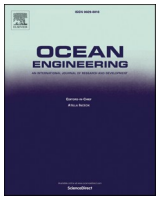
A spatial-temporal attention method for the prediction of multi ship time headways using AIS data

Ocean Engineering 311 (2024) 118927

Contents lists available at ScienceDirect

Ocean Engineering

journal homepage: www.elsevier.com/locate/oceaneng



Research paper

A spatial-temporal attention method for the prediction of multi ship time headways using AIS data

Quandang Ma [a], Xu Du [a], Mingyang Zhang [b,*], Hongdong Wang [c], Xiao Lang [d], Wengang Mao [d]

[a] Hubei Key Laboratory of Inland Shipping Technology, School of Navigation, Wuhan University of Technology, Wuhan, China
[b] Department of Mechanical Engineering, School of Engineering, Aalto University, Espoo, Finland
[c] MOE Key Laboratory of Marine Intelligent Equipment and System, Shanghai Jiao Tong University, Shanghai, 200240, China
[d] Department of Mechanics and Maritime Sciences, Chalmers University of Technology, Gothenburg, Sweden

ARTICLE INFO

Keywords:
Maritime traffic management
Ship time headway prediction
Deep learning
Graph neural network
Attention mechanism

ABSTRACT

Ship Time Headway (STH) is the time interval between two consecutive ships arriving in the same water area. It serves as a crucial indicator for visually measuring the probability of ship congestion and the frequency of passage in busy waterways. Accurately predicting the STH is crucial for effective maritime traffic management. In this paper, we propose a deep learning method aimed at simultaneously predicting the STH in multiple water areas (multi-STH). This method integrates the Variational Mode Decomposition (VMD) algorithm with the Spatial-Temporal Attention Graph Convolution Network (STAGCN) to deeply capture the complex spatial-temporal features between STHs of each water areas. STH sequences were obtained from Automatic Identification System (AIS) for each reach, ensuring that these sequences remained numerically continuous on the same timeline. The VMD algorithm was employed to decompose the sequences into multi-feature inputs for the STAGCN, training the model in conjunction with the inland waterway traffic network to capture the patterns of variation in STH between the water areas. Extensive experiments demonstrate that the proposed prediction method surpasses the accuracy and robustness of other existing methods, exhibiting excellent prediction performance in the STHs of various waterways. The multi-STH prediction study accounts for the inherent correlation between inland waterways, substantially improving prediction efficiency compared to single-waterway STH prediction. This study may have the potential to provide useful support for traffic management. This may be of practical significance in enhancing the safety of inland waterways navigation.

1. Introduction

Economic globalization has contributed to the continuous development of trade in goods between countries (Li et al., 2023a,b). Shipping has always been the primary mode of transporting goods worldwide. Compared to other modes of transport, shipping offers advantages such as low transport costs, large carrying capacity, and the ability to handle a wide variety of goods (Xing et al., 2023; S. Wang et al., 2023; Wang et al., 2023, 2023b; Y. Wang et al., 2023 b). These factors have led to the widespread use of maritime transportation on inland waterways. However, the frequent operations of maritime transportation have resulted in a continuous increase in the risks of ship navigation at sea and in inland waterways. Navigational accidents can lead to casualties, substantial property damage, and even the paralysis of navigation channels in inland waterways (Zhang et al., 2021; Shi et al., 2019; Chen et al., 2024). Therefore, ensuring the safety of ship navigation and effectively predicting potential navigational risks has become a critical issue, supporting the shift of maritime management from passive to proactive.

With the widespread adoption of Automatic Identification Systems (AIS) for ships, the vast amount of historical AIS data provides a key source of information for researchers to understand ship movement patterns, water traffic patterns, and other related aspects. Notable, a novel data-driven model was proposed to analyze navigation mode determination in ice-covered waters using AIS data, first quantitatively identifying and assessing the influencing factors for the need for icebreaker assistance, examining the impact of these factors on ships with different ship ice classes, and highlighting the innovation of influencing factor analysis through model performance comparison (C. Liu et al., 2024). Deep mining of this historical information plays an important role in optimizing port operations, preventing and controlling navigational risks, and improving navigational efficiency (Gao et al.,

* Corresponding author. Otakaari 4, Koneteknikka 1, 02150, Espoo, Finland.
E-mail address: Mingyang.0.zhang@aalto.fi (M. Zhang).

https://doi.org/10.1016/j.oceaneng.2024.118927
Received 22 June 2024; Received in revised form 3 August 2024; Accepted 5 August 2024
Available online 19 August 2024

Nomenclature

Variable	*definition*
AIS	Automatic Identification System
AM	Attention Mechanism
Bi-LSTM	Bidirectional LSTM
CNN	Convolutional Neural Network
DCC	Dilated Causal Convolution
EMD	Empirical Mode Decomposition
GAT	Graph Attention Network
GCN	Graph Convolutional Network
GNN	Graph Neural Network
GRU	Gate Recurrent Unit
HA	Historical Average
IMFs	Intrinsic Mode Functions
LR	Learning Rate
LSTM	Long Short-Term Memory
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
multi-STH	STH in multiple water areas
R2	Coefficient of Determination
RMSE	Root Mean Square Error
RNN	Recurrent Neural Network
S-T AM	Spatial-Temporal Attention Mechanism
STAGCN	Spatial-Temporal Attention Graph Convolution Network
STF	Ship Traffic Flow
STGCN	Spatial-Temporal Graph Convolution Network
STH	Ship Time Headway
SVR	Support Vector Regression
VMD	Variational Mode Decomposition
N	cross section position
$lon^N_{vir\alpha}, lon^N_{vir\beta}$	Longitude of the start of the cross section
$lat^N_{vir\alpha}, lat^N_{vir\beta}$	Latitude of the start of the cross section
α, β	Cross section start point
$P^N_{i,left}, P^N_{i,right}$	Nearest trajectory points of the cross section
$\alpha\vec{P}^N_{i,left}$, $\alpha\vec{P}^N_{i,right}$	Direction from point α in the cross section to the nearest trajectory point
$\bar{v}^N_i$	Average ship speed
ΔT^N_i	Time difference between the nearest trajectory points of the cross section
ΔD^N_i	Distance between the nearest trajectory points of the cross section
Δt^N_i	Time for the ship to reach the cross section from the nearest trajectory point of the cross section
T^N_i	Timestamp of ship's arrival at the cross section
t^N_i	STH value
X^N_h	Average level of STH in the h-th unit of time
F	The number of VMD decompositions
u_f	The decomposition mode
ω_f	Centre frequency of the mode
$\delta(t)$	The Dirac distribution
x^N	The decomposed sequence
ε	The penalty factor
λ	The Lagrange multiplier
τ	The update parameter of the Lagrange multiplier
ρ	The threshold for the VMD to stop iteration
G	The inland transport network
x	The model Inputs
T, T', S, S'	The attention matrix
U, W, V, b	The S-T AM Parameters
σ	The Sigmoid activation function
L	The Laplacian matrix
D	The diagonal matrix of node degrees
A	The adjacency matrix
I_N	The unit matrix
$\mathscr{U}$	The Fourier basis
Λ	The eigenvalue diagonal matrix
$\hat{x}$	The signal obtained by the graph Fourier transform
g_θ	The kernel of GCN
ψ	The Chebyshev polynomial coefficient vector
$T_k(x)$	The Chebyshev polynomial
K	The order of the Chebyshev polynomial
λ_{max}	The maximum eigenvalue of the Laplacian matrix
M	The observation time step
O	The prediction time step
H_n	The model hidden state
$\mathscr{F}(\bullet)$	The mapping function

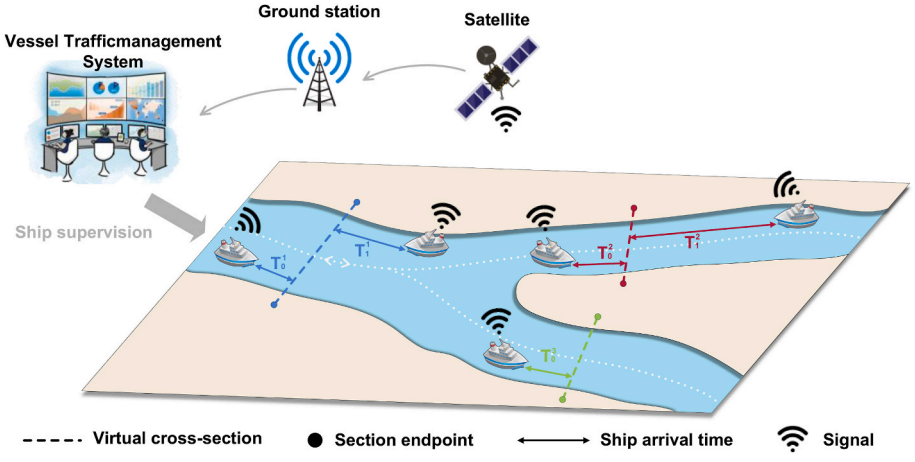


Fig. 1. Schematic diagram of multi-STH.

2023; Liu et al., 2022; R. W. Liu et al., 2024). In recent years, Ship Traffic Flow (STF) data mined from AIS by researchers has been able to provide a macroscopic assessment of ship traffic patterns in study waterways (Zhang et al., 2022). STF represents the total number of ships passing through a specific location within a specified time (Xiao et al., 2022). Changes in STF values can indicate the busyness and intensity of traffic in target waterways, with a consistently increasing STF in limited waterways suggesting a potentially higher navigational risk (Su et al., 2022). However, in busy and complex navigational environments, STF is limited in the information it can provide. In particular, the simultaneous existence of multiple ship behaviors such as sailing, anchoring, berthing, and de-berthing in a given water area may not accurately reflect the frequency and density of ship traffic. Therefore, analyzing macro-indicators such as STF alone is insufficient to determine the risks present in the waterways. It is necessary to introduce a micro-indicator, such as Ship Time Headway (STH), to provide more fine-grained feedback on the navigational risk in the target water areas (Ma et al., 2024a, b).

STH is the time interval between the arrivals of two consecutive ships in the same water area, as illustrated in Fig. 1. Analyzing changes in STH values can intuitively reflect the frequency of ship traffic and the probability of congestion in target water areas, particularly in busy or narrow channels with high ship density. Maintaining adequate ship safety spacing is essential to reduce navigational and congestion risks. The variation of STH as a micro-indicator is closely related to the behavior of individual ships and the overall structure of ship traffic. The specific roles of STH can be summarized as follows:

- Quantitative Assessment: STH can quantitatively assess the level of ship activity in a given water area and serve as a visual indicator of the frequency of ship traffic. A sudden reduction in STH indicates a higher risk of congestion or collision.
- Relation to STF: Changes in STH are related to changes in Ship Traffic Flow (STF). In heavy traffic waterways, controlling the appropriate STH can effectively alleviate traffic flow bottlenecks.
- Support for Maritime Supervision: STH can help maritime supervisors develop accurate and sound management measures and ship scheduling decisions for various water areas, including harbor water areas, fairway water areas, bridge water areas (Lei et al., 2021), and lock water areas (Deng et al., 2021). Additionally, STH provides reference information for developing subsequent voyage plans for ships in various waterways.

Understanding and studying STH is crucial for comprehending traffic conditions, improving ship traffic management, and reducing navigational risks in the water areas. Therefore, accurately and consistently predicting STH is essential, as it can provide reliable evaluation information for dynamically changing traffic patterns.

The nature of STH is time series data, making it similar to the STF prediction method (Wang et al., 2021). The most common and effective prediction methods for STH fall into two main categories: modeling-based machine learning prediction methods (Xu et al., 2017; Gao et al., 2023) and learning-based deep learning prediction methods (Kamilaris and Prenafeta-Boldú, 2018; Hinton and Salakhutdinov, 2006). However, STH is subject to non-linear variations due to uncontrollable factors (e.g., bad weather, traffic control, traffic accidents), resulting in non-stationary STH data. Modeling-based prediction methods may face difficulties in dealing with these irregularly fluctuating time-series data and are often unable to effectively capture the contextual relationships in the data (Semenoglou et al., 2023).

The increasing computational power and the clever design of Back Propagation algorithms have enabled deep learning methods to exhibit powerful learning capabilities. Initially, the emergence of Recurrent Neural Networks (RNN) proved effective in processing time series data (Cossu et al., 2021). By adding contextual association to the input, RNNs enable accurate prediction of time-series data. However, RNNs have limitations such as difficulty in learning long-term dependencies of complex data, susceptibility to gradient explosion or vanishing, and inability to capture spatial features (Bengio et al., 1994).

To address these issues, studies combined a variant of RNN, the Long Short-Term Memory (LSTM) neural network, with a Convolutional Neural Network (CNN). LSTM solves the problems of gradient explosion/vanishing and long-term dependency by adding forget gates, input gates, and output gates to the network (Yu et al., 2019). CNNs can extract spatial features from data (Alzubaidi et al., 2021). This combination has shown excellent predictive performance in spatial-temporal prediction (Kim and Cho, 2019). As deep learning continues to evolve, researchers have introduced graph theory into neural networks to form Graph Convolutional Networks (GCN), a type of convolutional neural network that can directly act on graphs and utilize their structural information (Zhang et al., 2019 c). GCNs can extract informative features in non-Euclidean data and have been successfully used in road traffic scenarios. In these networks, sensors and their interoperability are represented as nodes and edges in a graph, integrating spatial, temporal, and correlation information into the model. This approach allows for accurate and efficient predictions in more complex scenarios. Furthermore, many studies overlook the impact of data input forms on the performance of prediction methods, which relates to how much information the model receives from the data. For example, resetting a one-dimensional time series to high-dimensional data as an input strategy can effectively improve the model's ability to capture complex relationships between historical and current data (Chen and Sun, 2021). Sufficiently rich information input raises the upper limit of the model's capability to capture these relationships.

Therefore, this paper proposes a multi-STH prediction method based on deep learning. Multiple virtual cross-sections are set up within a continuous segment, and the times of real ship trajectories passing through these cross-sections are recorded to construct the multi-STH dataset.

The VMD algorithm, spatial-temporal attention module, GCN, and CNN are integrated into the VMD-STAGCN framework, enabling multi-STH prediction. The key innovations and contributions of this paper are summarized as follows:

- Development of an inland waterway transport network that leverages the connectivity and inherent interrelationships between inland waterways. This network combines spatial-temporal attention with a graph-based learning approach to deeply analyze the spatial-temporal variation patterns of STH across different water areas, facilitating simultaneous multi-water areas prediction.
- Extraction of temporal features from the multi-STH dataset using the VMD algorithm. These high-dimensional features enhance the model's ability to learn the evolutionary patterns of STH across different water areas.
- Provision of accurate and stable STH predictions that aid maritime regulators and ship pilots in understanding traffic patterns across extensive water areas, thereby offering robust support for safe navigation in these water areas.

The paper is structured as follows: Section 2 describes the issues raised in this paper. Section 3 presents the methodology and steps for the collection and prediction of multi-STH data to formulate the problem. Section 4 applies a real case to the method proposed in this paper, and the experimental results fully validate the effectiveness of the proposed method. Finally, Section 5 summarizes the main findings and limitations of the study and suggests directions for future research work.

2. Literature review and problem description

In this section, we provide a brief review of the relevant literature on both traffic flow analysis and traffic flow prediction and analyze the existing limitations. The research questions addressed in this paper are

also elaborated.

2.1. Literature review

Massive amounts of historical AIS data contain important information, and both quantitative and qualitative analyses help to better understand traffic flow patterns and provide insights for traffic safety management (Liang et al., 2021; Sui et al., 2024). In recent years, researchers have conducted extensive research on traffic flow analysis. For instance, Zhang et al. (2019) proposed an autocorrelation model to quantitatively describe the spatial-temporal dynamics of maritime traffic flows. Kujala et al. (2009) explored the relationship between traffic density and marine accidents in specific areas by analyzing historical AIS data. Ma et al. (2024b) utilized a spatial autocorrelation model and Moran's index to extract maritime traffic flow characteristics (average ship speed, density, etc.) and analyzed their spatial and temporal distribution. Kang et al. (2018) assessed the relationship between traffic flow speed and density in the target watershed using a basic traffic flow diagram and the least squares method. However, these macro-assessment methods often ignore individual ship behavior, preventing detailed analysis of traffic flows.

To explore individual ship behaviors out of maritime traffic, Xiao et al. statistically analyzed the dynamic values of AIS information (lateral position, speed, heading, and ship arrival interval) to characterize ship traffic (Xiao et al., 2015). Zhang et al. (2022) proposed a micro-prediction method for analyzing the complexity of traffic flow, highlighting the importance of ship arrival intervals for traffic flow assessment. In our previous study, the concept of STH was introduced and confirmed as predictable. The prediction results can effectively forecast ship arrivals, prevent congestion, and avoid accidental traffic accidents in inland waterways (Ma et al., 2024a,b). Deep learning models, with their sophisticated structures and numerous trainable parameters, can learn complex feature representations from data, improving the accuracy of regression and classification tasks and offering opportunities to handle complex time series data. Mainstream deep learning methods for traffic flow prediction include CNN (Zhang et al., 2019 b), RNN (Suo et al., 2020), Graph Neural Networks (GNN) (Yu et al., 2017), Attentional Mechanisms (AM) (Sun et al., 2020), and hybrid neural networks (Ma et al., 2022; Cao et al., 2022). For example, Belhadi et al. (2020) explored the application of RNN for long-term traffic flow prediction, proposing a method for long-term prediction using multiple data sources. Yang et al. (2019) developed a CNN-based multi-feature prediction model that integrates data from various periods with external factors to enhance prediction accuracy. Li et al. (2023) introduced a prediction architecture combining Graph Attention Network (GAT) and Dilated Causal Convolution (DCC), which considers the spatial-temporal variation patterns of traffic flows across multiple port areas to improve prediction accuracy. Hao et al. (2019) proposed an AM-based seq2seq framework for traffic flow prediction, an end-to-end architecture advantageous for capturing long-time dependencies, and extended its application to similar scenarios. Zhou et al. (2020) combined CNN, LSTM, and Bi-directional LSTM (Bi-LSTM) with CNN for traffic flow prediction, gridding the target water areas to predict traffic flow for each grid and validating the model's performance using a real dataset.

However, most existing traffic prediction studies focus on flow prediction and lack accurate prediction of micro traffic characteristics such as STH. This makes it difficult to effectively regulate ships traveling in target water areas under complex traffic conditions. Additionally, most studies have been limited to single-water forecasts, ignoring the inherent correlations that may exist between watersheds, resulting in reduced regulation efficiency and the inability to make accurate management decisions for a wide range of water areas. To fill this research gap, it is necessary to develop a method that can predict multi-STH, see more in Section 3.

2.2. Problem description

In previous studies, predictions based on maritime data have been limited to individual area. Expanding the prediction target to multiple water areas, while taking into account the inevitable correlation between water areas, would substantially increase the value of the study for real-world applications. However, similar closed water areas, such as inland rivers or longer waterways, typically exhibit more distinct patterns of traffic behavior.

For instance, in the case of inland rivers, as shown in Fig. 2, there are three primary scenarios for ship presence: (1) Sailing to upstream areas, (2) Sailing to downstream areas, and (3) Remaining in nearby areas (anchoring, entering, and exiting harbors, etc.). We used the Pearson correlation coefficient method to calculate the correlation of these multiple water areas traffic flows in the inland river. As shown in the correlation matrix in Fig. 2, there are positive correlations between traffic flows in different water areas within an inland river, and these correlations are related to the geographical location of the areas.

Therefore, it is feasible to make simultaneous forecasts for these waters by leveraging this inherent dependency. However, the correlation values are not very high, making the prediction of multi-STH challenging. To address this, the paper proposes a spatial-temporal attention method for the prediction of multi-STH using AIS data. This approach allows for a more comprehensive and accurate prediction model that reflects the interconnected nature of traffic flows in these waterways. By predicting multi-STH, it may enhance the effectiveness of traffic management and safety measures across multiple water areas, leading to improved navigation planning and congestion mitigation.

3. Methodology

In this section, a framework for the prediction of multi-STH using a spatial-temporal attention method is presented. Fig. 3 illustrates the proposed framework, which consists of three steps:

Step 1: **Collection of multi-STH.** After pre-processing the raw AIS data, the time of arrival at the proposed virtual cross-section is estimated based on the spatial motion of the ship historical trajectory to obtain the STH sub-sequence for each target water area. In order to make the data from these different water areas uniform on the timeline, multiple STH sub-data from each water in the same time window were averaged as STH data for the current time period.

Step 2: **Construction of multi-feature STH dataset using VMD algorithm.** The raw STH data has non-linear variations due to shipping uncertainty, with sharp fluctuations, making it difficult for the model to learn the pattern of change in the time series. The VMD algorithm decomposes the STH data of a certain area into several Intrinsic Mode Functions (IMFs) as multiple features to assist the model in learning the changing law of STH.

Step 3: **Deep learning method for multi-STH prediction.** Inland waterway transport networks are formed using transport relationships between multiple target water areas. Since these water areas are related to each other to varying degrees in both the temporal and spatial dimensions, weights were obtained by calculating attention scores for both the temporal and spatial dimensions. The inputs combined with the attention weights go into the spatial-temporal convolution module, which extracts spatial features between areas (nodes) by GCN. A conventional CNN operation is performed on the input temporal dimension to obtain temporal features. To enable the model to fully learn the spatial and temporal variations and connections between STH in multiple water areas.

3.1. Collection of multi-STH data

In this paper, virtual cross-sections are set up in several target water areas of an inland river based on previous research work (Ma et al., 2024a,b). The multi-STH data is collected from virtual cross-sections.

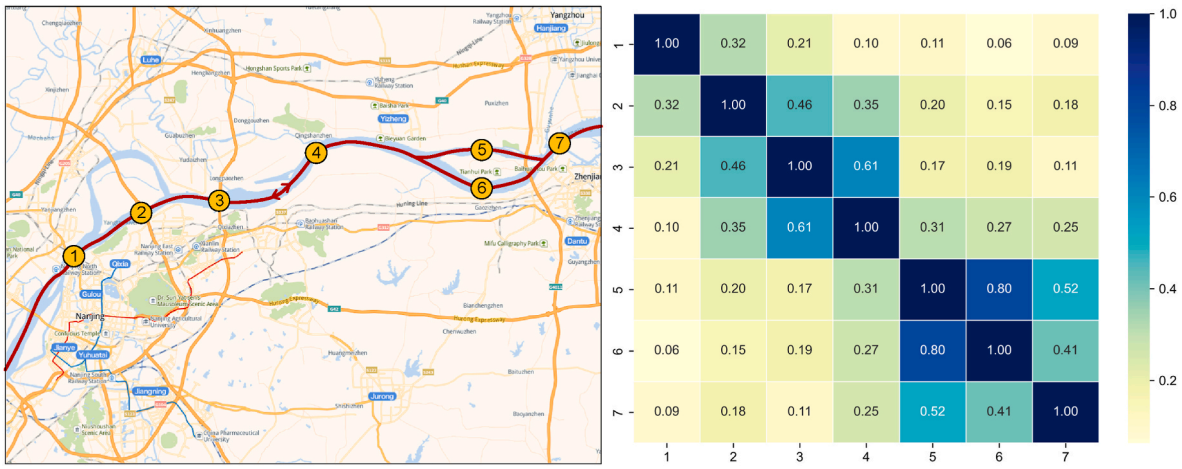


Fig. 2. Pearson correlation coefficients of STH between inland multiple water areas.

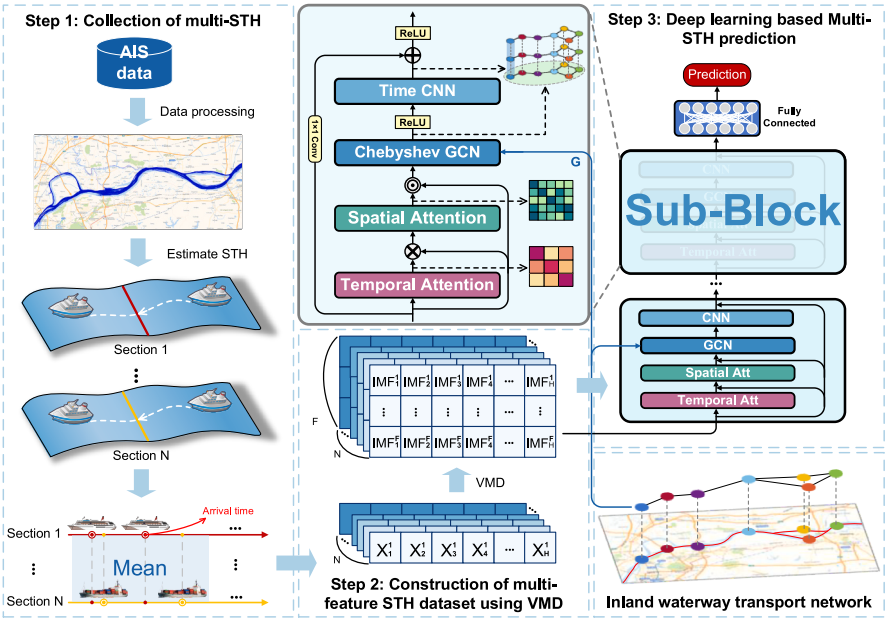


Fig. 3. The framework for multi-STH prediction using the proposed spatial-temporal attention method.

3.1.1. *Capturing sub-trajectories of ships passing through the cross-section*

In this section, all ships (upstream and downstream) passing through the cross-section are calculated to prevent a high number of missing values from being generated when the ship flow is too low at a given time in the inland waterway. The proposed virtual section consists of α, β connection $Virsection^N = \left(lon^N_{vir\alpha}, lon^N_{vir\beta}, lat^N_{vir\alpha}, lat^N_{vir\beta}\right)$, $N \in R$, N denotes the finite set of water area locations (nodes). If the ship passes through the section, then there must be a sub-trajectory that intersects the section. The sub-trajectory segment consists of a connection of two consecutive trajectory points of the closest cross-section of $P^N_{i,left}$ and $P^N_{i,right}$. i denotes the i-th trajectory. The two points are determined by Eq. (1) as

$$\alpha\vec{P}^N_{i,left} < \vec{Virsection} < \alpha\vec{P}^N_{i,right}, \tag{1}$$

where $\vec{Virsection}$ represents the direction from α to β, $\alpha\vec{P}^N_{i,left}$ and $\alpha\vec{P}^N_{i,right}$ represent the direction from α to the left and right neighbouring trajectory points, respectively.

3.1.2. *STH subsequence estimation*

To determine the time at which the ship finally arrives at the cross-section, the sub-trajectory's voyage ΔD^N_i is calculated using the time difference ΔT^N_i between the AIS recording times at points $P^N_{i,left}$ and $P^N_{i,right}$. Here the Euclidean distance calculation is used and the default speed of the ship is a uniform speed while passing through the segment. Eq. (2) is used to calculate the average speed $\bar{v}^N_i$ by

$$\bar{v}^N_i = \frac{\Delta D^N_i}{\Delta T^N_i} \tag{2}$$

The perpendicular distance Δd^N_i is calculated between $P^N_{i,right}$ (or $P^N_{i,left}$) and $Virsection^N$. Using $\bar{v}^N_i$ to derive the required time Δt^N_i for the ship to get to $Virsection^N$ at the position of $P^N_{i,left}$ (or $P^N_{i,right}$), the time of arrival of the ship, T^N_i, can be estimated by Eq. (3) and Eq. (4) as

$$Arrival\_time^N_i = \Delta t^N_i + T^N_{i,left}, \tag{3}$$

$$t^N_i = T^N_{i+1} - T^N_i, \tag{4}$$

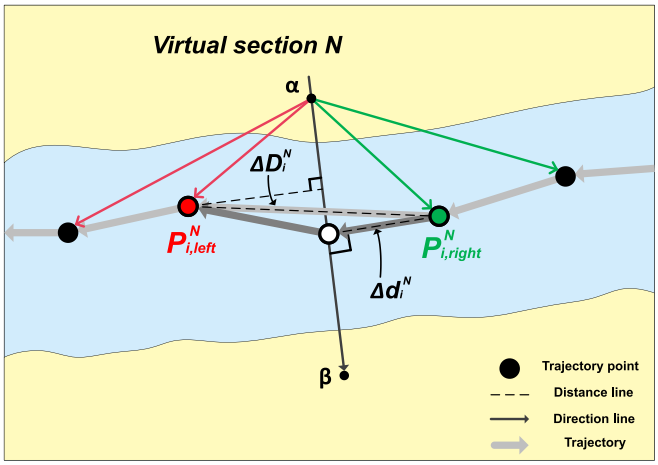


Fig. 4. STH estimation process (Ma et al., 2024a,b).

where $T^N_{i,left}$ refers to the time when the ship arrives at the position $P^N_{i,left}$. Repeating the above steps for all ships involved in the calculation, as shown in Fig. 4., we can get N ship arrival time series $\left(T^N_0, T^N_1, T^N_2, \ldots, T^N_M\right)$. The sequence consists of timestamps, and the STH subsequence is obtained by calculating the difference between the arrival times of two neighbouring ships in each sequence $\left(t^N_0, t^N_1, t^N_2, \ldots, t^N_m | m = M-1\right)$.

3.1.3. Computing continuous multi-STH data

The N STH subsequences estimated by the above process all possess a different number of elements and these different cross-sections are discrete from each other, which makes it extremely difficult to analyze the data and subsequent prediction. Therefore, in this section, these subsequences are reassigned, aiming to ensure that these cross-sections can appear as consecutive values on the same timeline. In this paper, a fixed time window (unit time) is set, as shown in Fig. 5. The time window slides from left to right along the time axis and is used to collect the elements of the STH subsequence in each unit of time $Window^N_h = \left(t^N_0, \ldots, t^N_n\right)$, with h denoting the h-th unit of time. By taking the mean value of these elements as STH per unit time, N equal length STH sequences will be obtained, $X^N = \left(X^N_1, \ldots, X^N_h, \ldots, X^N_H\right)$. At the h-th unit time, there exists a sequence of length N, $X_h = \left(X^1_h, X^2_h, \ldots, X^N_h\right)$, where X^N_h is defined by

$$X^N_h = \frac{\sum_{i=0}^{I} t^N_i}{n}, \tag{5}$$

where n denotes the number of ships passing through the cross-section in the h-th unit of time.

Due to some uncontrollable factors (i.e., bad weather, traffic accidents, channel maintenance, etc.) in the unit time of the ship flow is too small or no ship through the situation, especially in the tributaries, which may appear in the missing values or noise values lead to violent oscillations in the data, which have an impact on the prediction results. Therefore, it is necessary to introduce some thresholds to avoid similar situations: (1) $X^N_h = 0$ when no ship passes through the cross-section in the h-th unit of time, and (2) $X^N_h = 1$ when only one ship passes through the cross-section in the h-th unit of time, and X^N_h with thresholds is given by

$$X^N_h = \begin{cases} 0, if\ len\left(Window^N_h\right) = 0, \\ 1, if\ len\left(Window^N_h\right) = 1, \\ \dfrac{\sum_{i=0}^{I} t^N_i}{n}, otherwise. \end{cases} \tag{6}$$

3.2. Construction of multi-feature STH dataset using VMD algorithm

The originally acquired multi-STH data has a strong non-linear relationship among each water areas due to various uncertainties, and the numerical changes do not have an obvious pattern, which makes it difficult for the model to learn the spatial and temporal correlation of the data from it. Therefore, this paper uses the VMD algorithm to decompose the original sequence into F finite number of Intrinsic Mode Functions (IMFs) as multiple features of STH to improve the prediction accuracy. The VMD algorithm is an adaptive, fully non-recursive mode decomposition method (Humphrey et al., 1996). Complex sequence decomposition is achieved by finding the optimal solution of the constrained variational model, which overcomes the problems of end-point effects and modal component mixing in Empirical Mode Decomposition (EMD). Extract the centre frequency of the corresponding IMF such that the mode u_f revolves around the centre frequency ω_f. The VMD algorithm flow consists of constructing the variational problem and solving the variational problem (Zhang et al., 2020). The constructed constrained variational problem is defined as

$$\min_{\{u_f\}\{\omega_f\}} \left\{ \sum_{f=1}^{F} \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_f(t) \right] e^{-j\omega_f t} \right\|_2^2 \right\}, \tag{7}$$

$$x^N(t) = \sum_{f=1}^{F} u_f(t), \tag{8}$$

where u_f is the f-th mode, ω_f is the centre frequency of the f-th mode, F is the total number of modes, $\delta(t)$ is the Dirac distribution, and the composite of each mode is the decomposed sequence x^N.

To transform the constrained variational problem into an unconstrained variational problem, penalty factors and Lagrange multipliers are introduced (Lv et al., 2021) as

$$L\left(\{u_f\}, \{\omega_f\}, \lambda\right) = \varepsilon \sum_{f=1}^{F} \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_f(t) \right] e^{-j\omega_f t} \right\|_2^2 + \left\| H(t) - \sum_{f=1}^{F} u_f(t) \right\|_2^2 + \langle \lambda(t), H(t) - \sum_{f=1}^{F} u_f(t) \rangle, \tag{9}$$

where ε is the penalty factor and λ is the Lagrange multiplier.

The unconstrained variational problem is solved using the method of alternating multiplications, and then the minima of the generalised Lagrange expression are found by alternately updating u_f, ω_f and λ. The iterative formulation is as follows (Zhao et al., 2023)

$$\hat{u}_f^{p+1}(\omega) = \frac{\hat{H}(\omega) - \sum_{i<f} \hat{u}_i^n(\omega) - \sum_{i>f} \hat{u}_i^p(\omega) + \frac{\widehat{\lambda}^p(\omega)}{2}}{1 + 2\varepsilon\left(\omega - \omega_f^p\right)^2}, \tag{10}$$

$$\omega_f^{p+1} = \frac{\int_0^\infty \omega |u_f^{p+1}(\omega)|^2 d\omega}{\int_0^\infty |u_f^{p+1}(\omega)|^2 d\omega}, \tag{11}$$

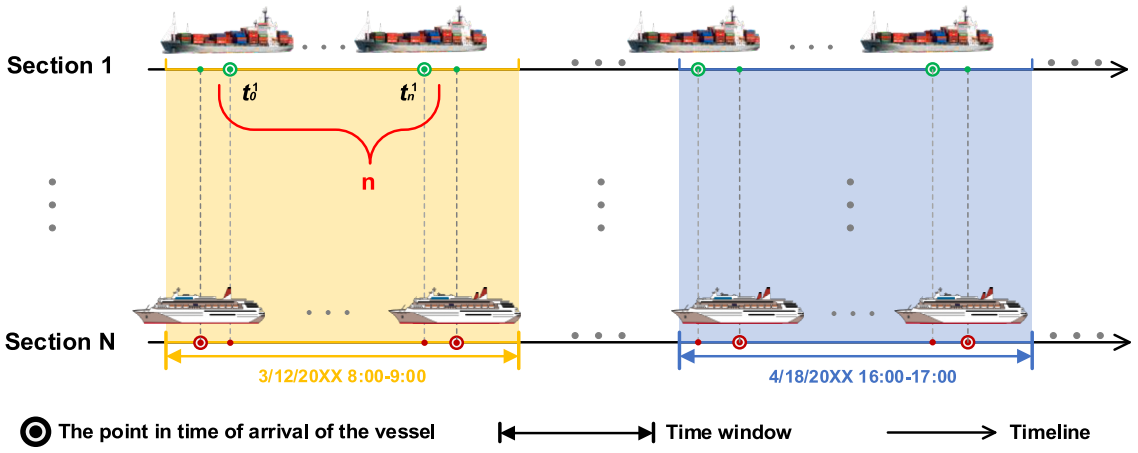


Fig. 5. Schematic diagram of multi-STH.

$$\hat{\lambda}^{p+1}(\omega)=\hat{\lambda}^{p}(\omega)+\tau\left(\hat{H}(\omega)-\sum_{f=1}^{F}\hat{u}_f^{p+1}(\omega)\right), \tag{12}$$

where $\hat{u}_f(\omega)$, $\hat{H}(\omega)$ and $\hat{\lambda}(\omega)$ are the spectra of $u_f(t)$, $H(t)$ and $\lambda(t)$, respectively. τ is the update parameter of the Lagrange multiplier, which determines whether or not to use the multiplier to enforce the constraint.

The specific steps of the VMD algorithm process are as follows:

a. Initialize u_f^1, ω_f^1 and λ^1, and make the number of iterations $p = 1$.
b. Set $p = p + 1$ and execute the loop.
c. For 1: f and all ω greater than or equal to 0, update $\hat{u}_f^{p+1}(\omega)$, ω_f^{p+1} and $\hat{\lambda}^{p+1}(\omega)$ according to the above equation.
d. For all ω greater than or equal to 0, Eq. (11) is executed.
e. Repeat steps b to d until the iteration stopping condition is satisfied, as shown in Eq. (13)

$$\sum_f \frac{\left\|\hat{u}_f^{p+1}-\hat{u}_f^{p}\right\|_2^2}{\left\|\hat{u}_f^{p}\right\|_2^2}<\rho, \tag{13}$$

where ρ is the threshold for determining the stopping of iterations.

When the whole decomposition process is finished, the STH sequence of a certain water area is decomposed into F feature sequences (IMF components) $X^N = \left(IMF_1^N, IMF_2^N, \ldots, IMF_F^N\right) \in \mathbb{R}^{H\times F}$. The VMD algorithm was used to decompose the STH sequences of all areas to obtain the multi-featured multi-STH dataset. At the h-th time unit, there exists the matrix $X_h = \left(X_h^1, X_h^2, \ldots, X_h^N\right)^T \in \mathbb{R}^{N\times F}$. The dataset can be represented as $X = (X_1, \ldots, X_h, \ldots, X_H)^T \in \mathbb{R}^{N\times F\times H}$.

3.3. Deep learning based multi-STH prediction

To enhance the interpretability of the proposed prediction framework, this section presents the construction of the inland waterway traffic network, the method for performing multi-STH prediction based on STAGCN, and the operational mechanism of each component in the model. The deep learning model primarily consists of the Spatial-Temporal Attention Mechanism (S-T AM), GCN, and CNN.

3.3.1. Inland waterway transport network

In this paper, we define an undirected graph $G = (N, E, A)$ inland river traffic network considering the characteristics of inland river that are closed and the water areas are connected to each other, which is used to construct a node-level prediction task (Jiang and Luo, 2022), where N denotes the finite set of nodes (water areas), E denotes the set of edges representing the connectivity of neighbouring nodes, and A denotes the adjacency matrix, $A \in \mathbb{R}^{N\times N}$. If nodes N_i and N_j are connected in G, the corresponding position of A_{ij} in the adjacency matrix takes the value of 1, otherwise, it is 0, as shown in Eq. (14)

$$A_{ij}=\begin{cases}1, \textit{if } N_i \textit{ connects to } N_j, i\neq j,\\ 0, \textit{otherwise},\end{cases} \tag{14}$$

where N_i and N_j denote the i-th and j-th nodes in the undirected graph, respectively. Each node in G records F measurements after VMD decomposition at the same sampling frequency, and each node generates a feature vector of length F at each unit time.

3.3.2. Spatial-temporal attention mechanism module

The emergence of AM provides more powerful learning capabilities for deep learning. It essentially mimics human vision and cognition and can help deep learning models focus on more important information. AM have been widely used in areas such as natural language processing, computer vision and time series prediction (Chorowski et al., 2015). In this section, multi-STH are subjected to attention score computation in the temporal and spatial dimensions separately, capturing the extent of time-space influence on the STH of different nodes. Unlike the common scaled dot product attention (Vaswani et al., 2017) and additive attention (Wu et al., 2021) computations. A dynamic computational method (Graves, 2016) is applied to construct an attention score function similar to a simple network (Guo et al., 2019). The advantage is that the output attention matrix can adaptively change the dynamic correlations between nodes in the spatial-temporal dimension as the inputs change and provide more trainable parameters to improve the learning ability of the model (Sukhbaatar et al., 2019). STH itself changes over time and interacts to varying degrees between different water areas over time. The input x is subjected to attention computation in the time dimension using the above method to obtain the attention matrix T. The attention weights in T are normalised and merged with the input x by means of the *softmax* function (Martins and Astudillo, 2016), as shown in Fig. 6 .The inputs are dynamically adjusted by applying the correlation of the time dimension, as shown by

$$T = V_t\cdot\sigma\left(\left(x^T U_1\right)U_2(U_3 x)+b_t\right), \tag{15}$$

$$T'_{i,j}=\frac{exp\left(T_{i,j}\right)}{\sum_{j=1}^{M} exp\left(T_{i,j}\right)}, \tag{16}$$

$$T' = softmax(T), \tag{17}$$

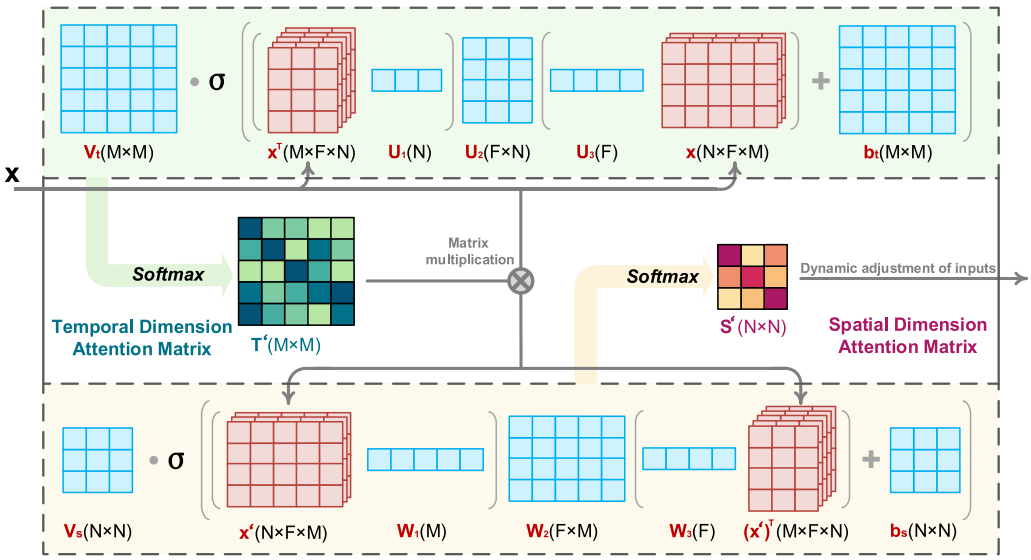


Fig. 6. Mechanisms for the operation of the S-T AM module.

$$x' = xT', \tag{18}$$

where $x = (X_{h-M+1}, X_{h-M+2}, \ldots, X_h) \in \mathbb{R}^{N\times F\times M}$ is the input, N is the number of target water areas (nodes), F is the number of features for each water areas in the input, M is the input time step, U_1, U_2, U_3, V_t, b_t are trainable parameters, $U_1 \in \mathbb{R}^N$, $U_2 \in \mathbb{R}^{F\times N}$, $U_3 \in \mathbb{R}^F$, $V_t, b_t \in \mathbb{R}^{M\times M}$. T is the output attention matrix, T' is the attention matrix through the *softmax* layer, $T, T' \in \mathbb{R}^{M\times M}$. $T'_{i,j}$ is a scalar for the correlation between moments i and j. x' is the final output, $x' \in \mathbb{R}^{N\times F\times M}$. σ is a Sigmoid activation function that provides nonlinearity to the model (Han and Moraga, 1995).

STH between different nodes all interact with each other, especially in closed water areas such as inland waterways. Ships in a given body of water can only sail upstream or downstream, and there is a strong dynamic correlation between the STHs of all neighbouring nodes. The output x' of temporal attention is taken in the spatial dimension to compute the attention matrix and the matrix is weight normalised by a *softmax* layer. The output S' of the spatial attention is applied to the graph convolution together with the input x and the adjacency matrix A to dynamically adjust the interactions between the nodes as shown in Fig. 6, through Eq. (19) to Eq. (21)

$$S = V_s \cdot \sigma\left(\left((x')^T W_1\right) W_2 (W_3 x') + b_s\right), \tag{19}$$

$$S'_{i,j} = \frac{exp\ (S_{i,j})}{\sum_{j=1}^{N} exp\ (S_{i,j})}, \tag{20}$$

$$S' = softmax(S), \tag{21}$$

where x' is the input of spatial attention, W_1, W_2, W_3, V_s, b_s are trainable parameters, $W_1 \in \mathbb{R}^M$, $W_2 \in \mathbb{R}^{F\times M}$, $W_3 \in \mathbb{R}^F$, V_s, $b_s \in \mathbb{R}^{N\times N}$. S is the output attention matrix. S' is the attention matrix through the *softmax* layer, $S, S' \in \mathbb{R}^{N\times N}$. $S'_{i,j}$ is a scalar for the correlation between moments i and j.

3.3.3. GCN

CNN are widely used in the fields of target detection, image classification, and image segmentation (Euclidean data) due to their powerful feature extraction capabilities. However, the existence of non-Euclidean data brings challenges for deep learning. For example, in chemical molecular structures, social networks, traffic networks, etc., it is challenging to complete classification or regression tasks using traditional CNNs in such graph data with topological properties (Liang et al., 2022). The emergence of GCN solves this problem by using spectral graph theory to extend the convolution operation from grid-based data to data based on graph structures and introducing filters to define GCN from the perspective of graph processing (Shuman et al., 2013). In this section the inland waterway traffic network signals are processed using spectral graph based GCN, making full use of the topology of the inland waterway traffic network to analyze the spatial features between the nodes, where the features of each node in the inland waterway traffic network are signals of the graph, as shown in Fig. 7. An undirected graph can be represented by a normalised Laplacian matrix L as:

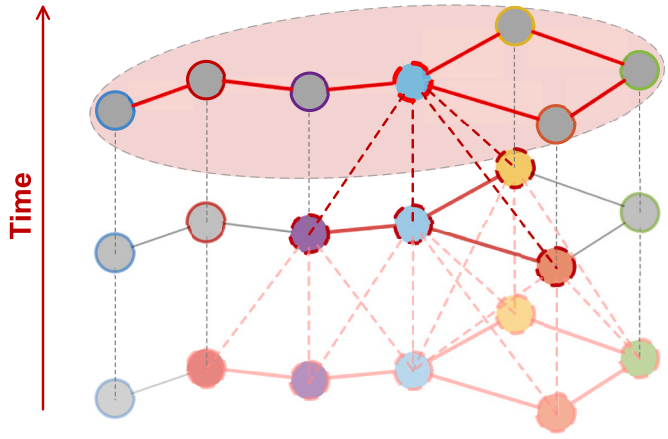


Fig. 7. Connectivity between nodes in the graph (elliptical regions indicate spatial graph convolution).

$$L = D - A, \tag{22}$$

$$L = I_N - D^{-\frac{1}{2}} A D^{-\frac{1}{2}} = \mathscr{U} \Lambda \mathscr{U}^T \in \mathbb{R}^{N\times N}, \tag{23}$$

where is the diagonal matrix of node degrees for $D \in \mathbb{R}^{N\times N}$, $D_{ii} = \sum_j (A_{i,j})$, $A \in \mathbb{R}^{N\times N}$ is the adjacency matrix, $I_N \in \mathbb{R}^{N\times N}$ is the unit matrix, $\mathscr{U} \in \mathbb{R}^{N\times N}$ is the Fourier basis, which is the matrix of eigenvectors ordered by eigenvalues, $\mathscr{U} = [u_0, u_1, \ldots, u_{N-1}]$. $\Lambda \in \mathbb{R}^{N\times N}$ is the eigenvalue diagonal matrix, $\Lambda_{ii} = \lambda_i$. Using the fact that the canonical Laplacian

matrix has the property of being symmetric semi-positive definite, it can be decomposed into $L = \mathscr{U}\Lambda\mathscr{U}^T$.

In graph signal processing, a graph signal $x \in \mathbb{R}^N$ is a certain characteristic of all nodes in an inland waterway traffic network. The graph Fourier transform of a signal x is defined as $\hat{x} = \mathscr{U}^T x$, the graph Fourier inverse transform is defined as $x = \mathscr{U}\hat{x}$, and $\hat{x}$ is the signal obtained by the graph Fourier transform. The graph Fourier transform projects the input signal into a standard orthogonal space with the basis formed by the eigenvectors of the normalised graph Laplacian operator, and the GCN can use linear operations diagonalised in the Fourier domain instead of traditional convolution operations (Henaff et al., 2015). The kernel g_θ and the input signal x are Fourier transformed and multiplied respectively, and the result of the multiplication is Fourier inverse transformed to get the result of the convolution operation. The spectrogram convolution can be defined as follows

$$g_\theta *_G x = g_\theta(L)x = g_\theta(\mathscr{U}\Lambda\mathscr{U}^T)x = \mathscr{U}g_\theta(\Lambda)\mathscr{U}^T x, \tag{24}$$

where g_θ is the kernel, $*_G$ is the graph convolution operator, and x is the input signal.

However, due to the fact that doing eigen-decomposition of the Laplacian matrix has an expensive computational cost, especially when the size of the graph becomes large. To improve the computational efficiency, the problem is solved using Chebyshev polynomials to approximate the kernel g_θ instead, i.e., $g_\theta = \sum_{k=0}^{K} \psi_k T_k(\widetilde{\Lambda})$. $\widetilde{\Lambda}$ is obtained by scaling the Laplace operator by Eq. (26). The Chebyshev polynomial is $T_k(x) = 2xT_{k-1}(x) - T_{k-2}(x)$, where $T_0(x) = 1, T_1(x) = x$. The Chebyshev spectrogram is convolved as

$$g_\theta *_G x = \mathscr{U}g_\theta(\Lambda)\mathscr{U}^T x = \mathscr{U}\left(\sum_{k=0}^{K-1}\psi_k T_k(\widetilde{\Lambda})\right)\mathscr{U}^T x = \sum_{k=0}^{K-1}\psi_k T_k(\widetilde{L})x, \tag{25}$$

$$\widetilde{\Lambda} = \frac{2}{\lambda_{max}}\Lambda - I_N, \tag{26}$$

where $\psi \in \mathbb{R}^K$ is the polynomial coefficient vector, $T_k(\widetilde{L}) \in \mathbb{R}^{N\times N}$ is a Chebyshev polynomial of order k computed on the scaled Laplace operator, λ_{max} is the maximum eigenvalue of the Laplacian matrix, and K is the order of the Chebyshev polynomial.

In this paper, in order to dynamically adjust the spatial-temporal correlation between the nodes in the input, the output S' from the spatial-temporal attention module is combined with $T_k(\widetilde{L})$. The output of the graph convolution is fed into the subsequent CNN through the *ReLU* activation function (Li and Yuan, 2017), as shown in Eq. (27)

$$ReLU(g_\theta *_G x) = ReLU\left(\sum_{k=0}^{K-1}\psi_k T_k(\widetilde{L}\odot S')x\right), \tag{27}$$

where $\odot$ is the Hadamar product.

3.3.4. CNN in the time dimension

After the Chebyshev spectral graph convolution process, the inputs have extracted the spatial correlation features between the nodes in the inland waterway traffic network. To further capture the features of the input in the time dimension for model learning, the above graph convolution output is subjected to a standard convolution operation in the time-varying direction (Shin et al., 2016), as shown in Fig. 8. To ensure that the stacking of sub-modules does not affect the model and to avoid the problem of vanishing gradients, which improves the nonlinearity of the model, residual connections are taken in the sub-modules (Szegedy et al., 2017). The final output of the sub-module H_1 is obtained by merging the input with the convolutional output of the time dimension after passing it through the convolutional operation of the 1 × 1 convolutional kernel as

$$H_1 = ReLU((\Gamma \circ ReLU(g_\theta *_G x)) + \gamma \circ x) \in \mathbb{R}^{N\times F\times M}, \tag{28}$$

where Γ is a convolution kernel in the time dimension, γ is a 1 × 1 convolution kernel, $\circ$ is a convolution operation.

3.3.5. Multi-dimensional time series prediction method for STAGCN

In this paper, the acquired multi-STH data are decomposed into

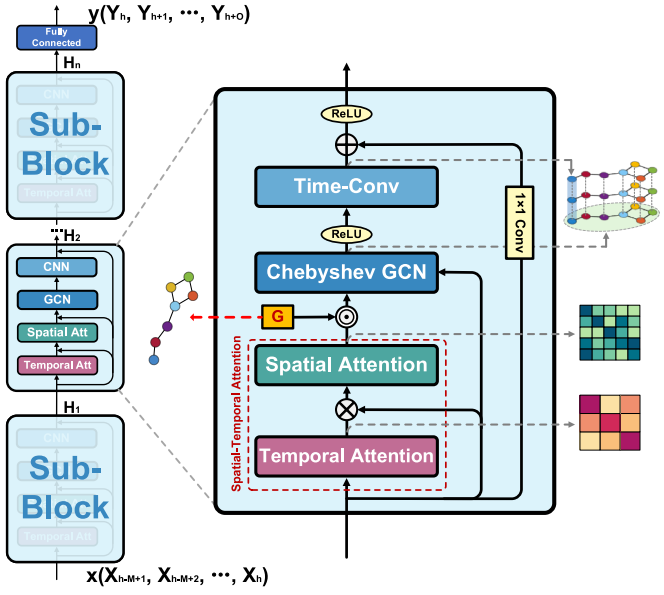


Fig. 9. STAGCN-based multi-feature temporal prediction framework map.

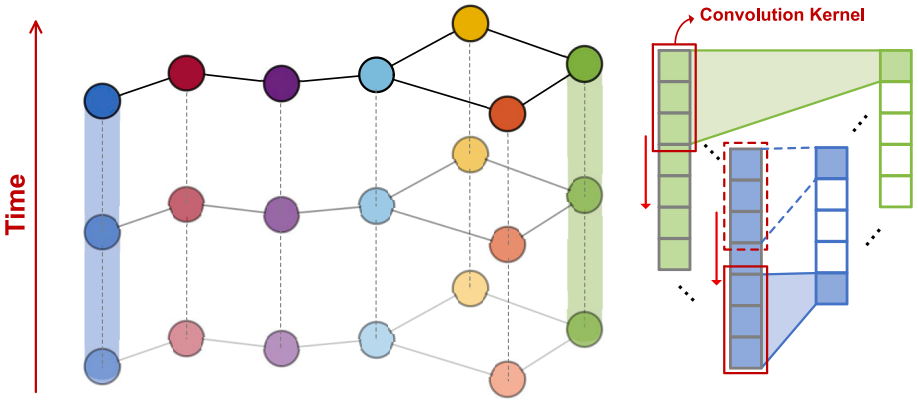


Fig. 8. Schematic of CNN in the time dimension.

multiple features input into the STAGCN model by the VMD algorithm. The constructed spatial-temporal attention module dynamically adjusts the inputs to make the model more attentive to important information based on the calculated attention score. The adjusted inputs are subsequently fed into a graph convolutional network to capture spatial features between nodes, making full use of the connectivity between nodes in the inland transport network. The standard convolution is then used for feature extraction along the temporal dimension to obtain the temporal dependence of the STH to get the output H_1 of the sub-Block. Multiple sub-Blocks are stacked to ensure a more powerful learning capability of the model and residuals are connected between the sub-modules, and the final output H_n is guaranteed to have the same dimensionality as the target value through a fully connected layer, as shown in Fig. 9. Multi-STH forecasting belongs to a classical multidimensional time series forecasting, which predicts the future O time steps based on the historical M time step observations. Therefore, the prediction method in this paper is shown as

$$\mathscr{F}([X_{h-M+1}, X_{h-M+2}, \ldots, X_h], G) = (Y_h, Y_{h+1}, \ldots, Y_{h+O}), \tag{29}$$

where $\mathscr{F}(\bullet)$ is the mapping function and G is the inland transport network.

Evaluation indicator scores are calculated to quantitatively measure the accuracy and robustness of the prediction results of the proposed method. However, the STH values contain several zeros, which would lead to a deviation of the assessment results from the actual situation if a score is calculated using Mean Absolute Percentage Error (MAPE) on the model prediction results. Therefore, in this paper, we use three other evaluation metrics commonly used for regression tasks, i.e., Root Mean Square Error (RMSE), Mean Absolute Error (MAE) and coefficient of determination (R^2). The model extension potential is measured by the scores of these metrics as

$$RMSE = \sqrt{\frac{\sum_{k=1}^{N}\left(y_{prediction} - y_{observation}\right)^2}{N}}, \tag{30}$$

$$MAE = \frac{1}{N}\sum_{k=1}^{N}\left|y_{prediction} - y_{observation}\right|, \tag{31}$$

$$R^2 = 1 - \frac{\sum_{k=1}^{N}\left(y_{prediction} - y_{observation}\right)^2}{\sum_{k=1}^{N}\left(y_{prediction} - \overline{y}_{observation}\right)^2}, \tag{32}$$

where N is the number of data, $y_{observation}$ is the historical observation and $y_{prediction}$ is the prediction. The smaller the RMSE and MAE, the better the predictive performance of the model, and the larger the R^2 indicates the better the fit of the model, i.e., the higher the prediction accuracy.

4. Case study

4.1. Data collection

In this paper, we use real AIS data to verify the feasibility and excellent performance of our proposed prediction framework. We collected historical ship AIS data from May 1, 2021, 00:00 to June 1, 2021, 00:00 for the Nanjing-Zhenjiang section of the Yangtze River in China, amounting to 24,363,072 trajectory points, as shown in Fig. 10. AIS devices in reality only have the function of receiving and transmitting signals and cannot automatically correct data information (Xing et al., 2023). During the transmission process, data can be lost or contain errors, leading to inconsistencies between the ship's historical trajectory

Table 1
Coordinates of the starting point of the virtual cross-section of the seven nodes.

Water Areas	Points	Longitude (°)	Latitude (°)
Node 1	N_α^1	118.72	32.12
	N_β^1	118.75	32.11
Node 2	N_α^2	118.83	32.17
	N_β^2	118.84	32.16
Node 3	N_α^3	118.94	32.19
	N_β^3	118.94	32.17
Node 4	N_α^4	119.16	32.25
	N_β^4	119.15	32.23
Node 5	N_α^5	119.35	32.25
	N_β^5	119.26	32.22
Node 6	N_α^6	119.36	32.22
	N_β^6	119.26	32.19
Node 7	N_α^7	119.52	32.28
	N_β^7	119.38	32.24

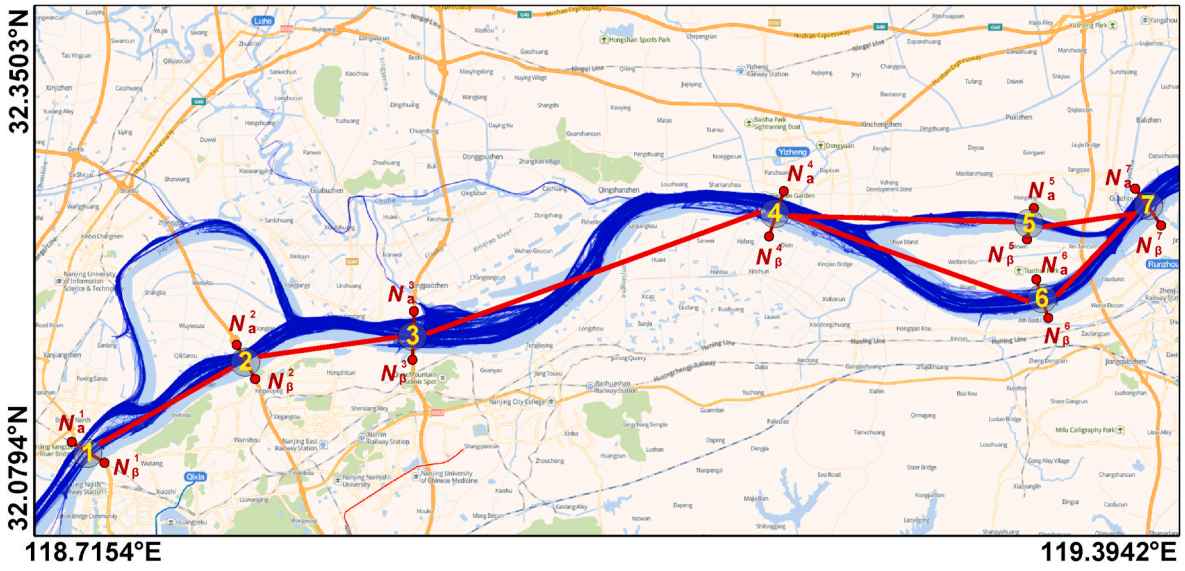


Fig. 10. Schematic of target water areas with multiple virtual cross-sections.

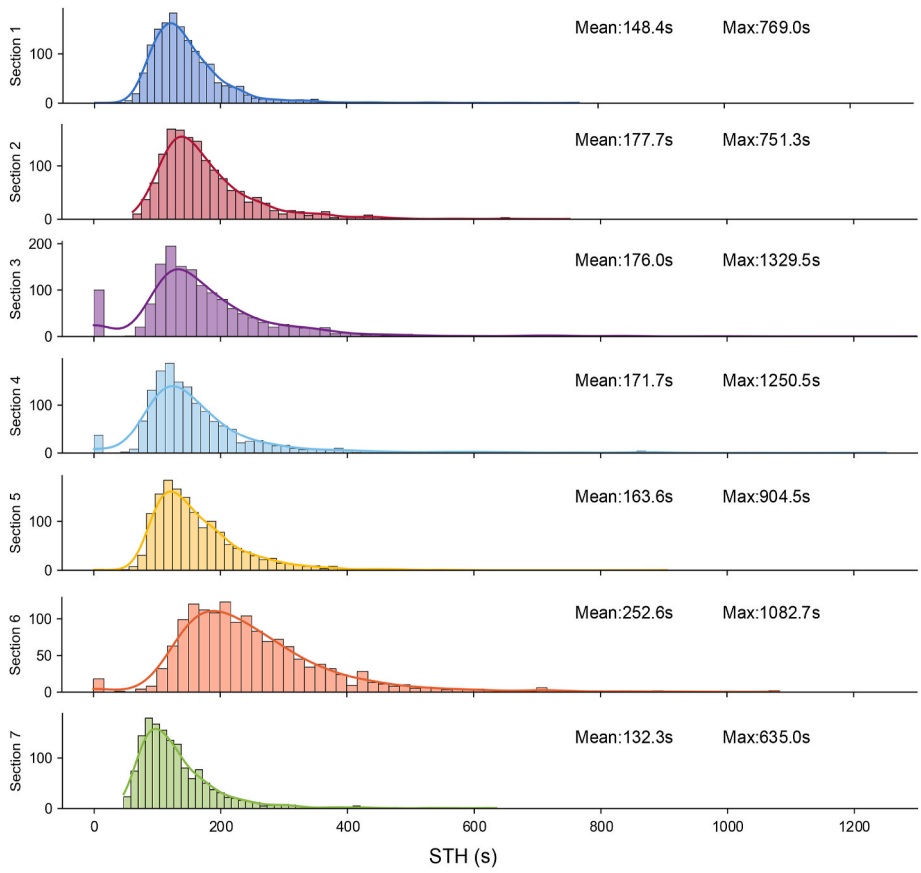


Fig. 11. Kernel density curves of the STH data samples for each cross-section, where the horizontal coordinate is the STH value, and the vertical coordinate is the number of STH values.

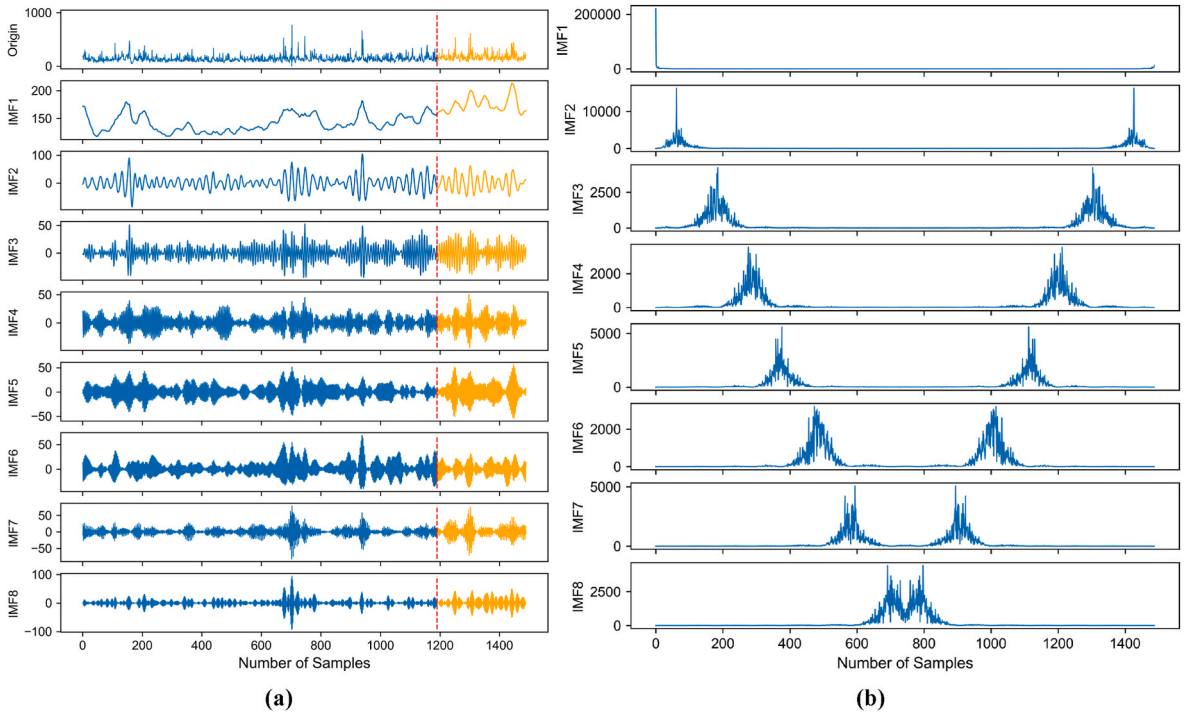


Fig. 12. Multi-feature decomposition based on VMD algorithm (taking node 1 as an example), (a) VMD decomposition results, where the blue part is the training set data and the yellow part is the testing set data, (b) spectrogram of each IMF component.

and the actual situation. Directly using the original AIS data to calculate the STH results in large deviations. Therefore, we use the AIS data preprocessing method proposed in a previous study to enhance the ship trajectory quality (Ma et al., 2024a,b).

In this paper, seven virtual cross-sections are set up to estimate the STH. These sections are all within the continuous river channel to ensure connectivity between the nodes, and the constructed inland waterway traffic network is shown in Fig. 10. The layout of the cross-sections should follow the direction of the cross-river bridge or be perpendicular to the navigation channel, avoiding areas with many ships crossing the channel vertically. Details of the virtual cross-section locations are shown in Table 1. Using the multi-STH calculation method proposed in this paper, the STH values of these nodes are calculated simultaneously at a unit time of 0.5 h to construct the Nanjing-Zhenjiang multi-STH dataset. The distribution of the generated sequences is shown in Fig. 11. The dataset is then decomposed into $F = 8$ IMF components as a multi-feature input model using the VMD algorithm, as shown in Fig. 12(a). Too few IMF components result in incomplete decomposition with differences from the original data, while excessive decomposition overburdens the model calculations. The output decomposed frequency domain data is used to determine whether modal aliasing or incomplete decomposition occurs. The more complete the decomposition of individual modes, the smaller the center frequency overlap, as shown in the center frequency distribution in Fig. 12(b) (Zhao et al., 2023). For example, the STH sequence of node 1 has no modal overlap at F = 8.

4.2. Deep learning based multi-STH prediction

In this section, the model is trained on data decomposed by the VMD algorithm. To ensure consistency across all experiments, they are conducted using the same computer, as specified in Table 2. We use the first 80% of all datasets as the training dataset and the remaining 20% as the test dataset. This allocation provides enough data for the model to learn and identify good model parameters. The Mean Square Error (MSE) is used as the model's loss function, measuring the error between the observed and predicted values. The training goal is achieved by continuously and iteratively updating the model parameters using the back-propagation algorithm (Li et al., 2023a,b). The Adam optimizer is employed to assist the loss function in continuously approximating the global minimum (Kingma and Ba, 2014).

4.2.1. Model parameter setting

The Learning Rate (LR) is a crucial hyperparameter for model training, determining whether the model's loss function can converge to a global minimum. The choice of LR has a decisive impact on the performance of the model. A small LR results in slow gradient descent, requiring more iterations and potentially causing the model to fall into a local optimum. Conversely, a large LR may cause the model to hover around the optimal solution or, in extreme cases, diverge and fail to capture features effectively. Therefore, in this paper, we experimented with different learning rates to determine the most appropriate value, as shown in Fig. 13. When LR = 0.01, the model converges faster and smoother, with the loss value eventually decreasing to the lowest and stabilizing after 60 iterations. To reduce unnecessary training time, the number of training iterations was set to 100 epochs.

The choice of batch size also influences the training results and affects the model's generalization. We tested different batch sizes (16, 32, 64, 128, 256) and compared the predictions, as shown in Fig. 14(a). The prediction accuracy for batch sizes 16 and 32 was similar; however, smaller batch sizes reduce the model's running speed and increase computational cost. Therefore, a batch size of 32 was used in all experiments.

Table 2
Experimental environment configuration.

CPU	Operating system	Python
Intel Core i5-13490 F	Windows 10 64-bit	3.11
GPU		Pytorch
GTX 1660 SUPER		2.0

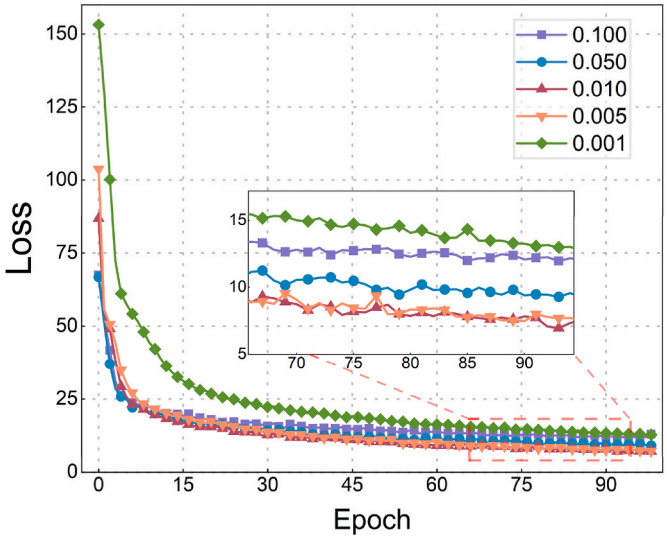


Fig. 13. Training loss of the model at different learning rates.

In time series forecasting, the samples of each input model form a mapping function to the target value. However, the historical information contained in the samples at different time steps (length of the input samples) has varying time-series associations with the target value. We selected different timesteps ∈ [2,4,6,8,10,12] for comparison experiments. As shown in Fig. 14(b), the best prediction accuracy of the model is achieved with a timestep of 10.

In deep learning, the training model is continuously updated using the backpropagation algorithm to form a mapping relationship between inputs and outputs for each parameter (weights and biases) of the model. The structure and scale of the model directly determine the accuracy and robustness of the prediction, with each component affecting the prediction results to varying degrees. We performed several comparison tests of the important parameters in the model to construct a suitable model structure, as shown in Fig. 14(c–f).

The highest prediction accuracy was achieved when the Chebyshev-based graph convolution output had a feature dimension of 32, and the time dimension convolution output had a feature dimension (number of convolution kernels) of 128. The convolution kernels in the time dimension had a dimension of 1×3 and slid along the direction of time change. The model performed best when the Chebyshev polynomial had the number of terms $K = 1$ and the number of blocks = 1. This is mainly due to the small number of nodes in the dataset, meaning the parameter size of the model does not need to be too large to form a good mapping relationship.

The hyperparameter configuration of the model in this paper is shown in Table 3, which has a total of 14,607 trainable parameters.

4.2.2. Qualitative evaluation of the proposed model

K-fold cross-validation is a commonly used method for model evaluation and selection in deep learning and machine learning. It is also a strategy for partitioning datasets (Zhang et al., 2024). The aim is to avoid overfitting due to the limited amount of data and to improve the generalization ability of the model, which is particularly advantageous for small datasets. The dataset is partitioned into k equal parts; each time, one part is used as the test dataset and the rest as the training dataset. This cross-validation is repeated k times, and the scoring results from these k iterations are averaged to determine the final

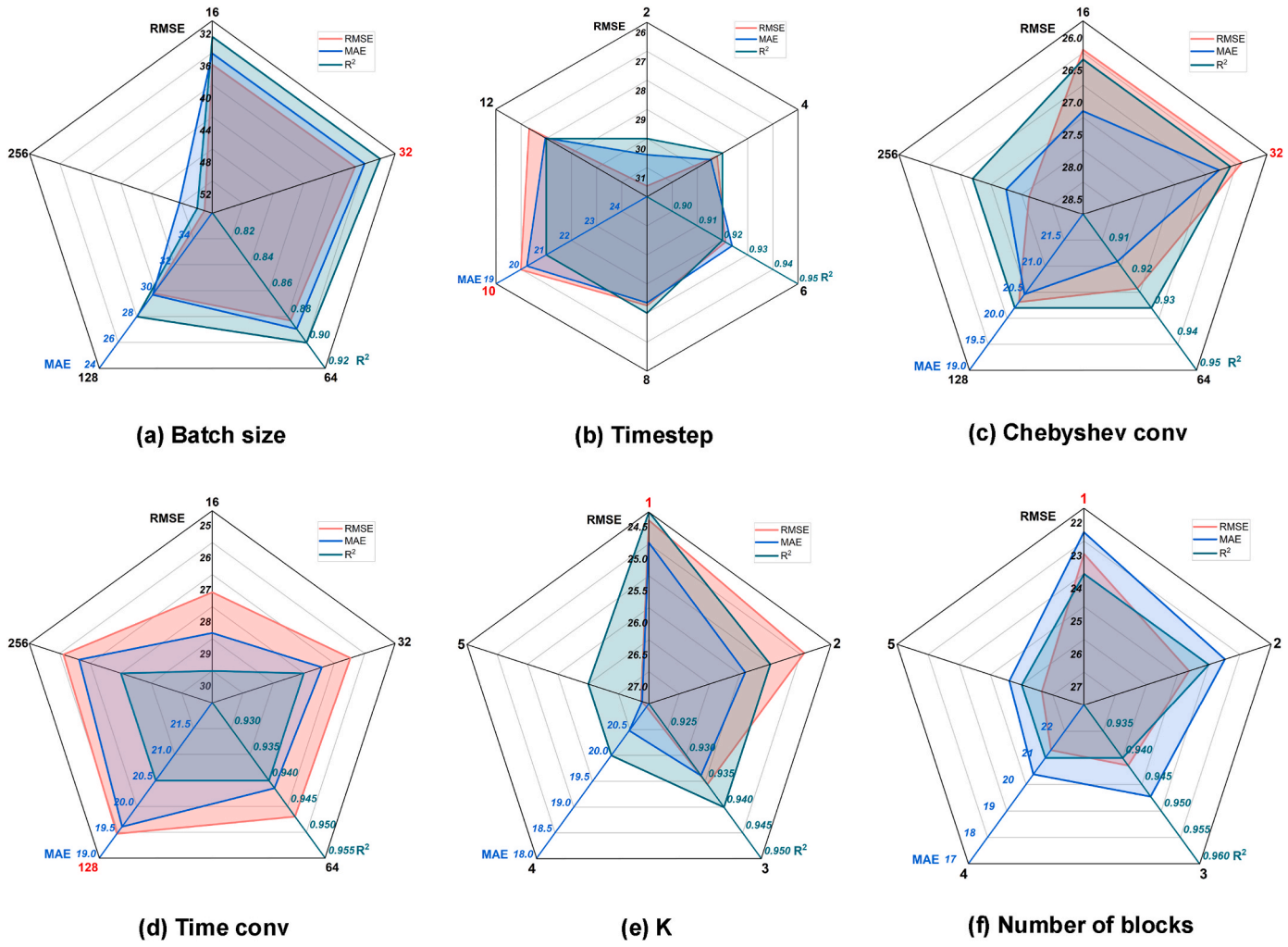


Fig. 14. Comparison of prediction scores for different model parameters.

Table 3
The model characteristics and optimal hyperparameters.

Model	Input variables	Output variable	Number of blocks
VMD-STAGCN	7*10*8 (node*timestep*feature)	7*1	1

Chebyshev conv	Time conv	K	Batch size
32	128	1	32

Learning rate	Epoch	Optimizer	Loss function
0.01	100	Adam	MSE

generalization error, as shown in Fig. 15.

In this paper, a 5-fold cross-validation strategy is used, meaning the ratio of the training dataset to the test dataset is kept at 4:1. The scores of the five trained models are averaged to determine the overall performance of the model when extended to new data, as shown in Table 4. The results indicate that the prediction method proposed in this paper demonstrates superior performance. However, it is important to note that the lower score for Split 4 is due to the presence of more noise in the dataset, caused by various uncertainties assigned to the test dataset, highlighting the model's limitations when facing unconventional variations in STH.

Based on the series of validation experiments mentioned above, the performance of the model on the validation set is finally derived. As

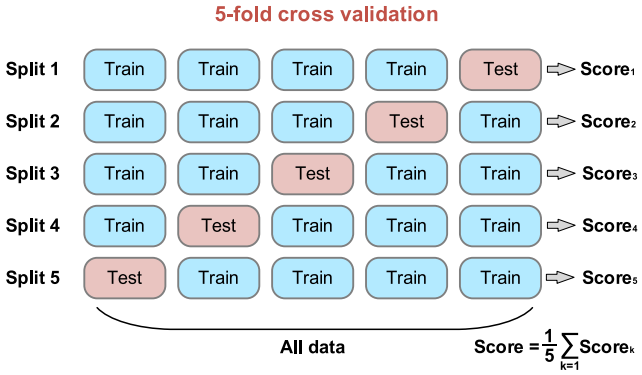


Fig. 15. 5-Fold cross-validation model evaluation.

Table 4
5-fold cross-validation of model predictions.

	$Score_1$	$Score_2$	$Score_3$	$Score_4$	$Score_5$	Score
RMSE(s)	23.39	18.85	18.37	57.54	24.88	**28.61**
MAE(s)	17.73	13.53	13.63	29.23	16.53	**18.13**
R^2	0.95	0.95	0.95	0.86	0.94	**0.93**

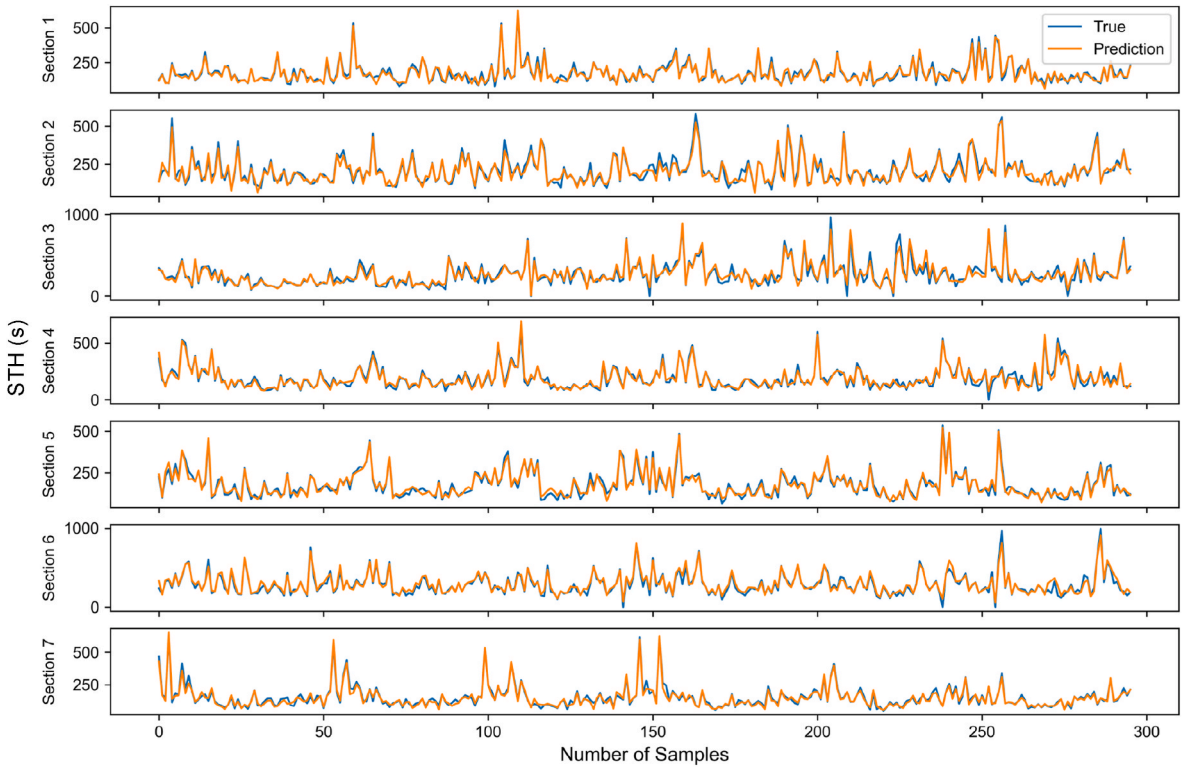


Fig. 16. Predictive performance of the model in the test dataset.

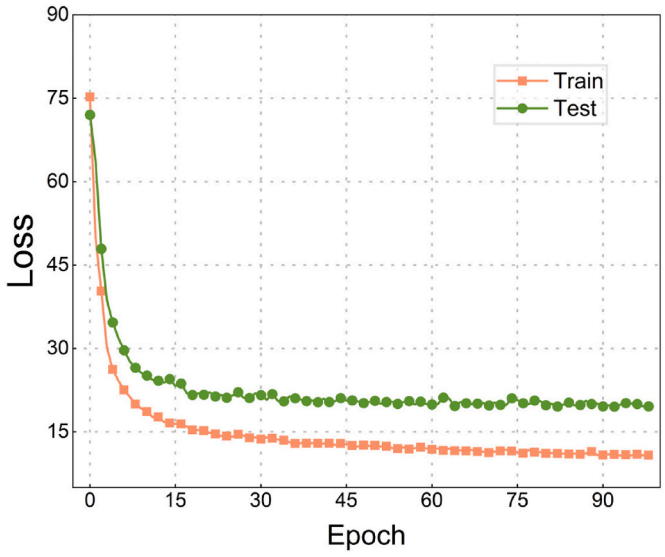


Fig. 17. Training loss and test loss of the model with optimal hyperparameters.

shown in Fig. 16, the model is able to fit the complex multi-STH data, demonstrating excellent feature extraction capabilities. The loss curves of the model were also output to judge the training of the model according to the optimal hyperparameters in Table 3, as shown in Fig. 17. The decrease in both training loss and test loss stabilised after the 60th epoch, indicating that the balance between model performance and training data converged and the model reached a state of good fit. To avoid overfitting, model training can be terminated early if no further improvement in performance occurs after this balance point, an operation that also saves unnecessary training time. In addition, the 5-fold cross-validation taken in the previous section can improve the generalisation ability of the model and effectively prevent the problem of model overfitting.

4.3. Validation and comparison

To verify the validity and excellent performance of this prediction method, we compared the proposed method with classical machine learning and deep learning baseline models. The baseline models included Historical Average (HA), Support Vector Regression (SVR), LSTM, Gate Recurrent Unit neural network (GRU), GCN, and Spatial-Temporal Graph Convolutional Network (STGCN), as shown in Table 5. All these baseline models struggle to learn key information from multi-STH data, whereas our prediction method is far superior to them. This is due to the small structure and scale of the baseline models, which cannot effectively extract timing features from the complex multi-STH

Table 5
Comparison of the performance of different prediction models.

Evaluation Metrics	Model								
	HA	SVR	LSTM	GRU	GCN	STGCN	STAGCN	VMD-STGCN	VMD-STAGCN
RMSE(s)	130.92	124.19	120.54	119.19	115.65	97.01	96.31	45.31	**28.61**
MAE(s)	87.69	76.55	75.93	75.08	72.66	63.10	62.48	29.33	**18.13**
R^2	*	0.01	0.01	0.01	0.02	0.15	0.16	0.88	**0.93**

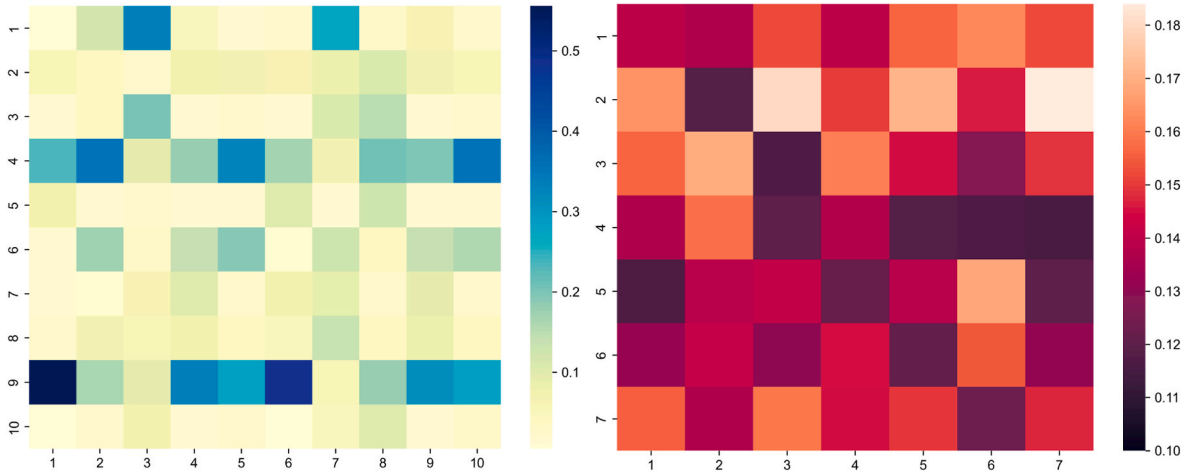


Fig. 18. Attention matrix output from the spatial-temporal attention module (temporal attention matrix on the left, spatial attention matrix on the right).

data. However, the STAGCN model performs exceptionally well because it not only considers the dynamic correlation of STH between nodes but also captures feature information in the time dimension.

We also conducted a set of ablation experiments to validate the importance of each component in the model, as shown in Table 5. When the VMD algorithm is introduced, the model's performance substantially improves. Transforming time-domain data to frequency-domain data provides the model with rich multi-STH feature information, enhancing its ability to detect spatial-temporal variations. Additionally, removing the spatial-temporal attention module results in decreased model performance, regardless of the presence of the VMD algorithm.

By observing the four sets of experiments, STGCN, STAGCN, VMD-STGCN, and VMD-STAGCN, it can be seen that the addition of the spatial-temporal attention module to the model will result in a better prediction performance with the same input data. The evaluation metrics RMSE and MAE were reduced by 36.9% and 38.2 %, respectively, and R^2 was improved by 5.7%. In addition, to visualize the operation mechanism of the spatial-temporal attention module, a random sample was used to output the temporal and spatial attention matrices. As shown in Fig. 18, the output matrices conduct the attention scoring from the temporal and spatial dimensions, according to the importance of the input information. Thus, the spatio-temporal attention module can effectively capture the complex dependencies from the data.

In this paper, the residual analysis of the proposed prediction method is carried out to examine the reasonableness and reliability of the method. In this part, residuals are defined as the difference between observed and predicted values. There may be differences in the model's predictive performance in the face of STH in different water areas. To verify the limitations of the model on different STH datasets, we plot the output of the model on the test set separately for different cross-sectional STH violin residuals, as shown in Fig. 19(a). Most of the residuals for each cross-section are spread around 0, with only a very small amount of data showing larger values, due to unavoidable noise in the data. Therefore, the method proposed in this paper can make good predictions in multiple water areas simultaneously. All the outputs are plotted against the residual distribution, as shown in Fig. 19(b), where most of the residual values are distributed on both sides of 0 and the mean value is only −0.52s. This indicates that VMD-STAGCN has successfully captured the spatial-temporal correlation of inland waterway traffic networks and extracted the complex STH relationship between different research areas.

To further evaluate the generalization potential of the method proposed in this paper and its application value in intelligent transport systems, we explore the multi-step prediction performance of the model

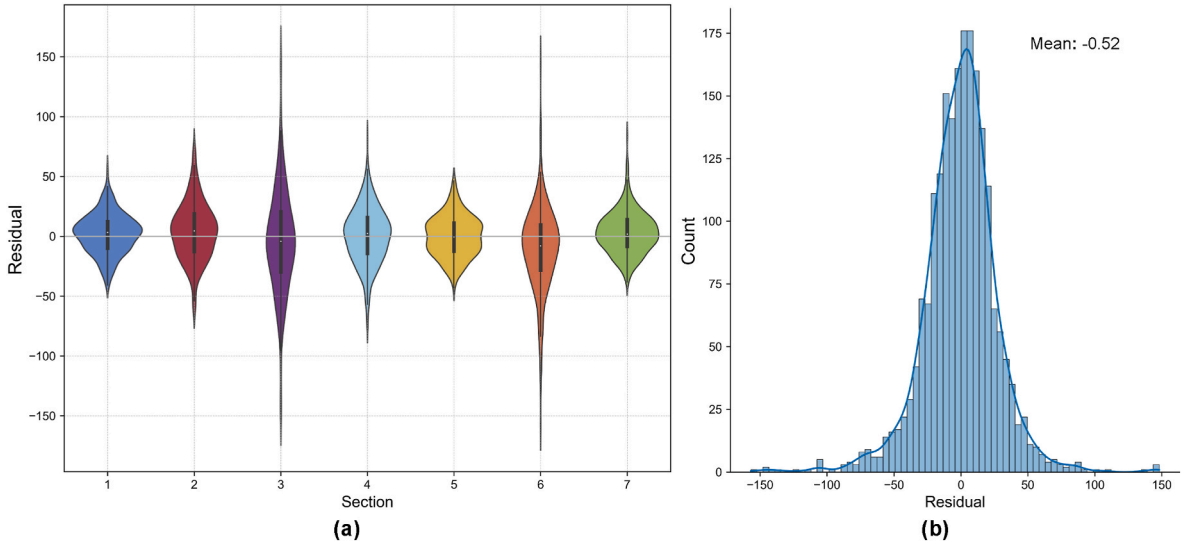


Fig. 19. Predictive performance of the model in the test set ((a) shows the violin plot of the residuals of the prediction results based on each cross-section, (b) shows the distribution of the residuals of all the prediction results).

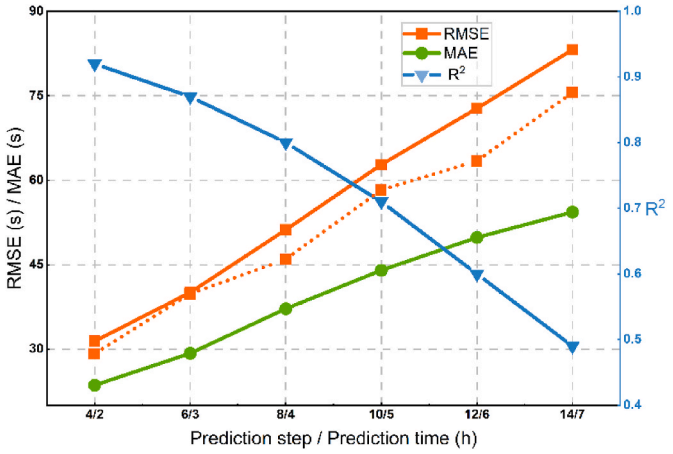


Fig. 20. Changes in model performance at different prediction step sizes.

(long-term prediction) in this section. The longer the model predicts while maintaining stable performance, the higher its application value. However, theoretically, as the prediction time horizon extends, the prediction difficulty increases dramatically, potentially limiting the model's performance (Liu et al., 2022).

As shown in Fig. 20, the model maintains high prediction performance under multi-step prediction tasks, even though prediction accuracy tends to decrease as the prediction time increases. Additionally, increasing the model input step size (input time step increased from 10 to 15) helps mitigate performance degradation, as indicated by the RMSE dashed line in Fig. 20. This is caused by the fact that the model needs inputs that contain more historical information to estimate the long-term STH.

5. Conclusion and future work

This paper proposes a deep learning method to predict multi-STH in inland waterways. The proposed framework consists of (1) estimating the STH: Estimate the STH of historical multiple water areas through the proposed set of cross-sectional ships and recount the STH sequences of these different cross-sections to ensure they are numerically continuous on a uniform timeline, (2) transforming STH Sequences: Transform the STH sequences into several IMF components as multi-features using the VMD algorithm, and (3) constructing a Traffic Network: Construct an inland waterway traffic network and train a graph-based deep learning model to dynamically capture the spatial and temporal characteristics of STH variations between each water areas.

To validate the feasibility of the proposed method, a large amount of historical AIS data was collected from the Nanjing-Zhenjiang section of the Yangtze River. The connectivity among the waterways was used to construct an inland waterway traffic network. The proposed method was subsequently compared with existing methods, and the validation results were analyzed from multiple perspectives to assess the method's potential for extension.

The main conclusions can be summarized as follows:

- A graph-based learning framework combined with a spatial-temporal attention mechanism module for predicting multi-STH is promising and a suitable choice. This approach efficiently processes data with a graphical structure while dynamically capturing the correlation between STH in the spatial-temporal dimension.
- Validation shows that using the VMD algorithm to decompose STH sequences into multiple features effectively improves the model's prediction performance.
- Comprehensive validation of real data and comparison experiments with existing baseline methods demonstrate that our method outperforms other baseline methods in prediction. It also shows high accuracy and robustness in predicting complex multidimensional STH. Additionally, the model exhibits excellent performance in multi-step prediction, offering great potential for method generalization.
- The ability to predict STH simultaneously, accurately, and stably in multiple inland waterways enables effective determination of ship arrival patterns and correlations across a wide range of waterways.
- Accurate and stable prediction of multi-STH may prevent maritime accidents in advance caused by ship congestion on inland waterways and may provide a solid basis for ship scheduling strategies and planning. This study may have important theoretical significance for studying maritime traffic.

This study applies not only to multiple contiguous segments of an inland river but also to areas with multiple correlated locks and bridge water areas. There is still room for improvement in multi-STH prediction to enhance its accuracy and potential for real applications. STH is subject to non-linear variations in different waters due to various abnormal events, including management rules of different segments, navigation regulations, and weather conditions. Additionally, the presence of outliers in the STH sequence can affect prediction accuracy. Ideally, noise influence should be minimized when collecting STH data.

Future work should expand the study range, for example, by applying STH prediction to busy coastal waterways. Integrating large-scale maritime shipping networks into the model will expand its application potential and demonstrate the significance of STH predictions. This integration will also support the establishment of traffic management schemes and ship scheduling decisions to improve shipping efficiency.

CRediT authorship contribution statement

Quandang Ma: Writing – review & editing, Writing – original draft, Supervision, Software, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Xu Du:** Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Methodology, Investigation, Formal analysis, Data curation. **Mingyang Zhang:** Writing – review & editing, Writing – original draft, Software, Resources, Methodology, Investigation, Formal analysis, Conceptualization. **Hongdong Wang:** Writing – review & editing, Visualization, Methodology. **Xiao Lang:** Writing – review & editing, Writing – original draft, Visualization, Methodology. **Wengang Mao:** Writing – review & editing, Visualization, Investigation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

All authors acknowledge the support from their employers. Dr. Quandang Ma and Mr. Xu Du acknowledge that research funding received from the Natural Science Foundation of Hubei Province (20221J0089).

References

Alzubaidi, L., Zhang, J., Humaidi, A.J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., et al., 2021. Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. J. big Data 8, 1–74.
Belhadi, A., Djenouri, Y., Djenouri, D., Lin, J.C.W., 2020. A recurrent neural network for urban long-term traffic flow forecasting. Appl. Intell. 50, 3252–3265.
Bengio, Y., Simard, P., Frasconi, P., 1994. Learning long-term dependencies with gradient descent is difficult. IEEE Trans. Neural Network. 5 (2), 157–166.

Cao, S., Wu, L., Wu, J., Wu, D., Li, Q., 2022. A spatio-temporal sequence-to-sequence network for traffic flow prediction. Inf. Sci. 610, 185–203.
Chen, X., Sun, L., 2021. Bayesian temporal factorization for multidimensional time series prediction. IEEE Trans. Pattern Anal. Mach. Intell. 44 (9), 4659–4673.
Chen, Y., Liu, Z., Zhang, M., Yu, H., Fu, X., Xiao, Z., 2024. Dynamics collision risk evaluation and early alert in busy waters: a spatial-temporal coupling approach. Ocean Eng. 300, 117315.
Chorowski, J.K., Bahdanau, D., Serdyuk, D., Cho, K., Bengio, Y., 2015. Attention-based models for speech recognition. Adv. Neural Inf. Process. Syst. 28.
Cossu, A., Carta, A., Lomonaco, V., Bacciu, D., 2021. Continual learning for recurrent neural networks: an empirical evaluation. Neural Network. 143, 607–627.
Deng, Y., Sheng, D., Liu, B., 2021. Managing ship lock congestion in an inland waterway: a bottleneck model with a service time window. Transport Pol. 112, 142–161.
Gao, R., Li, R., Hu, M., Suganthan, P.N., Yuen, K.F., 2023. Dynamic ensemble deep echo state network for significant wave height forecasting. Appl. Energy 329, 120261.
Graves, A., 2016. Adaptive computation time for recurrent neural networks. arXiv preprint arXiv:1603.08983.
Guo, S., Lin, Y., Feng, N., Song, C., Wan, H., 2019. Attention based spatial-temporal graph convolutional networks for traffic flow forecasting. Proc. AAAI Conf. Artif. Intell. 33 (1), 922–929.
Han, J., Moraga, C., 1995. The influence of the sigmoid function parameters on the speed of backpropagation learning. In: International Workshop on Artificial Neural Networks. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 195–201.
Hao, S., Lee, D.-H., Zhao, D., 2019. Sequence to sequence learning with attention mechanism for short-term passenger flow prediction in large-scale metro system. Transport. Res. C Emerg. Technol. 107, 287–300.
Henaff, M., Bruna, J., LeCun, Y., 2015. Deep convolutional networks on graph-structured data. arXiv preprint arXiv:1506.05163.
Hinton, G.E., Salakhutdinov, R.R., 2006. Reducing the dimensionality of data with neural networks. Sci. Technol. Humanit. 313 (5786), 504–507.
Humphrey, W., Dalke, A., Schulten, K., 1996. VMD: visual molecular dynamics. J. Mol. Graph. 14 (1), 33–38.
Jiang, W., Luo, J., 2022. Graph neural network for traffic forecasting: a survey. Expert Syst. Appl. 207, 117921.
Kamilaris, A., Prenafeta-Boldú, F.X., 2018. Deep learning in agriculture: a survey. Comput. Electron. Agric. 147, 70–90.
Kang, L., Meng, Q., Liu, Q., 2018. Fundamental diagram of ship traffic in the Singapore Strait. Ocean Eng. 147, 340–354.
Kim, T.Y., Cho, S.B., 2019. Predicting residential energy consumption using CNN-LSTM neural networks. Energy 182, 72–81.
Kingma, D.P., Ba, J., 2014. Adam: a method for stochastic optimization. arXiv preprint arXiv:1412.6980.
Kujala, P., Hänninen, M., Arola, T., Ylitalo, J., 2009. Analysis of the marine traffic safety in the Gulf of Finland. Reliab. Eng. Syst. Saf. 94 (8), 1349–1357.
Lei, J., Lei, L., Xiumin, C., Wei, H., Xinglong, L., Cong, L., 2021. Automatic identification system data-driven model for analysis of ship domain near bridge-waters. J. Navig. 74 (6), 1284–1304.
Li, Y., Yuan, Y., 2017. Convergence analysis of two-layer neural networks with relu activation. Adv. Neural Inf. Process. Syst. 30.
Li, Y., Li, Z., Mei, Q., Wang, P., Hu, W., Wang, Z., et al., 2023a. Research on multi-port ship traffic prediction method based on spatiotemporal graph neural networks. J. Mar. Sci. Eng. 11 (7), 1379.
Li, Y., Liang, M., Li, H., Yang, Z., Du, L., Chen, Z., 2023b. Deep learning-powered vessel traffic flow prediction with spatial-temporal attributes and similarity grouping. Eng. Appl. Artif. Intell. 126, 107012.
Liang, M., Liu, R.W., Li, S., et al., 2021. An unsupervised learning method with convolutional auto-encoder for vessel trajectory similarity computation. Ocean Eng. 225, 108803.
Liang, M., Liu, R.W., Zhan, Y., Li, H., Zhu, F., Wang, F.Y., 2022. Fine-grained vessel traffic flow prediction with a spatio-temporal multigraph convolutional network. IEEE Trans. Intell. Transport. Syst. 23 (12), 23694–23707.
Liu, C., Musharraf, M., Li, F., Kujala, P., 2022a. A data mining method for automatic identification and analysis of icebreaker assistance operation in ice-covered waters. Ocean Eng. 266, 112914.
Liu, C., Kulkarni, K., Suominen, M., Kujala, P., Musharraf, M., 2024. On the data-driven investigation of factors affecting the need for icebreaker assistance in ice-covered waters. Cold Reg. Sci. Technol. 221, 104173.
Liu, R.W., Liang, M., Nie, J., Yuan, Y., Xiong, Z., Yu, H., Guizani, N., 2022b. STMGCN: mobile edge computing-empowered vessel trajectory prediction using spatio-temporal multigraph convolutional network. IEEE Trans. Ind. Inf. 18 (11), 7977–7987.
Liu, R.W., Zheng, W., Liang, M., 2024. Spatio-temporal multi-graph transformer network for joint prediction of multiple vessel trajectories. Eng. Appl. Artif. Intell. 129, 107625.
Lv, L., Wu, Z., Zhang, J., Zhang, L., Tan, Z., Tian, Z., 2021. A VMD and LSTM based hybrid model of load forecasting for power grid security. IEEE Trans. Ind. Inf. 18 (9), 6474–6482.
Ma, C., Zhao, Y., Dai, G., Xu, X., Wong, S.C., 2022. A novel STFSA-CNN-GRU hybrid model for short-term traffic speed prediction. IEEE Trans. Intell. Transport. Syst. 24 (4), 3728–3737.
Ma, Q., Du, X., Liu, C., Jiang, Y., Liu, Z., Xiao, Z., Zhang, M., 2024a. A hybrid deep learning method for the prediction of ship time headway using automatic identification system data. Eng. Appl. Artif. Intell. 133, 108172.
Ma, Q., Tang, H., Liu, C., Zhang, M., Zhang, D., Liu, Z., Zhang, L., 2024b. A big data analytics method for the evaluation of maritime traffic safety using automatic identification system data. Ocean Coast Manag. 251, 107077.
Martins, A., Astudillo, R., 2016. From softmax to sparsemax: a sparse model of attention and multi-label classification. In: International Conference on Machine Learning. PMLR, pp. 1614–1623.
Semenoglou, A.A., Spiliotis, E., Assimakopoulos, V., 2023. Data augmentation for univariate time series forecasting with neural networks. Pattern Recogn. 134, 109132.
Shi, S., Zhang, D., Su, Y., Zhang, M., Sun, M., Yao, H., 2019. Risk factors analysis modeling for ship collision accident in inland river based on text mining. In: 2019 5th International Conference on Transportation Information and Safety (ICTIS). IEEE, pp. 602–607.
Shin, H.C., Roth, H.R., Gao, M., Lu, L., Xu, Z., Nogues, I., et al., 2016. Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. IEEE Trans. Med. Imag. 35 (5), 1285–1298.
Shuman, D.I., Narang, S.K., Frossard, P., Ortega, A., Vandergheynst, P., 2013. The emerging field of signal processing on graphs: extending high-dimensional data analysis to networks and other irregular domains. IEEE Signal Process. Mag. 30 (3), 83–98.
Su, Z., Wu, C., Xiao, Y., He, H., 2022. Study on the prediction model of accidents and incidents of cruise ship operation based on machine learning. Ocean Eng. 260, 111954.
Sui, Z., Wang, S., Wen, Y., Cheng, X., Theotokatos, G., 2024. Multi-state ship traffic flow analysis using data-driven method and visibility graph. Ocean Eng. 298, 117087.
Sukhbaatar, S., Grave, E., Bojanowski, P., Joulin, A., 2019. Adaptive Attention Span in Transformers arXiv preprint arXiv:1905.07799.
Sun, B., Sun, T., Zhang, Y., Jiao, P., 2020. Urban traffic flow online prediction based on multi-component attention mechanism. IET Intell. Transp. Syst. 14 (10), 1249–1258.
Suo, Y., Chen, W., Claramunt, C., Yang, S., 2020. A ship trajectory prediction framework based on a recurrent neural network. Sensors 20 (18), 5133.
Szegedy, C., Ioffe, S., Vanhoucke, V., Alemi, A., 2017. Inception-v4, inception-resnet and the impact of residual connections on learning. Proc. AAAI Conf. Artif. Intell. 31 (1).
Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., et al., 2017. Attention is all you need. Adv. Neural Inf. Process. Syst. 30.
Wang, D., Meng, Y., Chen, S., Xie, C., Liu, Z., 2021. A hybrid model for vessel traffic flow prediction based on wavelet and prophet. J. Mar. Sci. Eng. 9 (11), 1231.
Wang, S., Li, Y., Zhang, Z., Xing, H., 2023. Big data driven vessel trajectory prediction based on sparse multi-graph convolutional hybrid network with spatio-temporal awareness. Ocean Eng. 287, 115695.
Wang, Y., Liu, J., Liu, R.W., Liu, Y., Yuan, Z., 2023. Data-driven methods for detection of abnormal ship behavior: progress and trends. Ocean Eng. 271, 113673.
Wu, C., Wu, F., Qi, T., Huang, Y., Xie, X., 2021. Fastformer: additive attention can be all you need. arXiv preprint arXiv:2108.09084.
Xiao, F., Ligteringen, H., Van Gulijk, C., Ale, B., 2015. Comparison study on AIS data of ship traffic behavior. Ocean Eng. 95, 84–93.
Xiao, H., Zhao, Y., Zhang, H., 2022. Predict vessel traffic with weather conditions based on multimodal deep learning. J. Mar. Sci. Eng. 11 (1), 39.
Xing, W., Wang, J., Zhou, K., Li, H., Li, Y., Yang, Z., 2023. A hierarchical methodology for vessel traffic flow prediction using Bayesian tensor decomposition and similarity grouping. Ocean Eng. 286, 115687.
Xu, D.W., Wang, Y.D., Jia, L.M., Qin, Y., Dong, H.H., 2017. Real-time road traffic state prediction based on ARIMA and Kalman filter. Frontiers of Information Technology & Electronic Engineering 18, 287–302.
Yang, D., Li, S., Peng, Z., Wang, P., Wang, J., Yang, H., 2019. MF-CNN: traffic flow prediction using convolutional neural network and multi-features fusion. IEICE Trans. Info Syst. 102 (8), 1526–1536.
Yu, B., Yin, H., Zhu, Z., 2017. Spatio-temporal graph convolutional networks: a deep learning framework for traffic forecasting. arXiv preprint arXiv:1709.04875.
Yu, Y., Si, X., Hu, C., Zhang, J., 2019. A review of recurrent neural networks: LSTM cells and network architectures. Neural Comput. 31 (7), 1235–1270.
Zhang, L., Meng, Q., Fwa, T.F., 2019. Big AIS data based spatial-temporal analyses of ship traffic in Singapore port waters. Transport. Res. E Logist. Transport. Rev. 129, 287–304.
Zhang, W., Yu, Y., Qi, Y., Shu, F., Wang, Y., 2019. Short-term traffic flow prediction based on spatio-temporal analysis and CNN deep learning. Transportmetrica: Transport. Sci. 15 (2), 1688–1711.
Zhang, S., Tong, H., Xu, J., Maciejewski, R., 2019. Graph convolutional networks: a comprehensive review. Computational Social Networks 6 (1), 1–23.
Zhang, Y., Zhao, Y., Kong, C., Chen, B., 2020. A new prediction method based on VMD-PRBF-ARMA-E model considering wind speed characteristic. Energy Convers. Manag. 203, 112254.
Zhang, M., Conti, F., Le Sourne, H., Vassalos, D., Kujala, P., Lindroth, D., Hirdaris, S., 2021. A method for the direct assessment of ship collision damage and flooding risk in real conditions. Ocean Eng. 237, 109605.
Zhang, M., Zhang, D., Fu, S., Kujala, P., Hirdaris, S., 2022. A predictive analytics method for maritime traffic flow complexity estimation in inland waterways. Reliab. Eng. Syst. Saf. 220, 108317.
Zhang, M., Tsoulakos, N., Kujala, P., Hirdaris, S., 2024. A deep learning method for the prediction of ship fuel consumption in real operational conditions. Eng. Appl. Artif. Intell. 130, 107425.
Zhao, L., Li, Z., Qu, L., Zhang, J., Teng, B., 2023. A hybrid VMD-LSTM/GRU model to predict non-stationary and irregular waves on the east coast of China. Ocean Eng. 276, 114136.
Zhou, X., Liu, Z., Wang, F., Xie, Y., Zhang, X., 2020. Using deep learning to forecast maritime vessel flows. Sensors 20 (6), 1761.