

Article

Integrating Matrix Decomposition into Deep Learning for Vessel Traffic Flow Prediction in IoT Industries

Chao Zhang^{1,2,3,*}, Bi Yu Chen^{1,2,3} , Zehao Yuan^{1,2,3} and Qinyang Chen⁴

¹ State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China; chen.biyu@whu.edu.cn (B.Y.C.); yuanzehao@whu.edu.cn (Z.Y.)

² Collaborative Innovation Center of Geospatial Technology, Wuhan University, Wuhan 430079, China

³ Geocomputation Center for Social Sciences, Wuhan University, Wuhan 430079, China

⁴ Changjiang Waterway Bureau Survey Center, Wuhan 430010, China; cjhdccy@163.com

* Correspondence: zhangchao1@whu.edu.cn

Abstract

Accurate forecasting of vessel traffic flow (VTF) is essential for modern maritime and port management, as it improves route-planning efficiency, reduces congestion and collision risks, and optimizes port operations. This study proposes a novel deep learning framework, namely, the Bidimensional Empirical Mode Decomposition–Nocal Convolutional Neural Network–Transformer (BEMD–NocalCNN–Transformer), for high-precision VTF prediction. The proposed framework first applies the BEMD algorithm to decompose the original time-series data into high- and low-frequency components. The NocalCNN module is then employed to extract spatial features from each component, while the Transformer module captures temporal dependencies and predicts future traffic-flow trends. The final predictions are obtained by aggregating the outputs of the high- and low-frequency components. Sensitivity analyses are conducted on key parameters, including input sequence length, learning rate, number of iterations, and convolution kernel size, to optimize the model configuration. To comprehensively evaluate the proposed framework, SVM, BPNN, RNN, LSTM, GRU, Transformer, WVMA-LSTM, and NocalCNN–Transformer were implemented and evaluated using the same CFD and Wuhan datasets, data preprocessing procedures, training–testing partitions, prediction settings, and evaluation metrics. The experimental results demonstrate that the proposed model outperforms the benchmark models and achieves substantially lower prediction errors for both the Caofeidian Promontory (CFD) and Wuhan waterways. These findings demonstrate consistent prediction performance of the proposed framework and provide a robust technical foundation for intelligent maritime traffic management and port operation optimization.

Keywords: vessel traffic flow; Bidimensional Empirical Mode Decomposition; Nocal Convolutional Neural Network; Transformer



Academic Editor: Xinqiang Chen

Received: 2 July 2026

Revised: 25 July 2026

Accepted: 26 July 2026

Published: 29 July 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 rapid expansion of the global economy has led to a substantial increase in international trade activities [1]. Various modes of trade transport, such as air, land, and maritime transport, are used to support international trade [2,3]. Among these, maritime transport plays a pivotal role because of its ability to move large volumes of goods at relatively low cost, thereby attracting substantial attention [4,5]. According to the International Association of Maritime Economists, approximately 65% to 85% of global trade by weight is conducted by sea [6]. However, the rapid growth of the maritime sector and the increasing

number of vessels have significantly intensified traffic in waterways and port areas, posing considerable challenges to navigation management and maritime surveillance [7,8]. As a result, efficient navigation management and monitoring systems are essential to ensure maritime safety, maintain smooth traffic flow, and protect the marine environment. In this context, accurate forecasting of vessel traffic flow (VTF) is crucial for optimizing port operations, improving cargo handling efficiency, and reducing the risk of ship collisions and other maritime incidents.

VTF data comprises detailed records of vessel trajectories and operational characteristics, including position, speed, heading, and vessel type [9,10]. It quantifies the frequency of vessel passages through a specific location over a defined time interval [11]. The volume of VTF data reflects the density of maritime traffic in a given area, with elevated VTF values indicating a higher likelihood of vessel collisions and navigational congestion. Accordingly, accurate VTF forecasting is essential for efficient traffic management, enhanced operational performance, and the mitigation of collision risks. The acquisition of VTF data is enabled by the Automatic Identification System (AIS). Ships obtain their position and speed information via shore-based or satellite-based receiving stations, which transmit the data to maritime authorities and shipping companies [12,13]. Ground servers collect AIS data, including timestamps, vessel trajectories, ground headings, and speeds, to construct VTF datasets, which are crucial for traffic flow (TF) prediction [14]. AIS data with high spatio-temporal resolution enables real-time monitoring of navigation routes, thereby enhancing navigational safety and improving the maritime environment [15]. In addition to traffic monitoring and prediction, AIS data has been used to extract operational characteristics and support data-driven maritime management. Bai et al. derived short- to medium-term operational indicators, including relative trip distance, fleet repositioning flexibility, and trading diversity, from AIS records and investigated their effects on shipping risk using a Bayesian Belief Network. Their study demonstrated that AIS-derived information can support not only vessel movement analysis but also operational decision-making and risk management [16]. However, VTF data often exhibits non-stationary and nonlinear characteristics, making it challenging for predictive models to capture their intrinsic patterns and achieve high forecasting accuracy.

Deep learning techniques have demonstrated exceptional performance across various domains, including image recognition [17–20], semantic segmentation [21–23], and time-series prediction [24,25]. These methods effectively capture complex nonlinear relationships and extract advanced features through multilayer neural network architectures [26,27], making them particularly suitable for large-scale data processing and complex task modeling [28–30]. In VTF prediction, deep learning models can effectively represent the spatio-temporal dynamics of waterway networks and vessel operational behavior.

Numerous studies have applied deep learning techniques to VTF-related tasks. For example, Li introduced a novel learning-based prediction network that improves Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM) networks through similarity-based groupings, thereby revealing hidden relationships and improving prediction accuracy and stability [9]. Zhou et al. proposed three deep learning solutions for forecasting vessel traffic in upstream and downstream directions within a specific region, including a CNN, an LSTM network, and an integrated Bidirectional LSTM-CNN (BDLSTM-CNN) [31]. Mandalis et al. addressed the VTF problem by considering it as both a vessel route forecasting task and a flow sequence prediction challenge, using deep learning techniques on authentic datasets [32]. Li and Ren suggested a prediction method based on a Long Short-Term Memory Encoder–Decoder (LSTM-ED) for multi-step VTF prediction, demonstrating superior performance compared to baseline methods [33]. Despite the growing application of deep learning in VTF prediction, many existing studies do not

adequately account for the spatio-temporal complexity of the problem. Given the nonlinear and aperiodic nature of VTF data, it is essential to extract more stable and representative features to improve prediction accuracy.

In this study, we propose a novel prediction model called the Bidimensional Empirical Mode Decomposition–Nocal Convolutional Neural Network–Transformer (BEMD–NocalCNN–Transformer). The model first applies the BEMD algorithm to preprocess and decompose the original two-dimensional traffic flow data, effectively converting non-stationary signals into more stationary components. This decomposition enhances the model's ability to capture underlying patterns and significantly improves prediction accuracy. To address the challenges posed by data nonlinearity and multiple influencing factors, we construct a spatio-temporal prediction model based on the NocalCNN–Transformer architecture. By integrating NocalCNN for spatial feature extraction with a Transformer module for capturing temporal dependencies, the proposed model demonstrates enhanced generalization capability in VTF forecasting.

VTF prediction methods draw heavily on forecasting techniques used to predict road TF. Section 2 offers an overview of advanced TF prediction methods and their relevance to the current study. Key definitions, problem statements, and the methodology for VTF calculation. The framework and implementation steps of the proposed BEMD–NocalCNN–Transformer prediction method are detailed in Section 3. Section 4 presents comprehensive experiments: the first part assesses the impact of different network parameters on the predictive accuracy of the BEMD–NocalCNN–Transformer, while the second part conducts comparative analyses with other commonly employed prediction approaches. Finally, Section 5 concludes the paper by summarizing the findings and proposing future research directions.

2. Literature Review

Advancements in VTF forecasting are grounded in progress made in TF prediction research. This chapter reviews recent developments in TF forecasting and examines their influence on VTF prediction. Subsequently, it explores the application of deep learning methods in VTF forecasting. Finally, the chapter highlights the significance of the findings presented in this study.

2.1. Deep Learning-Based Traffic Flow Prediction

Accurate TF prediction plays a crucial role in modern transportation systems [34–37], enabling vehicles to avoid congestion and supporting real-time traffic data collection for the optimization of traffic operations [38–40]. The application of deep learning techniques in TF prediction has attracted considerable attention from researchers and has produced promising outcomes [9,41–44]. Deep learning approaches have also proven effective in extracting latent features from TF datasets, thereby improving prediction accuracy. For example, Duan et al. introduced a deep hybrid neural network augmented with a greedy algorithm to capture spatial and temporal features [45]. Their experimental findings demonstrated the superior performance of this enhanced deep learning model in prediction accuracy and computational efficiency compared with existing methods. Similarly, Sun et al. presented the Spatio-Temporal Dilated Convolutional Network (STDGCN) for extracting spatial and temporal characteristics from traffic flow data, outperforming conventional approaches by capturing intricate nonlinear patterns [46]. Li et al. proposed an innovative uncertainty quantification framework for long-term traffic flow prediction using a sequential deep learning model [47]. Their method exhibited enhanced reliability by at least 68.1%, with the average interval width increased by 3.4%, and improved sharpness by 1.7% compared with standard models. Moreover, Zhang et al. devised a graph neural network-based traffic

forecasting model that modeled traffic dynamics as a temporal evolution process [48]. Their results indicated a 30% to 40% improvement in prediction accuracy over baseline methods. In summary, deep learning has demonstrated strong potential in capturing the nonlinear characteristics of TF data, leading to substantial progress in traffic flow prediction research.

2.2. Overview of VTF Prediction Methods

Research in TF forecasting has significantly influenced VTF prediction, as both fields seek to enhance traffic management and resource allocation through the analysis and forecasting of complex transportation patterns. While TF primarily focuses on road networks, its methodologies offer valuable insights for VTF prediction. For instance, Liang et al. introduced the Spatio-Temporal Multi-Graph Convolutional Network (STMGCN) for VTF prediction, exploiting multiple intrinsic correlations in maritime traffic graphs to achieve superior accuracy and robustness [49]. Similarly, Zhao et al. proposed the Spatio-Temporal Dynamic Graph Neural Network (STDGNN) to capture spatio-temporal dependencies in regional VTF, outperforming the GRU, STGCN, and TGCN models by 37.7%, 17.23%, and 11.4%, respectively [50]. Furthermore, Huan et al. devised a VTF prediction methodology based on EMD and LSTM, demonstrating strong predictive performance [51]. Mandalis et al. introduced a distributed Unified Approach for VTF Forecasting (dUA-VTFF) that leverages Transformer models and the Apache Spark big data processing framework [52]. Their experimental findings demonstrated improved computational efficiency compared with existing representative benchmark traffic forecasting techniques. Zhou et al. developed an LSTM prediction model optimized through Particle Swarm Optimization (PSO), significantly improving prediction accuracy relative to competing models [53]. Wang et al. proposed the DWT-Prophet hybrid model, surpassing individual Prophet, LSTM, random forest, and Support Vector Regression (SVR) models [54]. Chen et al. provide a comprehensive summary of vessel traffic flow prediction methods, covering the development from conventional statistical approaches to machine learning, deep learning, and composite forecasting models [40]. Another recent study integrates multi-source maritime data to construct a high-resolution spatio-temporal ship carbon-footprint inventory and evaluates low-carbon shipping pathways from the perspectives of operational optimization and energy-technology upgrades, highlighting the broader contribution of maritime traffic data to environmental management and green shipping [37].

Signal-decomposition methods differ in their assumptions, input structures, and parameter requirements. The standard formulations of EMD, EEMD, and CEEMDAN are mainly designed for one-dimensional signals, whereas their application to the two-dimensional VTF matrix would generally require matrix vectorization or separate row-wise or column-wise decomposition. EEMD and CEEMDAN can alleviate mode mixing but introduce additional computational cost and stochasticity [55], while VMD requires the number of modes and penalty-related parameters to be specified in advance [56]. In contrast, BEMD can directly process the two-dimensional hour-day VTF matrix and adaptively extract multiscale components without destroying its two-dimensional structure [57]. Therefore, BEMD was selected because of its structural compatibility with the adopted VTF representation rather than because it was assumed to be universally optimal.

End-to-end deep learning provides an alternative to conventional multistage prediction frameworks. Liu et al. developed an end-to-end model that directly links system-level inputs with operational decisions for bike-sharing inventory management [58]. Although this application differs from VTF forecasting, it provides useful insights into integrating traffic prediction with downstream operational decision-making. The proposed BEMD-NocalCNN-Transformer is not strictly end-to-end trainable because BEMD is implemented as an independent deterministic preprocessing stage. Future work will investigate learn-

able decomposition modules and the joint optimization of VTF forecasting and maritime-management objectives.

2.3. Contributions of Our Study

The literature review on VTF prediction reveals notable limitations in existing deep learning approaches, particularly in two critical areas requiring improvement. First, VTF prediction inherently presents a spatio-temporal challenge. Second, fluctuations and anomalies in traffic data often lead to reduced forecast accuracy. To address these issues, this study develops an integrated VTF forecasting framework by adapting and combining the established BEMD, NocalCNN, and Transformer methods. The methodological contribution lies in the task-specific integration of multi-frequency decomposition, location-specific spatial feature extraction, temporal dependency modeling, and component-level forecast reconstruction. The main contributions of this study are summarized as follows:

- (1) The original one-dimensional VTF sequence is reorganized into a two-dimensional hour–day matrix, and the established BEMD method is adapted to decompose this matrix into multiscale components. This task-specific application preserves the joint structure of intra-day and inter-day traffic variations and provides more regular representations for subsequent forecasting.
- (2) An integrated forecasting architecture is constructed by adapting NocalCNN for location-specific spatial feature extraction and Transformer for temporal dependency modeling. The predicted frequency components are subsequently reconstructed to obtain the final VTF forecast.
- (3) The proposed integration is systematically evaluated under two different waterway environments and compared with representative forecasting models under a consistent experimental protocol.

3. Methodology

This section presents the proposed integrated BEMD–NocalCNN–Transformer framework for VTF forecasting. The individual BEMD, NocalCNN, and Transformer components are established methods derived from previous studies [59,60]. The contribution of the present study lies in adapting and integrating these components into a unified workflow for two-dimensional VTF decomposition, spatial feature extraction, temporal dependency modeling, and forecast reconstruction. Section 3.1 outlines the model’s overarching framework. Subsequently, Section 3.2 delves into the specifics of the BEMD–NocalCNN–Transformer model, encompassing the BEMD technique and the training procedures of NocalCNN and Transformer. Lastly, a detailed account of the VTF prediction process is provided.

3.1. The Overarching Framework

The schematic representation of the BEMD–NocalCNN–Transformer model is illustrated in Figure 1. The architecture consists of three primary modules: BEMD, NocalCNN, and Transformer. Since VTF data exhibits nonstationary and nonlinear characteristics, to enhance predictive accuracy, the original data matrix is first decomposed using the BEMD technique to enhance prediction accuracy, generating high-frequency and low-frequency components. Local sub-matrices from both components are then extracted based on temporal nodes and fed into the NocalCNN to capture spatial features of VTF. The resulting spatial feature vectors, obtained from multiple temporal snapshots, are subsequently passed into the Transformer module to model temporal dependencies and forecast future traffic states. Finally, the predicted values are derived by aggregating the outputs from both the high-frequency and low-frequency branches.

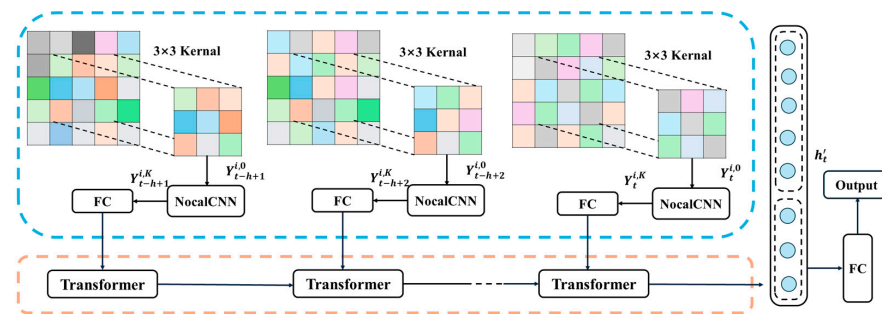


Figure 1. The VTF prediction model based on BEMD–NocalCNN–Transformer. FC represents the Fully Connected Layer.

3.2. The Proposed BEMD–NocalCNN–Transformer Model

VTF prediction presents a complex spatio-temporal challenge, influenced by diverse factors across distinct maritime channels. To address data instability, the BEMD algorithm is employed to decompose the original data into more stable frequency components. Following this, the NocalCNN–Transformer model is applied to enhance prediction accuracy by jointly capturing the spatial characteristics and temporal dependencies inherent in VTF data. Figure 2 illustrates the VTF prediction process using the BEMD–NocalCNN–Transformer framework.

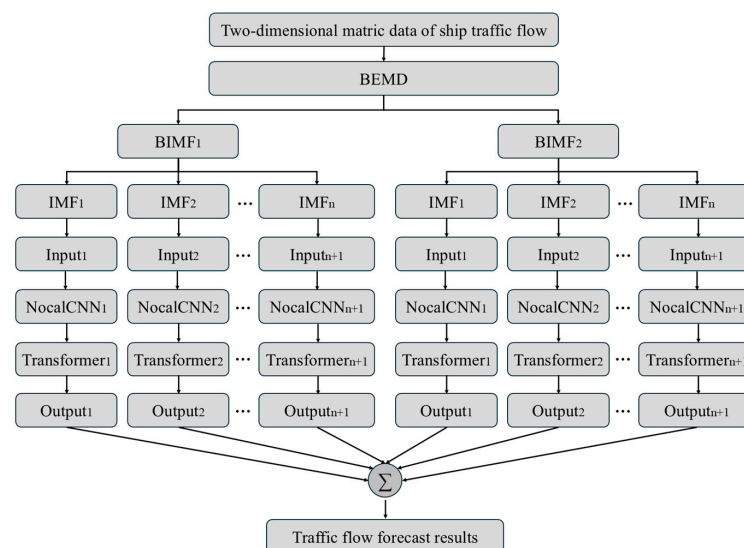


Figure 2. The VTF prediction flow based on BEMD–NocalCNN–Transformer.

BEMD, originally developed as a two-dimensional extension of empirical mode decomposition for separating spatial-frequency components, is adopted in this study to preprocess the two-dimensional VTF matrix [61]. It is adapted to decompose the non-stationary VTF matrix into multiple bidimensional intrinsic mode function components and a residual component, thereby providing more regular and interpretable representations for subsequent forecasting. To model the complex nonlinear and spatio-temporal patterns of VTF, the decomposed components are subsequently processed using the NocalCNN–Transformer architecture. NocalCNN extracts location-specific spatial features, while the self-attention mechanism of the Transformer captures temporal dependencies among consecutive traffic states. The integration of these spatial and temporal representations facilitates more effective feature learning and improves VTF prediction accuracy.

Unlike standard one-dimensional EMD, EEMD, and CEEMDAN, BEMD can directly process the two-dimensional VTF matrix without vectorizing the data or decomposing individual rows and columns separately. This property enables the model to jointly represent

intra-day and inter-day traffic variations. Compared with noise-assisted ensemble decomposition methods, BEMD also avoids repeated noise injection and ensemble averaging. In addition, unlike VMD, it does not require the number of modes to be specified in advance. These characteristics make BEMD a structurally appropriate decomposition method for the data representation adopted in this study, although they do not establish its universal optimality over all alternative decomposition methods.

VTF data exhibits non-stationary and nonlinear characteristics. To enhance prediction accuracy, the original VTF data is decomposed using the EMD technique. This process transforms the initial one-dimensional data matrix into a two-dimensional matrix (hours $\times$ days) using the EMD algorithm [62]. The matrix is then separated into high-frequency and low-frequency components. These components display varying frequencies and amplitudes across both time and space. The high-frequency component captures abrupt changes in VTF, while the low-frequency component represents regular traffic flow patterns. The specific decomposition procedure is outlined as follows:

$$e_m(x, y) = \frac{e_u(x, y) + e_l(x, y)}{2} \quad (1)$$

where $MD_{k-1}^l(x, y)$ is the original matrix; $e_u(x, y)$ and $e_l(x, y)$ represent the upper and lower parts of the envelope fitting, respectively; and $e_m(x, y)$ is the corresponding mean.

$$MD_k^l = MD_{k-1}^l(x, y) - e_m(x, y) \quad (2)$$

If MD_k^l meets the constraints, it is output as BIMF; if not, $MD_{k+1}^l(x, y) = MD_k^l(x, y)$, repeat the above calculation steps until meeting the requirements.

$$SD = \sum_{x=0}^m \sum_{y=0}^n \left(\frac{|MD_{k-1}^l(x, y) - MD_k^l(x, y)|}{MD_{k-1}^l(x, y)} \right)^2 < \tau \quad (3)$$

where m is the row; n represents the column; k is the number of calculations; $MD_k^l(x, y)$ is the l -th result; and $\tau \in [0.3, 0.5]$.

$$R_1(x, y) = MD_0^l(x, y) - B_l(x, y) \quad (4)$$

The decomposition ends when $R_1(x, y)$ has no extreme point, and the original matrix $M(x, y)$ can be expressed as the following formula:

$$M(x, y) = \sum_{q=1}^L B_q(x, y) + R_L(x, y) \quad (5)$$

where $B_q(x, y)$ is the q -th result and $R_L(x, y)$ represents the residual component.

CNNs are a class of deep feedforward neural networks widely used in visual processing tasks [63,64]. A typical CNN consists of convolutional layers, pooling layers, activation functions, and fully connected layers [65]. It can offer several advantages, including efficient learning of local data correlations, parameter sharing to reduce computational cost, and translation invariance, making it fundamental to applications such as image recognition, object detection, and image segmentation [66]. Unlike conventional CNNs, NocalCNN does not employ weight sharing, thereby enhancing its ability to capture location-specific features. In this study, NocalCNN is employed to extract spatial information from decomposed data matrices and generate target feature maps.

NocalCNN and Recurrent Neural Networks (RNNs) are designed for different data-processing tasks [67,68]. While NocalCNN is well-suited for structured grid-based data, RNNs are specialized for sequential data, making them particularly effective in applications such as natural language processing and time-series forecasting [69,70]. The typical architecture of an RNN consists of input, hidden, and output layers, enabling it to handle time-series problems with variable sequence lengths [66]. However, RNNs are prone to vanishing and exploding gradient problems, which limit their performance on long sequences [71,72]. To overcome these challenges, LSTM networks were introduced, incorporating memory cells and gating mechanisms to better capture long-term temporal dependencies [73]. In contrast, the Transformer model has demonstrated superior performance in sequence modeling tasks that require the learning of complex temporal dynamics and long-range dependencies [74–76].

NocalCNN has demonstrated considerable potential in spatial modeling and prediction tasks owing to its strong capability for feature extraction [77,78]. Unlike traditional recurrent architectures, the Transformer model uses self-attention mechanisms to effectively capture global temporal dependencies in sequence data, thereby enabling efficient parallel computation and more effective modeling of long-term trends and complex dynamic patterns in VTF [79,80]. By combining the spatial representation capabilities of NocalCNN with the temporal modeling advantages of the Transformer, this study proposes a BEMD-enhanced NocalCNN–Transformer framework for VTF prediction. The proposed model consists of six main components: BEMD decomposition, input, NocalCNN, Transformer, output, and reconstruction. Initially, the BEMD algorithm decomposes the two-dimensional TF data, transforming non-stationary signals into stationary components to facilitate more stable and accurate prediction.

- (1) BEMD decomposition layer. The BEMD algorithm is applied to decompose the two-dimensional TF matrix into several Intrinsic Mode Function (IMF) components.
- (2) Input layer. The decomposed two-dimensional IMF component matrix, $Y_t^i \in R^{3 \times 3 \times 1}$, is used as the input to the model.
- (3) NocalCNN layer. To effectively extract spatial features, the two-dimensional IMF component matrix is input into the CNN layer using a sliding-window approach with a specified step size. A 3×3 convolution operation is then performed to extract the spatial features of the VTF data.

$$Y_t^{i,K} = f\left(Y_t^{i,K-1} \otimes W_t^K + b_t^K\right) \quad (6)$$

where $\otimes$ is the convolutional operation; $f(\cdot)$ represents the activation function. After the convolution operation and tensor flattening operation, the spatial eigentensor $Y_t^{i,K} \in R^{3 \times 3 \times \lambda}$ is converted to a feature map $s_t^i \in R^{S^\lambda}$. Finally, the low-dimensional space features are obtained through the fully connected layer:

$$\hat{s}_t^i = f\left(W_t^{fc} s_t^i + b_t^{fc}\right) \quad (7)$$

- (4) Transformer layer. The Transformer architecture, originally introduced by Vaswani et al. [81], is adopted to capture temporal dependencies in the VTF sequence. In the proposed framework, the low-dimensional spatial features generated by NocalCNN from each decomposed VTF component are fed into the Transformer. The self-attention mechanism then learns temporal dependencies among consecutive traffic states and produces a spatio-temporal representation for component-level prediction. This application of the Transformer to spatial features extracted from decomposed VTF components constitutes a task-specific adaptation of the established architecture.

- (5) Output layer. The output layer generates the prediction result $\hat{y}_t^i$ for each decomposed component Y_t^i , thereby providing more accurate and stable component-level predictions.
- (6) Reconstruction layer. The predicted components are reconstructed through summation, yielding the final prediction results that integrate both temporal and spatial characteristics.
- (7) Training and overfitting control. To reduce the risk of overfitting, a finite-iteration stopping strategy was integrated into the model-training procedure. The maximum number of training iterations was determined through sensitivity analysis rather than being increased continuously. In addition, the input sequence length, learning rate, and convolutional kernel size were selected by jointly considering prediction accuracy and error stability. These procedures control training duration, input redundancy, convergence stability, and unnecessary model complexity.

4. Results and Discussion

To comprehensively evaluate the effectiveness of the BEMD–NocalCNN–Transformer model, VTF data from the CaofeiDian (CFD) and Wuhan waterways were utilized for comparative analysis. Sensitivity analyses were conducted to examine the effects of key model parameters on prediction accuracy, including the number of iterations, learning rate, input sequence length, and convolutional kernel size. Based on these analyses, the final parameter settings were determined. The predictive performance of the proposed model was then validated through comparison with alternative forecasting methods.

4.1. Dataset

In the paper, we used CFD and Wuhan water channels to assess the model's viability, as shown in Figure 3. This analysis was based on VTF data from various channels using authentic AIS data. Figure 3a displays the trajectory track density distribution in the CFD water, with ship trajectories depicted by red lines, where denser lines indicate higher traffic density. Figure 3b presents the ship track density distribution in the Wuhan water. The datasets were divided chronologically into training, validation, and test subsets using a ratio of 70%:15%:15%. The training subset was used to estimate model parameters, the validation subset was used for hyperparameter selection, and the test subset was reserved exclusively for final performance evaluation.

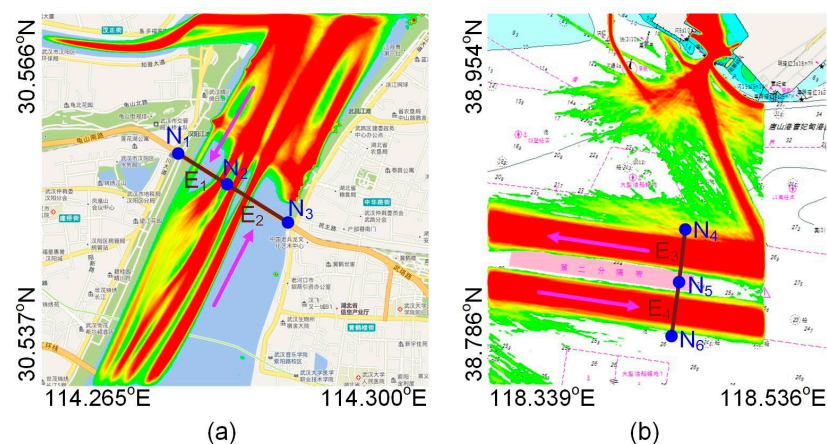


Figure 3. The visualization of vessel trajectory density in the Wuhan and CFD waterways: (a) vessel trajectory density in the Wuhan waterway, and (b) vessel trajectory density in the CFD waterway. The arrows indicate upstream and downstream vessel movements, respectively.

4.2. Evaluation Metrics

This study employed two evaluation metrics, the relative error mean (REM) and relative error standard deviation (RES), to evaluate the accuracy and stability of the VTF predictions generated by the BEMD–NocalCNN–Transformer model. Their mathematical formulations are given as follows:

$$e_b = \frac{|\tilde{y}_b - y_b|}{y_b + \varepsilon} \quad (8)$$

$$REM = \frac{1}{N} \sum_{b=1}^N e_b \quad (9)$$

$$RES = \sqrt{\frac{1}{N-1} \sum_{b=1}^N (e_b - REM)^2} \quad (10)$$

where N denotes the number of evaluated samples; y_b is the true value and $\tilde{y}_b$ represents the predicted value of the b -th sample; and ε is a small positive constant introduced to avoid division by zero; REM represents the relative error mean, and RES represents the relative error standard deviation, which reflects the dispersion degree of the relative prediction errors.

4.3. Network Parameter Settings

Because the proposed BEMD–NocalCNN–Transformer is trained using a non-convex optimization objective, a theoretically guaranteed global optimum cannot be obtained. Therefore, the final model configuration was determined empirically through systematic sensitivity analyses. Candidate settings were evaluated using REM and RES to jointly consider prediction accuracy and error stability, and the selected configuration should be interpreted as the empirically best-performing setting within the investigated ranges rather than as a globally optimal solution.

The predictive performance of the model is significantly affected by several key hyperparameters [66], including input sequence length, number of training iterations, learning rate, and convolutional kernel size [82–84]. A short input sequence may limit the model's capacity to learn temporal dependencies, whereas an excessively long sequence can introduce redundant information and increase training complexity [85]. Similarly, an excessively high learning rate or an excessive number of training iterations may lead to premature convergence and overfitting, highlighting the importance of sensitivity analysis in determining final learning and convergence parameters [86]. The convolutional kernel size directly affects the model's receptive field and its ability to extract multi-scale spatial features, making its selection critical for spatial representation [87]. In summary, careful hyperparameter tuning is essential for maximizing model performance, thereby reinforcing the need for comprehensive sensitivity analyses. In particular, the input sequence length and number of training iterations were restricted to reduce redundant input information and avoid unnecessary training after the prediction performance ceased to improve.

4.4. Prediction Results Analysis

Sensitivity analyses are conducted for key hyperparameters, including the input sequence length, number of training iterations, learning rate, and convolutional kernel size. Based on these analyses, the empirically best-performing configuration within the investigated parameter ranges is selected for the final model. This configuration helps limit redundant historical inputs, unnecessary training, unstable optimization, and excessive spatial complexity, thereby mitigating the risk of overfitting. The proposed model is subsequently compared with representative conventional machine learning, recurrent

neural network, self-attention-based, decomposition-based hybrid, and spatio-temporal benchmark models under a consistent experimental protocol.

4.4.1. Sensitivity Analysis and Empirical Optimization of Model Parameters

The term input sequence length refers to the number of time steps provided to the model as input [88]. Longer input sequences generally require greater computational resources during both training and inference, resulting in increased memory consumption and longer training time. However, they can capture richer contextual information, which may improve predictive accuracy. In contrast, shorter input sequences are computationally efficient but may fail to capture long-term dependencies and broader temporal patterns, making the model more susceptible to local fluctuations and increasing the likelihood of prediction errors. In this study, sensitivity analyses were performed using input sequence lengths of 3, 6, and 9 to evaluate their effects on model performance.

Figures 4 and 5 illustrate the model's REM values across various channels in the CFD and Wuhan waterways under varying input sequence lengths. The results indicate that input sequence length has a substantial effect on predictive accuracy, with consistent patterns observed in both regions. The model achieves the most stable performance when the input sequence length is set to 3, with REM values consistently minimized during both increasing and decreasing traffic-flow phases. These findings suggest that shorter historical input sequences can improve the model's prediction accuracy and stability.

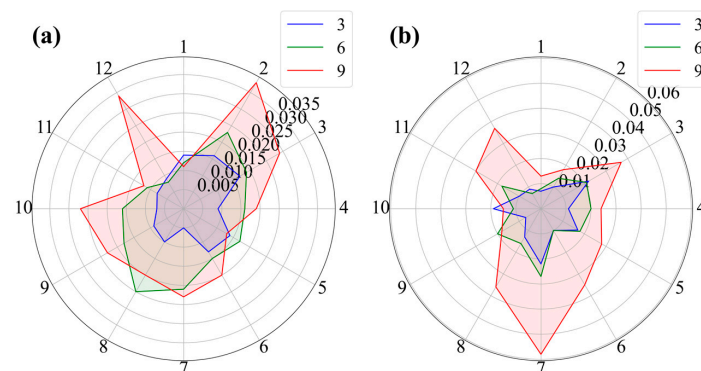


Figure 4. The REM values of VTF predictions across different channels under varying input sequence lengths in the CFD waterway. (a,b) represent the results for the upstream and downstream directions, respectively.

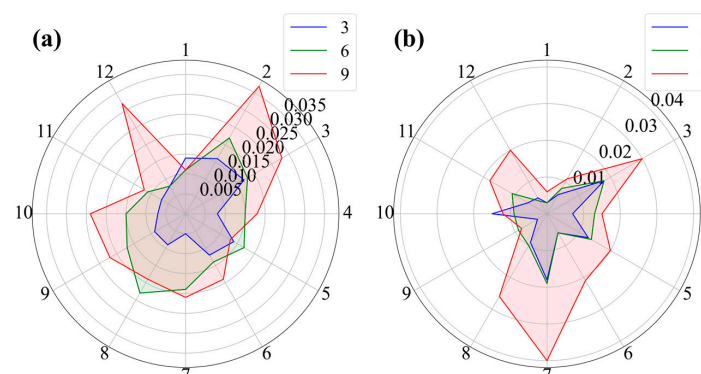


Figure 5. The REM values of VTF predictions across different channels under varying input sequence lengths in the Wuhan waterway. (a,b) represent the results for the upstream and downstream directions, respectively.

When the input sequence length is increased to 9, the model's prediction error increases markedly, with pronounced fluctuations observed at several time points, indicating possible

overfitting. This pattern is clearly shown in Figure 4a, where REM values exhibit notable spikes at the 2nd, 7th, 10th, and 12th time points. These results suggest that excessively long input sequences do not provide meaningful additional contextual information. Instead, they may introduce redundancy and noise, thereby hindering the model's ability to effectively capture short-term traffic dynamics.

Figures 6 and 7 present the model's RES values in the CFD and Wuhan waterways under varying input sequence lengths. The model achieved the most stable prediction performance in both regions when the input sequence length was set to 3. Specifically, in the CFD region, Figure 6a shows that error fluctuations were most pronounced when the sequence length was 9, particularly in the upstream direction. At several time points, namely the 2nd, 10th and 12th, RES values exceeded 0.025, indicating poor adaptability to late-night and evening traffic conditions. In contrast, when the sequence length was 3, error fluctuations remained consistently low across all time points, especially in the downstream direction, as shown in Figure 6b. Under this setting, RES values were predominantly below 0.01, reflecting minimal variation and improved predictive stability.

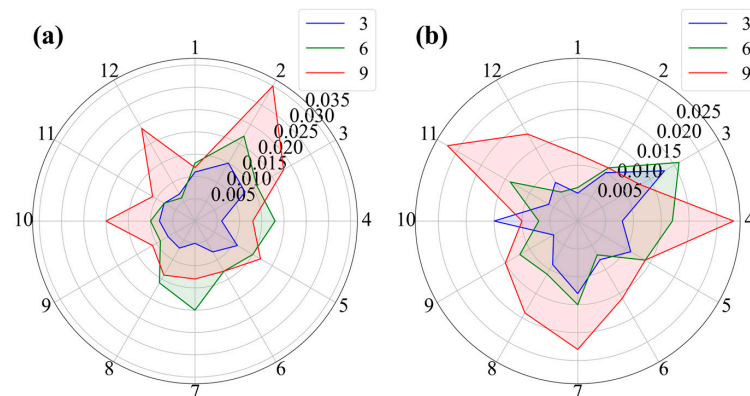


Figure 6. The RES values of VTF predictions across different channels under varying input sequence lengths in the CFD waterway. (a,b) represent the results for the upstream and downstream directions, respectively.

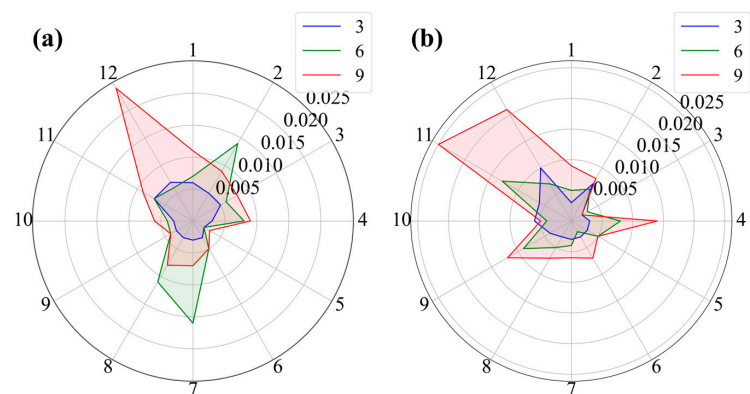


Figure 7. The RES values of VTF predictions across different channels under varying input sequence lengths in the Wuhan waterway. (a,b) represent the results for the upstream and downstream directions, respectively.

In contrast, a sequence length of 3 generally yields lower prediction errors, a narrower fluctuation range, and greater prediction stability across most time points. This advantage is observed under both the complex and variable traffic conditions of the Caofeidian waterway and the relatively regular traffic patterns of the Wuhan waterway, indicating relatively consistent performance across different waterway conditions. From the perspective of

overfitting control, the shorter input window limits redundant historical information and unnecessary input complexity, thereby reducing the risk that the model fits noise or local fluctuations rather than meaningful temporal dependencies. Although a sequence length of 6 slightly outperforms a sequence length of 3 at a few time points, its overall accuracy and stability are marginally lower. In the downstream direction, the performance curves for sequence lengths of 3 and 6 largely overlap, suggesting that the additional historical inputs provide limited useful information. These results indicate that an input sequence length of 3 is sufficient to capture the principal temporal characteristics required for short- to medium-term VTF forecasting within the datasets considered in this study.

In the Wuhan waterway, error fluctuations are noticeably lower than those observed in the CFD waterway, although some variation persists across different input sequence lengths. In the upstream direction, the model configured with a sequence length of 3 demonstrates minimal fluctuation, particularly during the critical 5th time period, where the blue curve closely aligns with the central reference line. Conversely, in the downstream direction, the model with a sequence length of 3 consistently yields the lowest RES values across most time intervals. The fluctuations are more uniformly distributed, indicating strong temporal consistency and enhanced prediction robustness.

In machine learning algorithms, the number of iterations refers to the total number of times model parameters are updated during the training or optimization process [89]. This parameter plays a critical role in training dynamics and model convergence. An insufficient number of iterations may hinder the model's ability to effectively learn from the training data, resulting in inadequate feature extraction and reduced prediction accuracy. Conversely, an excessive number of iterations may lead to overfitting, in which the model fits the training data well but generalizes poorly to unseen samples [90]. Therefore, selecting an appropriate number of iterations is essential for achieving balanced and effective model training. In this study, sensitivity analyses were conducted using a fixed input sequence length of 3, while the number of iterations varied from 100 to 900.

Figures 8 and 9 present the REM values of the model's VTF predictions across different channels under varying numbers of training iterations. The results indicate that the most stable performance, characterized by minimal prediction error, is achieved at 500 iterations. This reflects the model's ability to effectively learn temporal patterns without overfitting. In the CFD waterway, increasing the number of iterations from 100 to 500 leads to a substantial reduction in prediction error, particularly during the early-morning periods (1st–3rd time points) and nighttime periods (10th–12th time points) in the upstream direction. However, further increasing the number of iterations to 700 and 900 results in a slight increase in error during certain time intervals, suggesting performance degradation due to overtraining. A similar trend is observed in the downstream direction, although with smaller fluctuations, indicating relatively stable error behavior.

In contrast, the Wuhan waterway exhibits a faster convergence rate, as shown in Figure 9. The model achieves stable error levels after only 100 iterations, suggesting greater regularity in the data and easier model fitting. Notably, error levels remain well-controlled even beyond 500 iterations, especially during peak traffic periods in the early morning periods (3rd–4th time points) and evening rush periods (7th–8th time points) in the upstream direction, highlighting the model's robust performance in this region. Compared with the CFD setting, the Wuhan waterway shows less pronounced increases in error at 700 and 900 iterations, indicating a lower risk of overfitting and greater model stability.

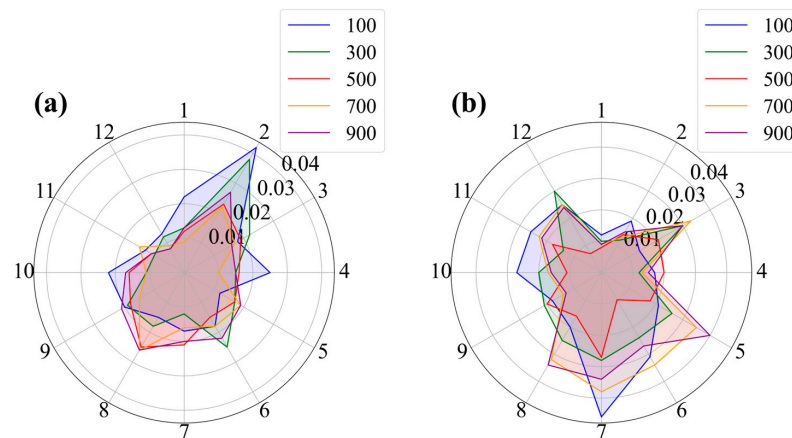


Figure 8. The REM values of VTF predictions across different channels under varying numbers of training iterations in the CFD waterway. (a,b) represent the results for the upstream and downstream directions, respectively.

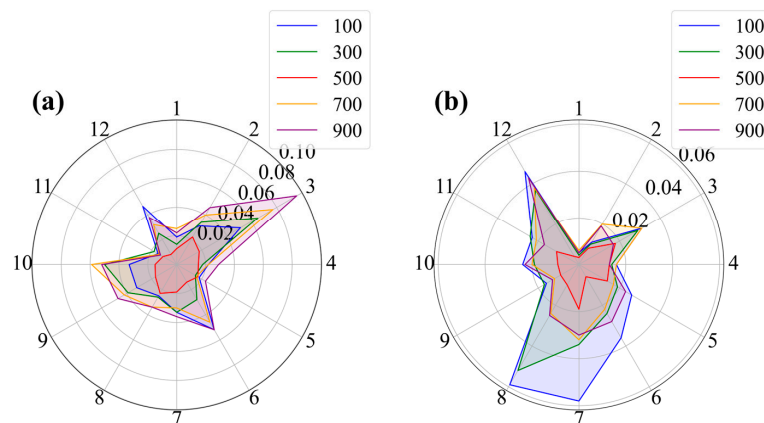


Figure 9. The REM values of VTF predictions across different channels under varying numbers of training iterations in the Wuhan waterway. (a,b) represent the results for the upstream and downstream directions, respectively.

The analysis indicates that 500 training iterations provide the best balance between prediction accuracy and model stability among the investigated settings. In the Caofeidian waterway, satisfactory model convergence requires a sufficient number of training iterations, whereas the model exhibits lower sensitivity to changes in the number of iterations in the Wuhan waterway. Further increasing the number of iterations does not consistently improve forecasting performance and leads to greater error fluctuations in the Caofeidian waterway, indicating an increased risk of overfitting. Based on these observations, the maximum number of training iterations was fixed at 500 in all subsequent experiments. This controlled training-iteration strategy avoids unnecessary training after the prediction performance ceases to improve and helps mitigate the risk of overfitting.

To further validate the experimental results, Figure 10 presents the RES values of the model's VTF forecasts under different numbers of iterations. In the CFD waterway, the model exhibits the lowest error fluctuations at 500 iterations for both upstream and downstream flows. Notably, during critical time intervals, such as early-morning and nighttime periods, the fluctuations are markedly reduced. In contrast, at only 100 iterations, the model shows comparatively larger fluctuations, indicating insufficient training and a limited ability to capture temporal traffic patterns, thereby compromising prediction stability. When the number of iterations increases to 700 or 900, RES values increase during specific periods, such as the 1st and 12th periods, suggesting mild overfitting. This effect is

reflected in the model's heightened sensitivity to anomalies or low-frequency variations in traffic flow.

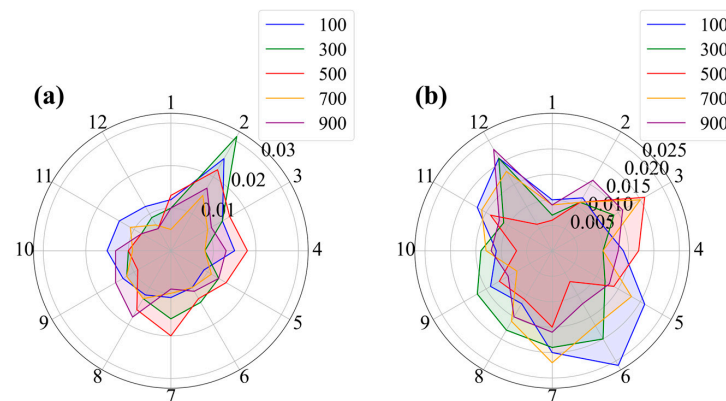


Figure 10. The RES values of VTF predictions across different channels under varying numbers of training iterations in the CFD waterway. (a,b) represent the results for the upstream and downstream directions, respectively.

Figure 11 presents the RES values of the model's VTF predictions under varying numbers of iterations in the Wuhan waterway. Compared with the CFD waterway, the model exhibits markedly lower overall volatility in Wuhan, indicating more consistent traffic-flow patterns and greater stability in model fitting. Notably, even at 500 training iterations, the model achieves consistently low RES values across all time intervals, particularly in the upstream direction. Increasing the number of iterations to 700 and 900 introduces only minimal fluctuations in error variance, highlighting the robustness of the model in the Wuhan setting. However, at specific time points, such as the 12th time point, the curve corresponding to 900 iterations shows a localized increase, suggesting that excessive training, even in relatively stable data environments, may lead to unstable predictions.

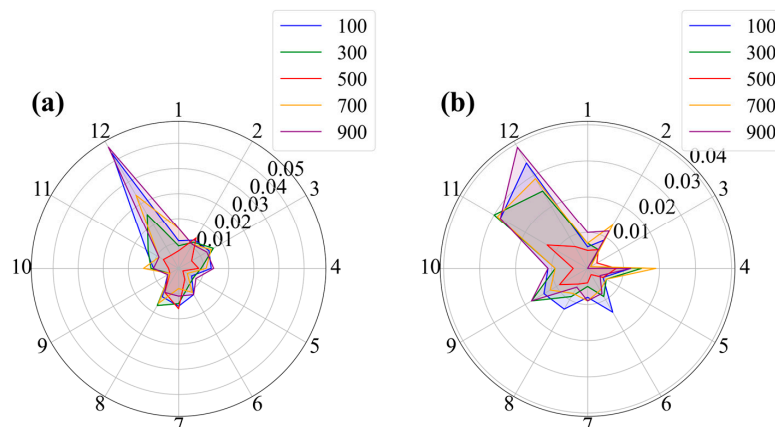


Figure 11. The RES values of VTF predictions across different channels under varying numbers of training iterations in the Wuhan waterway. (a,b) represent the results for the upstream and downstream directions, respectively.

The learning rate, a hyperparameter that controls parameter updates, plays a critical role in model convergence during training [91]. An excessively high learning rate may hinder convergence or cause training instability, whereas an overly low learning rate may prolong training and increase the risk of convergence to a suboptimal local minimum [92]. Therefore, selecting an appropriate learning rate is essential for achieving stable convergence and favorable prediction performance. In this study, sensitivity analyses

were conducted using a fixed input sequence length of 3 and 500 training iterations, with learning rates set to 0.01, 0.005, 0.001, 0.0005, and 0.0001.

Figures 12 and 13 illustrate the REM values of the model's VTF predictions in the upstream and downstream directions within the CFD and Wuhan waterways under different learning rate configurations. The model demonstrates the most stable error control at a learning rate of 0.001, with consistent performance over time, particularly during nighttime and early-morning intervals, such as the 1st, 2nd, and 3rd time points, as well as during peak hours, such as the 7th and 8th time points. At this learning rate, the model outperforms the other configurations and exhibits superior temporal generalization. In contrast, increasing the learning rate to 0.005 or 0.01 markedly increases the prediction error, especially in the upstream direction of the CFD waterway, as shown in Figure 12a. At a learning rate of 0.01, the error approaches 0.1 at the first time point, indicating that an excessively high learning rate may cause model oscillation and training instability. Conversely, a very low learning rate, such as 0.0001, results in relatively high overall error levels, suggesting slow convergence and possible underfitting.

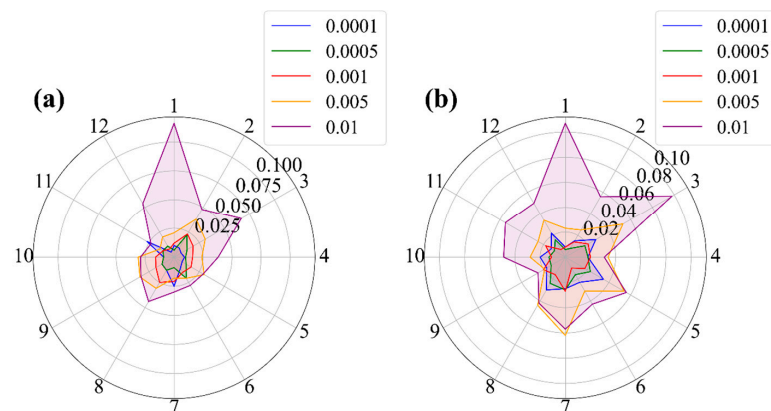


Figure 12. The REM values of VTF predictions across different channels under varying learning rates in the CFD waterway. (a,b) represent the results for the upstream and downstream directions, respectively.

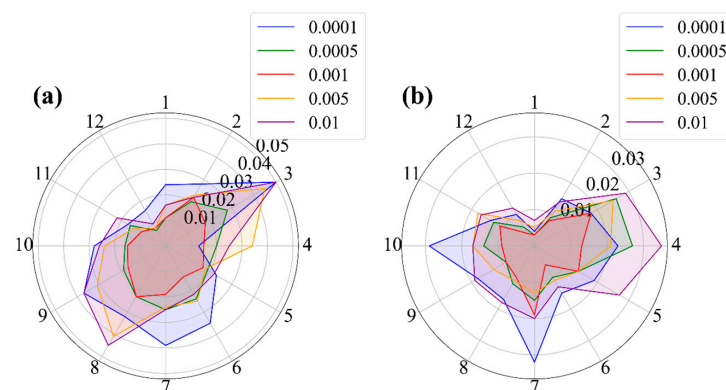


Figure 13. The REM values of VTF predictions across different channels under varying learning rates in the Wuhan waterway. (a,b) represent the results for the upstream and downstream directions, respectively.

To investigate the impact of the learning rate on prediction stability, Figures 14 and 15 illustrate the RES values of the model's VTF predictions in the upstream and downstream directions under varying learning rates in the CFD and Wuhan waterways. RES reflects the variability of the model's prediction errors, with lower values indicating greater stability and robustness. A comprehensive analysis of both scenarios indicates that a learning rate of 0.001 provides the best model performance. This learning rate produces minimal error

fluctuations under both increasing and decreasing traffic-flow trends, thereby supporting stable training and robust adaptability to temporal variations in traffic flow. Notably, during periods of irregular traffic patterns, such as early-morning and evening periods, the RES values obtained at a learning rate of 0.001 are markedly lower than those obtained under alternative configurations, highlighting its consistent performance across two different waterway conditions across varying time intervals. In contrast, extremely small learning rates, such as 0.0001, lead to erratic fluctuations during specific periods, indicating slow convergence and insufficient learning. Conversely, excessively large learning rates, such as 0.01, result in greater error variability, with elevated RES values particularly evident in the initial phase of upstream traffic and between the 11th and 12th time points of downstream traffic. These RES spikes indicate oscillatory behavior and instability during model training.

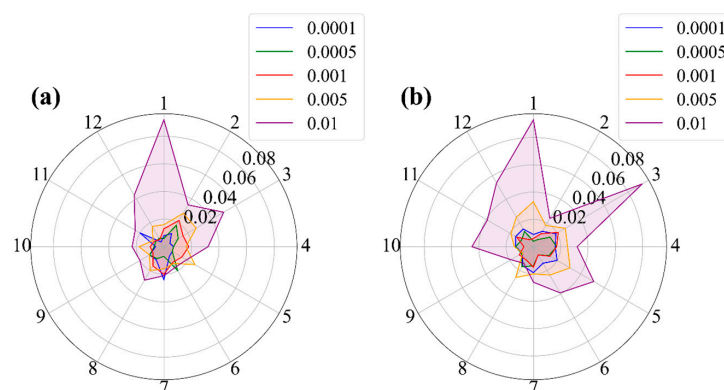


Figure 14. The RES values of VTF predictions across different channels under varying learning rates in the CFD waterway. (a,b) represent the results for the upstream and downstream directions, respectively.

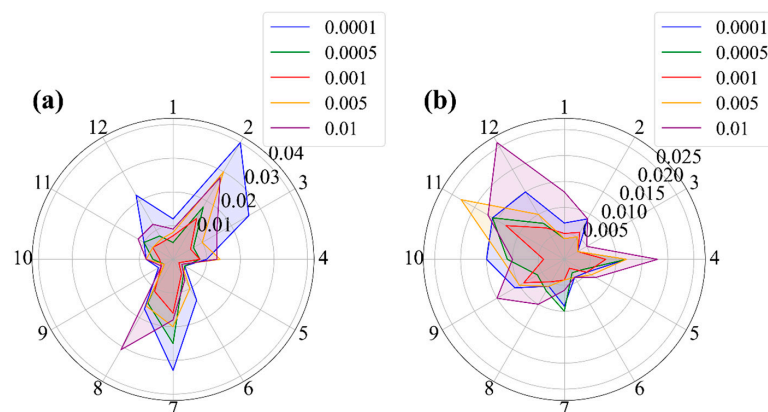


Figure 15. The RES values of VTF predictions across different channels under varying learning rates in the Wuhan waterway. (a,b) represent the results for the upstream and downstream directions, respectively.

Despite differences in data structure and traffic patterns between the two waterways, the overall trend of the RES curve remains remarkably consistent. This consistency highlights the strong cross-regional uniformity in the influence of learning rate on model stability. Specifically, a learning rate of 0.001 consistently reduces error fluctuations across both regions and flow directions and achieves the most stable performance among the tested values within the current model framework.

The convolutional kernel size is a key hyperparameter because it directly determines the receptive field and affects the model's feature extraction capability [93]. Smaller kernels, such as 3×3 , are more effective in capturing local features, whereas larger kernels, such as 5×5 or 7×7 , are better suited for extracting broader spatial patterns [94]. Therefore,

selecting an appropriate kernel size enables the model to capture traffic flow features at different spatial scales more effectively. In this study, convolutional kernel sizes of 3×3 , 5×5 , and 7×7 were evaluated for comparative analysis. Figures 16 and 17 present the VTF prediction errors of the model under different kernel size settings.

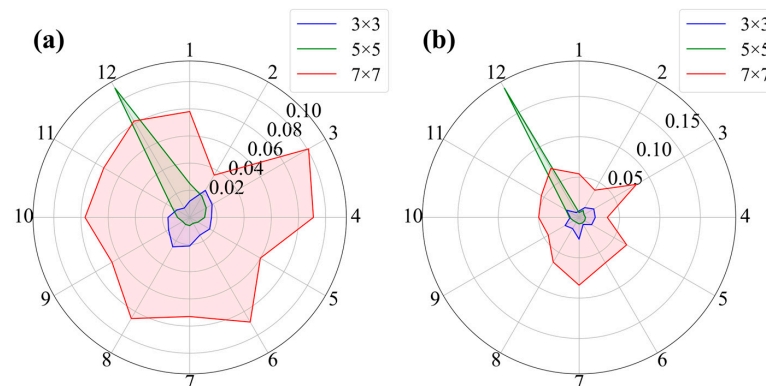


Figure 16. The REM values of VTF predictions across different channels under varying convolution kernel sizes in the CFD waterway. (a,b) represent the results for the upstream and downstream directions, respectively.

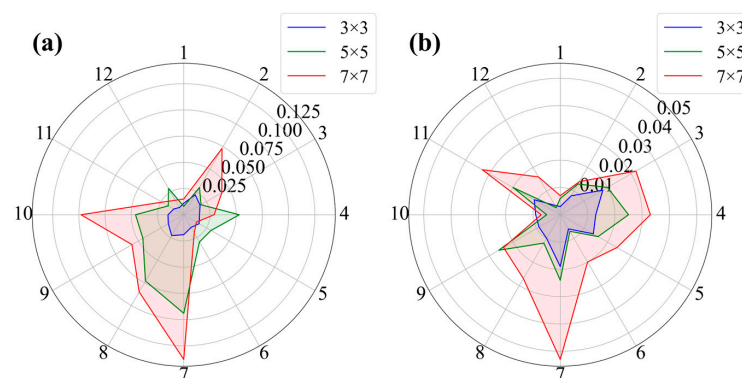


Figure 17. The REM values of VTF predictions across different channels under varying convolution kernel sizes in the Wuhan waterway. (a,b) represent the results for the upstream and downstream directions, respectively.

The model achieves the lowest REM values across both study areas and flow directions when using a 3×3 convolutional kernel. In the CFD waterway, particularly in the upstream direction, as shown in Figure 16a, the overall error obtained with the 3×3 kernel remains below 0.02, outperforming the 5×5 and 7×7 kernels. By contrast, the REM increases substantially when the 7×7 kernel is used, reaching 0.08–0.10 across several time intervals. This indicates that an excessively large receptive field may weaken the model's ability to capture local TF features, thereby reducing prediction accuracy.

A similar pattern is observed in the Wuhan waterway. The 3×3 kernel consistently produces the lowest REM values across most time intervals, regardless of flow direction, suggesting stronger sensitivity to and better fitting capacity for temporal traffic flow variations. In addition, the overall REM values in Wuhan are generally lower than those in the CFD waterway, reflecting lower data volatility and more regular traffic patterns. In the downstream direction, as shown in Figure 17b, the REM values mostly remain below 0.02. However, the 7×7 kernel also results in a clear increase in prediction error in the Wuhan waterway, further confirming that an overly large receptive field can impair model performance across different regions.

Figures 18 and 19 illustrate the RES values of the model's VTF predictions under varying convolution kernel sizes. The forecasting results show slight variations across different channels and time intervals. Among the tested configurations, the 3×3 convolution kernel demonstrates the most stable performance across both waterways and navigation directions. In the CFD waterway, the RES values associated with the 7×7 kernel increase markedly at most time points, particularly in the upstream direction, where fluctuations exceed 0.05 in several instances. This suggests that an excessively large receptive field may introduce redundant or noisy information, thereby reducing model stability. In contrast, the RES values for the 3×3 kernel consistently range from 0.01 to 0.02 across most time intervals, indicating strong temporal consistency and prediction robustness.

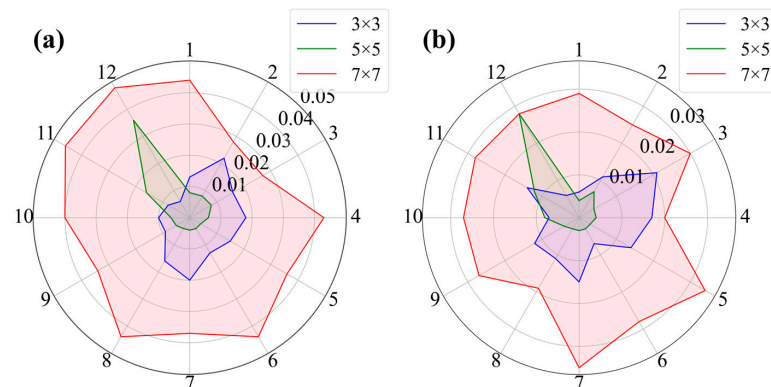


Figure 18. The RES values of VTF predictions across different channels under varying convolution kernel sizes in the CFD waterway. (a,b) represent the results for the upstream and downstream directions, respectively.

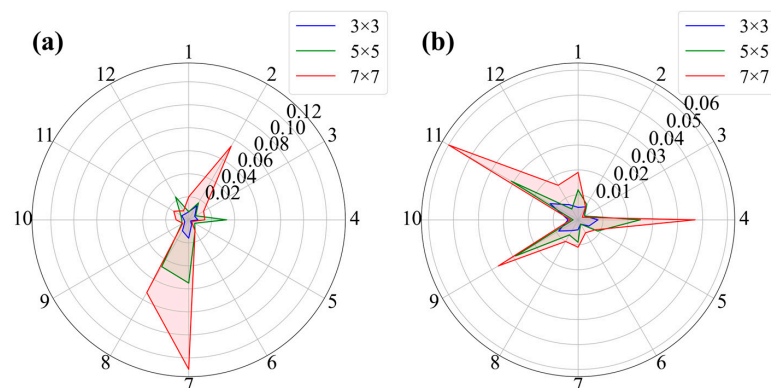


Figure 19. The RES values of VTF predictions across different channels under varying convolution kernel sizes in the Wuhan waterway. (a,b) represent the results for the upstream and downstream directions, respectively.

In the Wuhan waterway, although the overall fluctuations are relatively small, differences among kernel sizes remain evident. The 3×3 kernel maintains low and centrally distributed RES values in both directions, particularly during critical time intervals, such as the 4th, 7th, and 10th periods, as shown in Figure 19. Conversely, the 7×7 kernel exhibits more pronounced fluctuations during these intervals, further highlighting its lower stability compared with the smaller receptive-field configuration.

The sensitivity analyses demonstrate that model performance varies systematically with the selected hyperparameters. Within the investigated ranges, the combination of an input sequence length of 3, 500 training iterations, a learning rate of 0.001, and a 3×3 convolutional kernel achieved the most favorable balance between prediction accuracy and stability. The input sequence length and convolutional kernel size are related to

the temporal and local structural characteristics of VTF. Specifically, a sequence length of 3 captures the continuity of recent traffic states while avoiding excessive historical redundancy, whereas the 3×3 kernel covers neighboring elements along the intra-day and inter-day dimensions of the VTF matrix, enabling the model to extract local traffic variations. In contrast, the learning rate and number of training iterations are numerical optimization parameters rather than direct representations of physical traffic behavior. The learning rate of 0.001 provided stable convergence, while 500 iterations were sufficient for model convergence without unnecessary prolonged training. Because the optimization objective of the deep learning model is non-convex, these results indicate the empirically best-performing configuration within the examined search space rather than theoretical global optimality.

4.4.2. Comparison with Representative Previous Forecasting Methods

To evaluate the proposed framework against representative methods used in previous traffic-flow and VTF forecasting studies, Support Vector Machine (SVM), Backpropagation Neural Network (BPNN), RNN, LSTM, Gated Recurrent Unit (GRU), Transformer, and NocalCNN–Transformer were implemented as benchmark models. To strengthen the comparison with representative previous vessel traffic flow forecasting studies, WVMA–LSTM was included as an additional advanced benchmark [95]. This model integrates the Whale Optimization Algorithm (WOA), Variational Mode Decomposition (VMD), and a Multi-Head Attention Mechanism. It was selected because it represents a recent decomposition-based hybrid framework specifically developed for ship traffic flow prediction under complex and high-density traffic conditions. The selected benchmark models provide controlled reference points for evaluating conventional nonlinear prediction, recurrent temporal modeling, self-attention-based forecasting, and spatio-temporal feature learning. They are not intended to represent all recently developed decomposition-based hybrid networks. All models were trained and evaluated using the same CFD and Wuhan datasets, data preprocessing procedures, training–testing partitions, input sequence settings, prediction targets, and evaluation metrics. Therefore, the differences reported in this section primarily reflect the feature representation and prediction capabilities of the models rather than differences in the datasets or evaluation procedures. Model-specific parameters were determined using the same validation principle where applicable.

Among the selected benchmarks, NocalCNN–Transformer provides the most direct ablation comparison with the proposed framework because it retains the same spatial and temporal prediction architecture but excludes BEMD preprocessing. Therefore, the comparison between NocalCNN–Transformer and BEMD–NocalCNN–Transformer directly evaluates the incremental contribution of the decomposition stage.

Figures 20 and 21 illustrate the VTF prediction errors across different model configurations. The BEMD–NocalCNN–Transformer achieves the lowest REM values, thereby highlighting its excellent stability and accuracy. In the CFD waterway, the proposed model consistently outperforms all other approaches across all time intervals, as shown in Figure 20. Notably, at the 1st, 3rd, 8th, and 12th time points, its prediction errors are substantially lower than those of traditional neural networks such as RNN and GRU, and classical machine learning models, such as SVM and BPNN. Furthermore, it outperforms the NocalCNN–Transformer, demonstrating the added benefit of incorporating the BEMD component. In contrast, models such as SVM and BPNN exhibit considerable error variability over time, with pronounced spikes at the 1st and 8th time points, indicating limited adaptability to abrupt fluctuations in traffic flow.

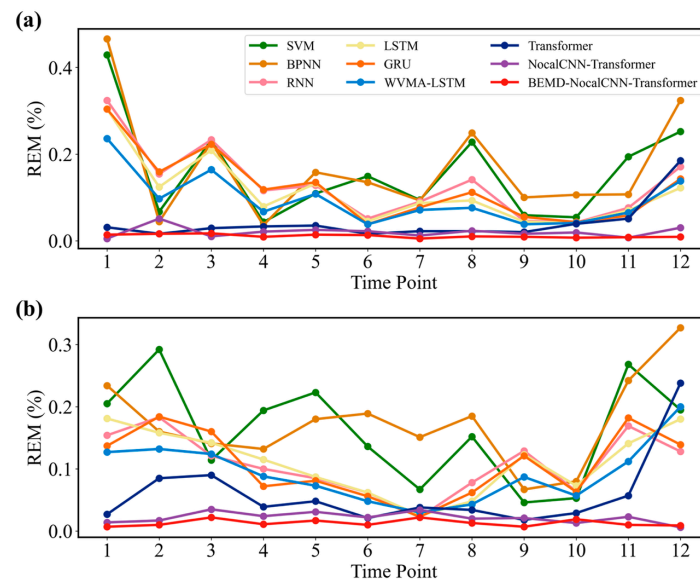


Figure 20. The REM values of different methods for VTF forecasting in the CFD waterway. (a,b) represent the results for the upstream and downstream directions, respectively.

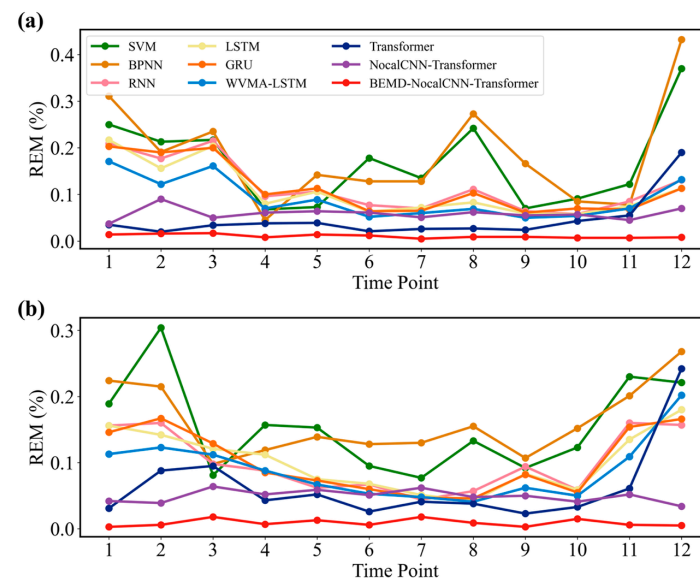


Figure 21. The REM values of different methods for VTF forecasting in the Wuhan waterway. (a,b) represent the results for the upstream and downstream directions, respectively.

Among the benchmark models, NocalCNN–Transformer provides the most direct comparison with the proposed framework because it retains the same spatial and temporal prediction architecture but excludes BEMD preprocessing. Therefore, the performance difference between the two models can be mainly attributed to the decomposition stage. The consistently lower REM and RES values of BEMD–NocalCNN–Transformer indicate that separating the original non-stationary traffic-flow matrix into frequency components facilitates the learning of both regular traffic patterns and abrupt fluctuations.

In the Wuhan waterway, the BEMD–NocalCNN–Transformer demonstrates superior performance in both upstream and downstream flow directions, consistently maintaining the lowest REM values across all time intervals, as shown in Figure 21. These results highlight the effectiveness of the model in capturing temporal variations in traffic-flow dynamics. In contrast, conventional models exhibit increased errors during nighttime or peak traffic periods, particularly at the 2nd, 11th, and 12th time points, with models

such as SVM and the standard Transformer showing pronounced instability. Comparative analysis confirms that the BEMD–NocalCNN–Transformer outperforms both traditional machine learning and deep learning approaches in terms of prediction accuracy and cross-regional generalization. Notably, even the NocalCNN–Transformer also outperforms most baseline models, highlighting the advantages of combining local convolutional structures with Transformer-based architectures. The integration of the BEMD component further enhances the model's capacity to capture nonlinear and multi-scale temporal features, thereby enabling high-precision VTF forecasting.

Compared with the conventional SVM, BPNN, RNN, LSTM, and GRU models, WVMA-LSTM generally achieves lower prediction errors, confirming the effectiveness of combining signal decomposition, attention-based feature weighting, and recurrent temporal modelling. This result also demonstrates that WVMA-LSTM provides a more competitive benchmark than conventional single-structure forecasting methods. Nevertheless, the proposed BEMD–NocalCNN–Transformer achieves lower overall REM and RES values than WVMA-LSTM. At most time points, the proposed model produces the lowest or near-lowest prediction errors and exhibits smaller fluctuations across the forecasting period. These results indicate that the proposed framework provides more stable performance under different traffic directions and varying traffic-flow conditions.

The comparison between NocalCNN–Transformer and BEMD–NocalCNN–Transformer provides further insight into the mechanisms underlying the observed performance improvement. The original VTF matrix contains relatively smooth traffic trends, periodic variations, and abrupt short-term fluctuations at different frequency scales. When these patterns are modeled jointly, dominant low-frequency variations may obscure weaker high-frequency changes, resulting in delayed responses to sudden fluctuations or excessive smoothing of local peaks and valleys. BEMD alleviates this problem by decomposing the non-stationary VTF matrix into multiple BIMF components and a residual component, thereby reducing cross-scale interference and enabling more effective component-level forecasting. NocalCNN further mitigates prediction errors associated with the spatial weight-sharing assumption of conventional CNNs. Because different positions in the two-dimensional hour–day VTF matrix correspond to different temporal and operational contexts, similar local patterns occurring at different positions may have different predictive meanings. By learning location-specific weights, NocalCNN preserves local heterogeneity and improves the representation of localized peaks, valleys, and irregular traffic states. Therefore, BEMD reduces interference among multiscale traffic variations before feature extraction, whereas NocalCNN reduces location-dependent representation bias during spatial feature learning, and their complementary roles contribute to the improved prediction accuracy and stability of the proposed framework.

In conclusion, the BEMD–NocalCNN–Transformer demonstrates superior predictive performance, achieving the lowest mean relative error in both upstream and downstream directions across the CFD and Wuhan waterways. Among all compared models, it is identified as the most effective approach, exhibiting strong spatio-temporal generalization capability and substantial potential for practical deployment.

To evaluate the prediction stability of each model, Figures 22 and 23 present a comparative analysis of RES values across different models in both upstream and downstream directions within the CFD and Wuhan waterways. The BEMD–NocalCNN–Transformer consistently exhibits minimal error fluctuations across both navigation directions and regions, indicating a high degree of prediction stability. In the CFD waterway, RES values remain consistently below 0.01% for both flow directions, significantly outperforming conventional models such as GRU, LSTM, and RNN. The superior stability of the proposed model is particularly evident at the 2nd, 3rd, and 12th time points. Similarly, in the Wuhan

waterway, the BEMD–NocalCNN–Transformer maintains low volatility across all time points, further highlighting its robust, consistent performance under varying spatial and temporal conditions.

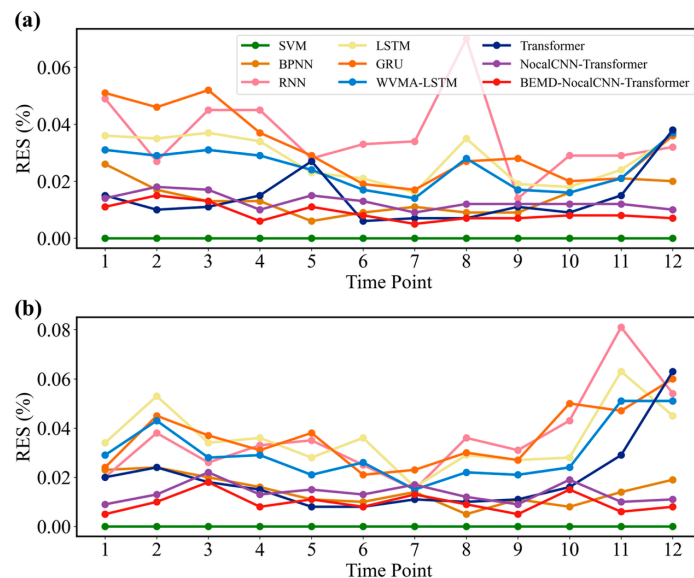


Figure 22. The RES results of various methods for VTF forecasting in the CFD waterway. (a,b) represent the results for the upstream and downstream directions, respectively.

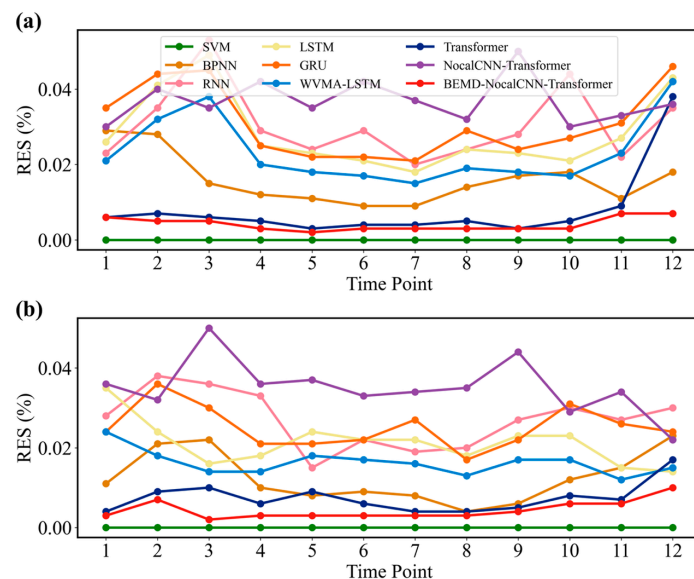


Figure 23. The RES results of various methods for VTF forecasting in the Wuhan waterway. (a,b) represent the results for the upstream and downstream directions, respectively.

In contrast, traditional neural network models such as GRU and RNN exhibit pronounced error fluctuations at several time intervals, particularly at the 2nd and 8th time points, where RES values approach or exceed 0.04, as shown in Figure 22a. These models demonstrate limited adaptability to abnormal traffic conditions and sudden variations. Although the NocalCNN–Transformer performs well at selected time points, its overall volatility remains substantially higher than that of the BEMD-enhanced version, underscoring the critical importance of multi-scale temporal feature extraction in enhancing model stability, as shown in Figure 23b. Although the Transformer model achieves strong predictive accuracy, it still exhibits considerable volatility, which limits its robustness under varying traffic conditions. By comparison, BPNN and LSTM show moderate performance:

they remain relatively stable during periods of low traffic variability but exhibit heightened fluctuations during intervals characterized by abrupt changes or high variability.

A similar performance pattern is observed in the Caofeidian waterway. Although WVMA-LSTM generally outperforms the conventional SVM, BPNN, RNN, LSTM, and GRU models, the proposed BEMD–NocalCNN–Transformer achieves lower overall REM and RES values in both traffic directions. In particular, the proposed model obtains the lowest or near-lowest prediction errors at most time points and exhibits relatively small error fluctuations, even during periods of pronounced traffic-flow variation. The consistent performance improvements observed in both the Caofeidian and Wuhan waterways indicate that the advantages of the proposed model are not confined to a single study area. Instead, the model demonstrates good robustness and cross-regional consistency under waterway environments characterized by different traffic patterns and operational conditions.

Although the Transformer also achieves relatively high prediction accuracy, its error curves show noticeable fluctuations, indicating reduced stability under rapidly changing traffic conditions. BPNN and LSTM exhibit moderate overall performance and remain relatively stable during periods of limited traffic variation; however, their prediction errors increase substantially when the traffic flow changes abruptly. In comparison, the proposed BEMD–NocalCNN–Transformer maintains relatively low and stable REM and RES values under both regular and highly variable traffic conditions. This result suggests that the integration of BEMD-based multi-frequency decomposition, NocalCNN-based local feature extraction, and Transformer-based temporal modelling improves the model's ability to capture complex and non-stationary vessel traffic patterns.

Overall, using the same datasets and evaluation procedures, the proposed BEMD–NocalCNN–Transformer achieved better prediction accuracy and stability than the selected representative benchmark models. Its consistent performance in the CFD coastal-port waterway and the Wuhan inland waterway demonstrates the effectiveness of combining signal decomposition with spatio-temporal learning under two distinct traffic environments. However, because the model was trained and evaluated only on these two regional datasets, the current results do not establish its global applicability.

4.4.3. Robustness to Variations in Input Conditions

In addition to the sensitivity analysis of model hyperparameters, the robustness of the proposed model was examined under different input-side problem conditions while maintaining the same model configuration and evaluation protocol. Specifically, the CFD and Wuhan datasets represent two distinct waterway environments with different traffic patterns and operational characteristics. Model performance was also evaluated separately for upstream and downstream vessel traffic, as well as across different temporal periods.

The results show that the proposed model maintains relatively stable prediction performance across both waterways and navigation directions. The Wuhan waterway generally exhibits more regular traffic patterns and smaller error fluctuations, whereas the more complex and variable traffic conditions in the CFD waterway result in greater prediction variability. Nevertheless, the proposed model consistently maintains comparatively low REM and RES values under both conditions. Similarly, the model performs reliably for both upstream and downstream traffic, indicating that its prediction capability is not restricted to a single navigation direction.

Temporal variations also affect prediction performance. Error fluctuations are generally more pronounced during nighttime periods than during daytime periods, which may be associated with greater variability in vessel operations and traffic organization at night. These observations indicate that the proposed model is reasonably robust to variations in

waterway environment, navigation direction, and temporal traffic conditions, although irregular nighttime traffic remains a more challenging input scenario.

It should be noted that the present robustness analysis is based on naturally occurring variations in the available datasets. Controlled perturbation experiments involving artificial AIS noise, missing observations, and changes in sampling frequency were not conducted. These input-quality variations will be systematically investigated in future work.

5. Conclusions

Accurate VTF forecasting is critical for optimizing route planning, reducing congestion and collision risks, and enhancing the efficiency and safety of maritime operations. It also facilitates the optimal allocation of port facilities and resources, thereby improving overall operational performance and service quality. To improve VTF prediction accuracy, this study proposes a novel deep learning model, named the BEMD–NocalCNN–Transformer. The proposed approach first converts the original one-dimensional traffic data into a matrix format. The BEMD algorithm is then applied to decompose the matrix into high- and low-frequency components. These component matrices are fed into NocalCNN to extract spatial features, which are subsequently passed to a Transformer network to capture temporal dependencies and predict future traffic values. The final prediction is obtained by aggregating the outputs of the high- and low-frequency branches, enabling the model to effectively integrate multi-scale features.

A series of comparative experiments were conducted to evaluate the effects of key model parameters, including input sequence length, number of training iterations, learning rate, convolution kernel size, and similarity matrix weights. The final parameter settings were selected based on the lowest prediction errors within the investigated ranges. Specifically, input sequence lengths of 3, 6, 9, 12, and 15 were tested, along with training iteration counts of 100, 300, 500, 700, and 900. Learning rates of 0.01, 0.005, 0.001, 0.0005, and 0.0001 were evaluated, and convolution kernel sizes of 3×3 , 5×5 , and 7×7 were compared. The results indicate that the optimal configuration consists of an input sequence length of 3, 500 training iterations, a learning rate of 0.001, and a 3×3 convolution kernel. This configuration consistently yielded the lowest prediction error. To further validate the effectiveness of the model, the BEMD–NocalCNN–Transformer was compared with several baseline models, including SVM, BPNN, RNN, LSTM, GRU, Transformer, and NocalCNN–Transformer, under consistent parameter settings. These results demonstrate that the proposed framework provides improved prediction accuracy and stability compared with the selected benchmark methods under the experimental settings considered in this study.

Despite the encouraging results, this study has several limitations. First, the evaluation is limited to the CFD coastal-port waterway and the Wuhan inland waterway and therefore does not establish the global applicability of the proposed model. Future work will use multi-region AIS datasets and cross-region transfer experiments to assess its spatial transferability under different traffic densities, vessel compositions, and waterway conditions. Second, the present study assumes a centralized data setting. A federated-learning extension will be investigated to enable local BEMD decomposition and model training while exchanging only protected model updates among participating organizations. Third, although the selected benchmarks provide controlled comparisons across representative forecasting paradigms, they do not include independently developed modern decomposition-based hybrid networks. Future work will conduct systematic comparisons with EMD-, EEMD-, CEEMDAN-, and VMD-based hybrid models under the same datasets and evaluation protocol.

Author Contributions: C.Z.: data acquisition; data analysis; data interpretation; study design; writing—original draft; Q.C. and Z.Y.: data analysis; data interpretation; B.Y.C.: writing—review and editing. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

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.

References

1. Krugman, P.; Cooper, R.N.; Srinivasan, T.N. Growing World Trade: Causes and Consequences. *Brook. Pap. Econ. Act.* **1995**, *1995*, 327–377. [\[CrossRef\]](#)
2. Behar, A.; Venables, A.J. Transport Costs and International Trade. In *A Handbook of Transport Economics*; Edward Elgar Publishing: Cheltenham, UK, 2011.
3. Wessel, J. Evaluating the Transport-Mode-Specific Trade Effects of Different Transport Infrastructure Types. *Transp. Policy* **2019**, *78*, 42–57. [\[CrossRef\]](#)
4. Panayides, P.M.; Song, D.-W. Maritime Logistics as an Emerging Discipline. *Marit. Policy Manag.* **2013**, *40*, 295–308. [\[CrossRef\]](#)
5. Del Mondo, G.; Peng, P.; Gensel, J.; Claramunt, C.; Lu, F. Leveraging Spatio-Temporal Graphs and Knowledge Graphs: Perspectives in the Field of Maritime Transportation. *ISPRS Int. J. Geo-Inf.* **2021**, *10*, 541. [\[CrossRef\]](#)
6. Christiansen, M.; Fagerholt, K.; Nygreen, B.; Ronen, D. Chapter 4 Maritime Transportation. In *Handbooks in Operations Research and Management Science*; Barnhart, C., Laporte, G., Eds.; Transportation; Elsevier: Amsterdam, The Netherlands, 2007; Volume 14, pp. 189–284.
7. Cheng, Z.; Chen, P.; Mou, J.; Chen, L. Novel Collision Risk Measurement Method for Multi-Ship Encounters via Velocity Obstacles and Temporal Proximity. *Ocean Eng.* **2024**, *302*, 117585. [\[CrossRef\]](#)
8. Graziano, A.; Teixeira, A.P.; Guedes Soares, C. Classification of Human Errors in Grounding and Collision Accidents Using the TRACER Taxonomy. *Saf. Sci.* **2016**, *86*, 245–257. [\[CrossRef\]](#)
9. Li, Y.; Liang, M.; Li, H.; Yang, Z.; Du, L.; Chen, Z. Deep Learning-Powered Vessel Traffic Flow Prediction with Spatial-Temporal Attributes and Similarity Grouping. *Eng. Appl. Artif. Intell.* **2023**, *126*, 107012. [\[CrossRef\]](#)
10. Xiao, Y.; Chen, Y.; Liu, X.; Yan, Z.; Cheng, L.; Li, M. Oil Flow Analysis in the Maritime Silk Road Region Using AIS Data. *ISPRS Int. J. Geo-Inf.* **2020**, *9*, 265. [\[CrossRef\]](#)
11. Nie, Y.; Liu, K.; Xin, X.; Yu, Q. A Novel Through Capacity Model for One-Way Channel Based on Characteristics of the Vessel Traffic Flow. *TransNav Int. J. Mar. Navig. Saf. Sea Transp.* **2017**, *11*, 495–502. [\[CrossRef\]](#)
12. Sang, L.; Wall, A.; Mao, Z.; Yan, X.; Wang, J. A Novel Method for Restoring the Trajectory of the Inland Waterway Ship by Using AIS Data. *Ocean Eng.* **2015**, *110*, 183–194. [\[CrossRef\]](#)
13. Lv, T.; He, C.; Zhang, J.; Song, Z. Massive AIS Data Storage and Query Based on Hadoop Platform. *J. Phys. Conf. Ser.* **2021**, *1948*, 012016. [\[CrossRef\]](#)
14. Schwing, F.B. Modern Technologies and Integrated Observing Systems Are “Instrumental” to Fisheries Oceanography: A Brief History of Ocean Data Collection. *Fish. Oceanogr.* **2023**, *32*, 28–69. [\[CrossRef\]](#)
15. Le Tixerant, M.; Le Guyader, D.; Gourmelon, F.; Queffelec, B. How Can Automatic Identification System (AIS) Data Be Used for Maritime Spatial Planning? *Ocean Coast. Manag.* **2018**, *166*, 18–30. [\[CrossRef\]](#)
16. Bai, X.; Cheng, L.; Iris, Ç. Data-Driven Financial and Operational Risk Management: Empirical Evidence from the Global Tramp Shipping Industry. *Transp. Res. Part E Logist. Transp. Rev.* **2022**, *158*, 102617. [\[CrossRef\]](#)
17. Wu, M.; Chen, L. Image Recognition Based on Deep Learning. In Proceedings of the 2015 Chinese Automation Congress (CAC), Wuhan, China, 27–29 November 2015; pp. 542–546.
18. Chen, C.; Li, O.; Tao, D.; Barnett, A.; Rudin, C.; Su, J.K. This Looks Like That: Deep Learning for Interpretable Image Recognition. In *Proceedings of the Advances in Neural Information Processing Systems*; Curran Associates, Inc.: Red Hook, NY, USA, 2019; Volume 32.
19. Li, Y. Research and Application of Deep Learning in Image Recognition. In Proceedings of the 2022 IEEE 2nd International Conference on Power, Electronics and Computer Applications (ICPECA), Shenyang, China, 21–23 January 2022; pp. 994–999.
20. Wang, X.; Liu, G.; He, J.; Zhou, X.; Luo, Z. A Large Language Model for Traffic Flow Prediction Based on Stationary Wavelet Transform and Graph Convolutional Networks. *ISPRS Int. J. Geo-Inf.* **2026**, *15*, 166. [\[CrossRef\]](#)
21. Garcia-Garcia, A.; Orts-Escolano, S.; Oprea, S.; Villena-Martinez, V.; Garcia-Rodriguez, J. A Review on Deep Learning Techniques Applied to Semantic Segmentation. *arXiv* **2017**, arXiv:1704.06857.

22. Lateef, F.; Ruichek, Y. Survey on Semantic Segmentation Using Deep Learning Techniques. *Neurocomputing* **2019**, *338*, 321–348. [\[CrossRef\]](#)
23. Hao, S.; Zhou, Y.; Guo, Y. A Brief Survey on Semantic Segmentation with Deep Learning. *Neurocomputing* **2020**, *406*, 302–321. [\[CrossRef\]](#)
24. Hua, Y.; Zhao, Z.; Li, R.; Chen, X.; Liu, Z.; Zhang, H. Deep Learning with Long Short-Term Memory for Time Series Prediction. *IEEE Commun. Mag.* **2019**, *57*, 114–119. [\[CrossRef\]](#)
25. Han, Z.; Zhao, J.; Leung, H.; Ma, K.F.; Wang, W. A Review of Deep Learning Models for Time Series Prediction. *IEEE Sens. J.* **2021**, *21*, 7833–7848. [\[CrossRef\]](#)
26. Asghari, V.; Leung, Y.F.; Hsu, S.-C. Deep Neural Network Based Framework for Complex Correlations in Engineering Metrics. *Adv. Eng. Inform.* **2020**, *44*, 101058. [\[CrossRef\]](#)
27. Schulz, M.-A.; Yeo, B.T.T.; Vogelstein, J.T.; Mourao-Miranada, J.; Kather, J.N.; Kording, K.; Richards, B.; Bzdok, D. Different Scaling of Linear Models and Deep Learning in UKBiobank Brain Images versus Machine-Learning Datasets. *Nat. Commun.* **2020**, *11*, 4238. [\[CrossRef\]](#) [\[PubMed\]](#)
28. Chen, X.-W.; Lin, X. Big Data Deep Learning: Challenges and Perspectives. *IEEE Access* **2014**, *2*, 514–525. [\[CrossRef\]](#)
29. Gheisari, M.; Wang, G.; Bhuiyan, M.Z.A. A Survey on Deep Learning in Big Data. In Proceedings of the 2017 IEEE International Conference on Computational Science and Engineering (CSE) and IEEE International Conference on Embedded and Ubiquitous Computing (EUC), Guangzhou, China, 21–24 July 2017; Volume 2, pp. 173–180.
30. Miotto, R.; Wang, F.; Wang, S.; Jiang, X.; Dudley, J.T. Deep Learning for Healthcare: Review, Opportunities and Challenges. *Brief. Bioinform.* **2018**, *19*, 1236–1246. [\[CrossRef\]](#) [\[PubMed\]](#)
31. Zhou, X.; Liu, Z.; Wang, F.; Xie, Y.; Zhang, X. Using Deep Learning to Forecast Maritime Vessel Flows. *Sensors* **2020**, *20*, 1761. [\[CrossRef\]](#) [\[PubMed\]](#)
32. Mandalis, P.; Chondrodima, E.; Kontoulis, Y.; Pelekis, N.; Theodoridis, Y. Machine Learning Models for Vessel Traffic Flow Forecasting: An Experimental Comparison. In Proceedings of the 2022 23rd IEEE International Conference on Mobile Data Management (MDM), Paphos, Cyprus, 6–9 June 2022; pp. 431–436.
33. Li, Y.; Ren, H. Vessel Traffic Flow Prediction Using LSTM Encoder-Decoder. In *Proceedings of the 2022 5th International Conference on Signal Processing and Machine Learning*; Association for Computing Machinery: New York, NY, USA, 2022; pp. 1–7.
34. Chen, Y.; Shu, L.; Wang, L. Poster Abstract: Traffic Flow Prediction with Big Data: A Deep Learning Based Time Series Model. In Proceedings of the 2017 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), Atlanta, GA, USA, 1–4 May 2017; pp. 1010–1011.
35. Chen, C.; Liu, Z.; Wan, S.; Luan, J.; Pei, Q. Traffic Flow Prediction Based on Deep Learning in Internet of Vehicles. *IEEE Trans. Intell. Transp. Syst.* **2021**, *22*, 3776–3789. [\[CrossRef\]](#)
36. Zhou, Y.; Wang, X.; Jia, J. A Traffic Flow Forecasting Method Based on Transfer-Aware Spatio-Temporal Graph Attention Network. *ISPRS Int. J. Geo-Inf.* **2025**, *14*, 459. [\[CrossRef\]](#)
37. Chen, X.; Zhang, Y.; Feng, Z.; Zhang, H.; Biancardo, S.A.; Shang, W.-L. Spatiotemporal Ship Carbon Emission Exploitation and Low-Carbon Shipping Pathway Evaluation1. *Appl. Energy* **2026**, *422*, 128334. [\[CrossRef\]](#)
38. Shi, Q.; Abdel-Aty, M. Big Data Applications in Real-Time Traffic Operation and Safety Monitoring and Improvement on Urban Expressways. *Transp. Res. Part C Emerg. Technol.* **2015**, *58*, 380–394. [\[CrossRef\]](#)
39. Feng, J.; Du, C.; Mu, Q. Traffic Flow Prediction Based on Federated Learning and Spatio-Temporal Graph Neural Networks. *ISPRS Int. J. Geo-Inf.* **2024**, *13*, 210. [\[CrossRef\]](#)
40. Chen, D.; Huang, C.; Fan, T.; Lau, H.C.; Yan, X. Predictive Modelling for Vessel Traffic Flow: A Comprehensive Survey from Statistics to AI. *Transp. Saf. Environ.* **2025**, *7*, tda022. [\[CrossRef\]](#)
41. Shao, Y.; Zhao, Y.; Yu, F.; Zhu, H.; Fang, J. The Traffic Flow Prediction Method Using the Incremental Learning-Based CNN-LTSM Model: The Solution of Mobile Application. *Mob. Inf. Syst.* **2021**, *2021*, 5579451. [\[CrossRef\]](#)
42. Bilotta, S.; Collini, E.; Nesi, P.; Pantaleo, G. Short-Term Prediction of City Traffic Flow via Convolutional Deep Learning. *IEEE Access* **2022**, *10*, 113086–113099. [\[CrossRef\]](#)
43. Liu, D.; Qu, Q.; Chen, X. STAE-BiSSSM: A Traffic Flow Forecasting Model with High Parameter Effectiveness. *ISPRS Int. J. Geo-Inf.* **2025**, *14*, 388. [\[CrossRef\]](#)
44. Li, Z.; Han, Y.; Xu, Z.; Zhang, Z.; Sun, Z.; Chen, G. PMGCN: Progressive Multi-Graph Convolutional Network for Traffic Forecasting. *ISPRS Int. J. Geo-Inf.* **2023**, *12*, 241. [\[CrossRef\]](#)
45. Duan, Z.; Yang, Y.; Zhang, K.; Ni, Y.; Bajgain, S. Improved Deep Hybrid Networks for Urban Traffic Flow Prediction Using Trajectory Data. *IEEE Access* **2018**, *6*, 31820–31827. [\[CrossRef\]](#)
46. Sun, X.; Li, J.; Lv, Z.; Dong, C. Traffic Flow Prediction Model Based on Spatio-Temporal Dilated Graph Convolution. *KSII Trans. Internet Inf. Syst. TIIS* **2020**, *14*, 3598–3614. [\[CrossRef\]](#)
47. Li, Y.; Chai, S.; Wang, G.; Zhang, X.; Qiu, J. Quantifying the Uncertainty in Long-Term Traffic Prediction Based on PI-ConvLSTM Network. *IEEE Trans. Intell. Transp. Syst.* **2022**, *23*, 20429–20441. [\[CrossRef\]](#)

48. Zhang, Q.; Yu, K.; Guo, Z.; Garg, S.; Rodrigues, J.J.P.C.; Hassan, M.M.; Guizani, M. Graph Neural Network-Driven Traffic Forecasting for the Connected Internet of Vehicles. *IEEE Trans. Netw. Sci. Eng.* **2022**, *9*, 3015–3027. [\[CrossRef\]](#)
49. Liang, M.; Liu, R.W.; Zhan, Y.; Li, H.; Zhu, F.; Wang, F.-Y. Fine-Grained Vessel Traffic Flow Prediction with a Spatio-Temporal Multigraph Convolutional Network. *IEEE Trans. Intell. Transp. Syst.* **2022**, *23*, 23694–23707. [\[CrossRef\]](#)
50. Zhao, C.; Li, X.; Zuo, M.; Mo, L.; Yang, C. Spatiotemporal Dynamic Network for Regional Maritime Vessel Flow Prediction amid COVID-19. *Transp. Policy* **2022**, *129*, 78–89. [\[CrossRef\]](#) [\[PubMed\]](#)
51. Huan, Y.; Kang, X.; Zhang, Z.; Zhang, Q.; Wang, Y.; Wang, Y. AIS-Based Vessel Traffic Flow Prediction Using Combined EMD-LSTM Method. In *Proceedings of the 4th International Conference on Advanced Information Science and System*; Association for Computing Machinery: New York, NY, USA, 2023; pp. 1–8.
52. Mandalis, P.; Chondrodima, E.; Kontoulis, Y.; Pelekis, N.; Theodoridis, Y. A Transformer-Based Method for Vessel Traffic Flow Forecasting. *GeoInformatica* **2024**, *29*, 149–173. [\[CrossRef\]](#)
53. Zhou, H.; Zuo, Y.; Li, T.; Li, S. Application of PSO-LSTM for Forecasting of Ship Traffic Flow. In *Proceedings of the 2021 International Conference on Security, Pattern Analysis, and Cybernetics (SPAC)*, Chengdu, China, 18–20 June 2021; pp. 298–302.
54. Wang, D.; Meng, Y.; Chen, S.; Xie, C.; Liu, Z. A Hybrid Model for Vessel Traffic Flow Prediction Based on Wavelet and Prophet. *J. Mar. Sci. Eng.* **2021**, *9*, 1231. [\[CrossRef\]](#)
55. Teja, K.; Tiwari, R.; Mohanty, S. Adaptive Denoising of ECG Using EMD, EEMD and CEEMDAN Signal Processing Techniques. *J. Phys. Conf. Ser.* **2020**, *1706*, 012077. [\[CrossRef\]](#)
56. Zhang, J.; Qi, H.; Ju, X.; Wang, H.; Ye, G.; Huang, B.; Qi, M.; Tang, Y. Short-Term Photovoltaic Power Forecasting Based on EEMD Residual Secondary IWOA-VMD Decomposition and ISSA-Optimized BiGRU. *Sustainability* **2026**, *18*, 2234. [\[CrossRef\]](#)
57. Chen, H.-G.; Shen, J.-Y.; Chen, W.-H.; Huang, C.-S.; Yi, Y.-Y.; Qian, J.-C. Grinding Chatter Detection and Identification Based on BEMD and LSSVM. *Chin. J. Mech. Eng.* **2019**, *32*, 1. [\[CrossRef\]](#)
58. Liu, R.; Çağatay, I.; Xu, J.; Chen, J. End-to-End Deep Learning for Inventory Management in Capacitated Bike-Sharing Systems with Customer Roaming Behaviour. *Transp. Res. Part E Logist. Transp. Rev.* **2026**, *211*, 104852. [\[CrossRef\]](#)
59. Chen, Y.; Lopez-Moreno, I.; Sainath, T.N.; Visontai, M.; Alvarez, R.; Parada, C. Locally-Connected and Convolutional Neural Networks for Small Footprint Speaker Recognition. In *Interspeech*; International Speech Communication Association (ISCA): Grenoble, France, 2015; pp. 1136–1140.
60. Tang, L.; Zhang, C.; Li, T.; Li, L. A Novel BEMD-Based Method for Forecasting Tourist Volume with Search Engine Data. *Tour. Econ.* **2021**, *27*, 1015–1038. [\[CrossRef\]](#)
61. Nunes, J.C.; Bouaoune, Y.; Delechelle, E.; Niang, O.; Bunel, P. Image Analysis by Bidimensional Empirical Mode Decomposition. *Image Vis. Comput.* **2003**, *21*, 1019–1026. [\[CrossRef\]](#)
62. Bhuiyan, S.M.A.; Adhami, R.R.; Khan, J.F. A Novel Approach of Fast and Adaptive Bidimensional Empirical Mode Decomposition. In *Proceedings of the 2008 IEEE International Conference on Acoustics, Speech and Signal Processing*, Las Vegas, NV, USA, 31 March–1 April 2008; pp. 1313–1316.
63. Nandhini Abirami, R.; Durai Raj Vincent, P.M.; Srinivasan, K.; Tariq, U.; Chang, C.-Y. Deep CNN and Deep GAN in Computational Visual Perception-Driven Image Analysis. *Complexity* **2021**, *2021*, 5541134. [\[CrossRef\]](#)
64. Xiao, L.; Lan, T.; Xu, D.; Gao, W.; Li, C. A Simplified CNNs Visual Perception Learning Network Algorithm for Foods Recognition. *Comput. Electr. Eng.* **2021**, *92*, 107152. [\[CrossRef\]](#)
65. O'Shea, K.; Nash, R. An Introduction to Convolutional Neural Networks. *arXiv* **2015**, arXiv:1511.08458.
66. Jiang, S.; Fan, H.; Wang, C. Improvement of Typhoon Intensity Forecasting by Using a Novel Spatio-Temporal Deep Learning Model. *Remote Sens.* **2022**, *14*, 5205. [\[CrossRef\]](#)
67. Yin, W.; Kann, K.; Yu, M.; Schütze, H. Comparative Study of CNN and RNN for Natural Language Processing. *arXiv* **2017**, arXiv:1702.01923.
68. Salehinejad, H.; Sankar, S.; Barfett, J.; Colak, E.; Valaee, S. Recent Advances in Recurrent Neural Networks. *arXiv* **2018**, arXiv:1801.01078.
69. Giles, C.L.; Lawrence, S.; Tsoi, A.C. Noisy Time Series Prediction Using Recurrent Neural Networks and Grammatical Inference. *Mach. Learn.* **2001**, *44*, 161–183. [\[CrossRef\]](#)
70. Hewamalage, H.; Bergmeir, C.; Bandara, K. Recurrent Neural Networks for Time Series Forecasting: Current Status and Future Directions. *Int. J. Forecast.* **2021**, *37*, 388–427. [\[CrossRef\]](#)
71. Shewalkar, A.; Nyavanandi, D.; Ludwig, S.A. Performance Evaluation of Deep Neural Networks Applied to Speech Recognition: RNN, LSTM and GRU. *J. Artif. Intell. Soft Comput. Res.* **2019**, *9*, 235–245. [\[CrossRef\]](#)
72. Ribeiro, A.H.; Tiels, K.; Aguirre, L.A.; Schön, T. Beyond Exploding and Vanishing Gradients: Analysing RNN Training Using Attractors and Smoothness. In *Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics*, Online, 26–28 August 2020; pp. 2370–2380.
73. Gu, A.; Gulcehre, C.; Paine, T.; Hoffman, M.; Pascanu, R. Improving the Gating Mechanism of Recurrent Neural Networks. In *Proceedings of the 37th International Conference on Machine Learning*, Virtual Event, 13–18 July 2020; pp. 3800–3809.

74. Lu, Y.; Liu, D.; Liang, Z.; Liu, R.; Chen, P.; Liu, Y.; Li, J.; Feng, Z.; Li, L.M.; Sheng, B.; et al. A Pretrained Transformer Model for Decoding Individual Glucose Dynamics from Continuous Glucose Monitoring Data. *Natl. Sci. Rev.* **2025**, *12*, nwaf039. [[CrossRef](#)] [[PubMed](#)]
75. Ren, L.; Wang, H.; Mo, T.; Yang, L.T. A Lightweight Group Transformer-Based Time Series Reduction Network for Edge Intelligence and Its Application in Industrial RUL Prediction. *IEEE Trans. Neural Netw. Learn. Syst.* **2025**, *36*, 3720–3729. [[CrossRef](#)] [[PubMed](#)]
76. Sun, W.; Wu, H.; Yan, Y.; Zhang, X.; Yao, X.; Li, R.; Zuo, J.; Li, W.; Ouyang, H. Living Joint Prosthesis with In-Situ Tissue Engineering for Real-Time and Long-Term Osteoarticular Reconstruction. *Bioact. Mater.* **2025**, *48*, 431–442. [[CrossRef](#)] [[PubMed](#)]
77. Khosravi, K.; Panahi, M.; Golkarian, A.; Keesstra, S.D.; Saco, P.M.; Bui, D.T.; Lee, S. Convolutional Neural Network Approach for Spatial Prediction of Flood Hazard at National Scale of Iran. *J. Hydrol.* **2020**, *591*, 125552. [[CrossRef](#)]
78. Saha, S.; Saha, A.; Hembram, T.K.; Mandal, K.; Sarkar, R.; Bhardwaj, D. Prediction of Spatial Landslide Susceptibility Applying the Novel Ensembles of CNN, GLM and Random Forest in the Indian Himalayan Region. *Stoch. Environ. Res. Risk Assess.* **2022**, *36*, 3597–3616. [[CrossRef](#)]
79. Han, Y.; Yu, X.; Yuan, J.; Zhu, M.; Xie, S. Small-Sample InSAR Time-Series Data Prediction Method Based on Generative Models. *Nat. Resour. Res.* **2025**, *34*, 815–838. [[CrossRef](#)]
80. Zhao, Y.; Ban, Y. Near Real-Time Wildfire Progression Mapping with VIIRS Time-Series and Autoregressive SwinUNETR. *Int. J. Appl. Earth Obs. Geoinf.* **2025**, *136*, 104358. [[CrossRef](#)]
81. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, L.; Polosukhin, I. Attention Is All You Need. *arXiv* **2017**, arXiv:1706.03762v7.
82. Yang, J.; Yang, G. Modified Convolutional Neural Network Based on Dropout and the Stochastic Gradient Descent Optimizer. *Algorithms* **2018**, *11*, 28. [[CrossRef](#)]
83. Phong, N.H.; Santos, A.; Ribeiro, B. PSO-Convolutional Neural Networks with Heterogeneous Learning Rate. *IEEE Access* **2022**, *10*, 89970–89988. [[CrossRef](#)]
84. Li, Z.; Liu, F.; Yang, W.; Peng, S.; Zhou, J. A Survey of Convolutional Neural Networks: Analysis, Applications, and Prospects. *IEEE Trans. Neural Netw. Learn. Syst.* **2022**, *33*, 6999–7019. [[CrossRef](#)] [[PubMed](#)]
85. Urošević, V.; Dimitrijević, S. Optimum Input Sequence Size for a Sliding Window-Based LSTM Neural Network Used in Short-Term Electrical Load Forecasting. In Proceedings of the 2021 29th Telecommunications Forum (TELFOR), Belgrade, Serbia, 23–24 November 2021; pp. 1–4.
86. Smith, L.N. A Disciplined Approach to Neural Network Hyper-Parameters: Part 1—Learning Rate, Batch Size, Momentum, and Weight Decay. *arXiv* **2018**, arXiv:1803.09820. [[CrossRef](#)]
87. Öztürk, Ş.; Özkaya, U.; Akdemir, B.; Seyfi, L. Convolution Kernel Size Effect on Convolutional Neural Network in Histopathological Image Processing Applications. In Proceedings of the 2018 International Symposium on Fundamentals of Electrical Engineering (ISFEE), Bucharest, Romania, 1–3 November 2018; pp. 1–5.
88. Moosavi, V.; Gheisoori Fard, Z.; Vafakhah, M. Which One Is More Important in Daily Runoff Forecasting Using Data Driven Models: Input Data, Model Type, Preprocessing or Data Length? *J. Hydrol.* **2022**, *606*, 127429. [[CrossRef](#)]
89. Bottou, L.; Curtis, F.E.; Nocedal, J. Optimization Methods for Large-Scale Machine Learning. *SIAM Rev.* **2018**, *60*, 223–311. [[CrossRef](#)]
90. Raftery, A.; Lewis, S. The Number of Iterations, Convergence Diagnostics and Generic Metropolis Algorithms. *Pract. Markov Chain Monte Carlo* **1970**, *7*, 763–773.
91. Smith, S.L.; Kindermans, P.-J.; Ying, C.; Le, Q.V. Don't Decay the Learning Rate, Increase the Batch Size. *arXiv* **2017**, arXiv:1711.00489. [[CrossRef](#)]
92. Smith, L.N. Cyclical Learning Rates for Training Neural Networks. In Proceedings of the 2017 IEEE Winter Conference on Applications of Computer Vision (WACV), Santa Rosa, CA, USA, 24–31 March 2017; pp. 464–472.
93. Luo, W.; Li, Y.; Urtasun, R.; Zemel, R. Understanding the Effective Receptive Field in Deep Convolutional Neural Networks. In *Proceedings of the Advances in Neural Information Processing Systems*; Curran Associates, Inc.: Red Hook, NY, USA, 2016; Volume 29.
94. Li, X.; Wang, W.; Hu, X.; Yang, J. Selective Kernel Networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*; IEEE: New York, NY, USA, 2019; pp. 510–519.
95. Luo, L.; Wang, M.; Qiu, C.; Kan, R.; Shen, X.; Feng, L. Ship Traffic Flow Analysis and Prediction in High-Traffic Areas Under Complex Environments. *Appl. Sci.* **2025**, *15*, 11776. [[CrossRef](#)]

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.