monitoring and Assessment* 195, 1415, 1–19, https://doi.org/10.1007/s10661-023-11947-7
16. Tran, T.T., Pham, N.H., Pham, Q.B., X.Q. Ngo, D.L. Nguyen, P.N. Nguyen, B.K. Veettil (2022). Performances of different machine learning algorithms for predicting saltwater intrusion in the Vietnamese Mekong Delta using limited input data: A study from Ham Luong River. *Water Resourcces,* 49, 391–401, https://doi.org/10.1134/S0097807822030198
17. Tran, D.D., Pham, T.B.T., Park, E., Phan, T.T.H., Duong, B.M., Wang, J. (2024). Extent of saltwater intrusion and freshwater exploitability in the coastal Vietnamese Mekong Delta assessed by gauging records and numerical simulations. *Journal of Hydrology*, 630, 130655, https://doi.org/10.1016/j.jhydrol.2024.130655
18. Torrence, C., Compo, G P. (1998). A practical guide to wavelet analysis. *Bulletin of the American Meteorological Society*, 79, 61–78. https://doi.org/10.1175/1520-0477(1998)079<0061:APGTWA>2.0.CO;2
19. Saccotelli, L, Verri, G., de Lorenzis, A., Cherubini, C., Caccioppoli, R., Coppini, G., Maglietta, R. (2024). Enhancing estuary salinity prediction: A machine learning and deep learning base approach. *Applied Computing and Geoscience* 23, 100173, 1–12, https://doi.org/10.1016/j.acags.2024.100173
20. Smajgl, A., Toan, T., Nhan, D.K., Ward, J., Trung, N.H., Tri, L.Q., Tri, V.P.D., Vu, P.T. (2015). Responding to rising sea levels in the Mekong Delta. *Nature Clim Change* 5, 167–174, https://doi.org/10.1038/nclimate2469
21. Vu, M.T., Luu, C., Bui, D.Q., Vu, Q.H., Pham, M.Q. (2024). Simulation of hydrodynamic changes and salinity intrusion in the Lower Vietnamese Mekong Delta under climate change-induced sea level rise and upstream river discharge. *Regional Studies in Marine Science,* 103749, https://doi.org/10.1016/j.rsma.2024.103749