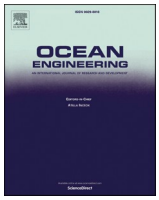
Ocean Engineering 286 (2023) 115687

Contents lists available at ScienceDirect

Ocean Engineering

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



A hierarchical methodology for vessel traffic flow prediction using Bayesian tensor decomposition and similarity grouping

Wenbin Xing [a,1], Jingbo Wang [a,1], Kaiwen Zhou [a,1], Huanhuan Li [b], Yan Li [c,*], Zaili Yang [b,**]

[a] School of Computer Science and Artificial Intelligence, Wuhan University of Technology, Wuhan, China
[b] Liverpool Logistics, Offshore and Marine (LOOM) Research Institute, Liverpool John Moores University, Liverpool, UK
[c] State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan, China

ARTICLE INFO

Handling Editor: Prof. A.I. Incecik

Keywords:
Vessel traffic flow prediction
Navigation safety
Tensor decomposition
Automatic identification system (AIS)
Similarity grouping

ABSTRACT

Accurate vessel traffic flow (VTF) prediction can enhance navigation safety and economic efficiency. To address the challenge of the inherently complex and dynamic growth of the VTF time series, a new hierarchical methodology for VTF prediction is proposed. Firstly, the original VTF data is reconfigured as a three-dimensional tensor by a modified Bayesian Gaussian CANDECOMP/PARAFAC (BGCP) tensor decomposition model. Secondly, the VTF matrix (hour × day) of each week is decomposed into high- and low-frequency matrices using a Bidimensional Empirical Mode Decomposition (BEMD) model to address the non-stationary signals affecting prediction results. Thirdly, the self-similarities between VTF matrices of each week within the high-frequency tensor are utilised to rearrange the matrices as different one-dimensional time series to solve the weak mathematical regularity in the high-frequency matrix. Then, a Dynamic Time Warping (DTW) model is employed to identify grouped segments with high similarities to generate more suitable high-frequency tensors. The experimental results verify that the proposed methodology outperforms the state-of-the-art VTF prediction methods using real Automatic Identification System (AIS) datasets collected from two areas. The methodology can potentially optimise relation operations and manage vessel traffic, benefiting stakeholders such as port authorities, ship operators, and freight forwarders.

1. Introduction

The shipping industry has accounted for more than 80% of the total international trade in recent years, contributing to the prosperity of the world economy (Li and Yang, 2023; Zheng et al., 2022). The vigorous growth of the shipping industry has caused an increase in the frequency of water traffic, making it a common occurrence (Xin et al., 2023a,b; M. Zhang et al., 2023). Nevertheless, this growth has also led to a more complex water traffic situation, as well as the risk of water traffic accidents (Jiang et al., 2022; H. Li et al., 2023; Li et al., 2022; Li and Yang, 2023; M. Li et al., 2023). According to the global annual reports of international maritime traffic accidents by the European Maritime Safety Agency (EMSA), the statistics of maritime traffic accidents and the number of accidents based on ship types are displayed in Fig. 1 (a) and (b), respectively. The trends from 2014 to 2020 indicate that cargo ships are at a much higher risk of accidents than passenger, fishing, service, and other ships (EMSA, 2022). Therefore, how to effectively enhance the safety of ship navigation has become a crucial issue in the research of water transportation.

Vessel Traffic Flow (VTF) represents the number of vessels passing through a specific research zone within a designated time interval (Xiao et al., 2023). The interpretation of VTF data may differ significantly within the study area (Li et al., 2023). For example, VTF data on port waters can reflect the port's busyness and economic indicators. With the outbreak of COVID-19, especially in 2020, the VTF of ports in various countries showed a downward trend, reflecting the gradual decline in import and export trade (Liu et al., 2023; Zhao et al., 2022). On the other hand, VTF data of channel waters can indicate traffic intensity. An increase in VTF data expresses a higher number of ships sailing. However, due to the limited capacity of the channel, an excessive flow of ships can lead to an increased probability of accidents (Su et al., 2022; Xu and Zhang, 2022). In summary, VTF can reflect the traffic situation of a

* Corresponding author.
** Corresponding author.
E-mail addresses: liyanWHU@whu.edu.cn (Y. Li), Z.Yang@ljmu.ac.uk (Z. Yang).
[1] Equal contribution.

https://doi.org/10.1016/j.oceaneng.2023.115687
Received 6 April 2023; Received in revised form 2 August 2023; Accepted 21 August 2023
Available online 31 August 2023

specific research area to some extent, and the magnitude of the flow can also provide insights into determining the probability of accidents.

In recent years, due to continuous innovations in data mining technology, research on VTF has been conducted from various perspectives, including VTF simulation (Goerlandt and Kujala, 2011; Park et al., 2002), VTF feature extraction (Gao and Shi, 2019; Rong et al., 2022), and VTF prediction (D. Wang et al., 2021; Zhang et al., 2022). More specifically, VTF prediction enables the acquisition of flow data within a study area for a certain period in the future (Xiao et al., 2020, 2022). Hence, VTF-based prediction data becomes a vital element in forecasting the likelihood of accidents in that particular area. In other words, it is crucial to ensure the efficiency and safety of maritime transportation systems by achieving accurate information about near-term VTF data in realistic water traffic networks.

It is of utmost importance to forecast future trends by analysing the development characteristics of historical data for VTF prediction. Automatic Identification System (AIS) transmitters have been mandated to be installed on all international voyage vessels of no less than 300 gross tonnages, as well as passenger vessels (Chen et al., 2022; Li et al., 2023a; Robards et al., 2016; Rong et al., 2020). Consequently, the AIS data of the vessels passing through a channel and a port, including vessel location, course, and speed, can be received through the AIS base stations and/or satellites to calculate the VTF in the study area, as illustrated in Fig. 2. Therefore, data-driven methods become a feasible technical solution to VTF prediction. However, the accurate and reliable prediction of VTF data is challenging. A considerable challenge in obtaining reliable prediction data is that the dynamic growth of VTF often exhibits complex and non-stationary characteristics. Additionally, emergencies such as accidents, channel closures, or severe weather significantly impact the reliability of VTF data prediction. Consequently, recent efforts have focused on exploring new and revised methods to enhance prediction capacity and address the challenges associated with VTF data prediction.

VTF belongs to time series data, and hence conventional Machine Learning (ML) (Jin et al., 2021; Tang et al., 2019), Neural Networks (NN) (Dikshit et al., 2022; Do et al., 2019), and Deep Learning (DL) (Wang et al., 2019; Yin et al., 2022) are the most general and effective prediction methods. In particular, when dealing with a substantial volume of historical VTF data and facing high data volatility, DL methods tend to outperform other approaches in predicting future data (Liang et al., 2022). However, a common drawback of most time series prediction methods is their limited ability to effectively capture the internal relationship between the specific historical and current data during the modelling process. The above problem no doubt reduces the accuracy and stability of VTF prediction. Following big data technology, large-scale and multi-dimensional spatiotemporal datasets are increasingly prevalent in real applications. Thus, a growing number of studies focus on predicting high-dimensional data. For example, the original time series (i.e., one-dimensional (1D) time series) could be reconfigured into a three-dimensional tensor. The essence is that spatiotemporal time series data often display strong relevance and shared potential patterns, such as VTF time series with recurring temporal peaks (Chen and Sun, 2022). As a result, tensor factorisation in time series prediction has garnered significant attention. Specifically, Salakhutdinov and Mnih (2008) proposed the Bayesian matrix method. Moreover, Chen et al. (2019b) designed a Bayesian probabilistic matrix factorisation method for higher dimensional tensors, called a Bayesian Gaussian CANDECOMP/PARAFAC (BGCP) method, which could effectively handle tasks of imputing and predicting three-dimensional tensor data. Therefore, the BGCP method is selected as the foundational framework for in-depth exploration of the three-dimensional (3D) VTF prediction problem.

This study aims to develop a hierarchical method that enables high prediction accuracy and strong stability for VTF data prediction in the maritime industry. The following content is outlined as follows. The state-of-the-art methods for VTF prediction are reviewed, and the contributions of our study are summarised in Section 2. Section 3 is the preparatory part, which introduces the principles of Bidimensional Empirical Mode Decomposition (BEMD) (Chen et al., 2013; Hou et al., 2019), BGCP, and Dynamic Time Warping (DTW) methods to provide the necessary background information. This paper proposes a new optimised hierarchical prediction methodology based on these three algorithms. Section 4 describes the technical principles of the proposed hierarchical prediction methodology. A comprehensive experiment in Section 5 using four different VTF datasets is conducted to compare the VTF prediction performance with 17 established methods. Section 6 outlines the conclusions and outlines future directions.

2. Literature review

It is necessary to develop VTF prediction to explore knowledge discovery in maritime traffic management. Especially in recent years, the popularisation of AIS equipment and the maturity of data acquisition technology has allowed researchers to analyse VTF in target waters using AIS data. Hence, the data-driven VTF prediction methods have gained significant attention and become the mainstream prediction scheme.

In this section, a critical analysis of the VTF prediction research is conducted from three perspectives: traditional ML-based methods in Section 2.1, NN and DL-based methods in Section 2.2, as well as tensor factorisation methods in Section 2.3. All three types of methods can be used to predict future traffic through the characteristics of historical VTF

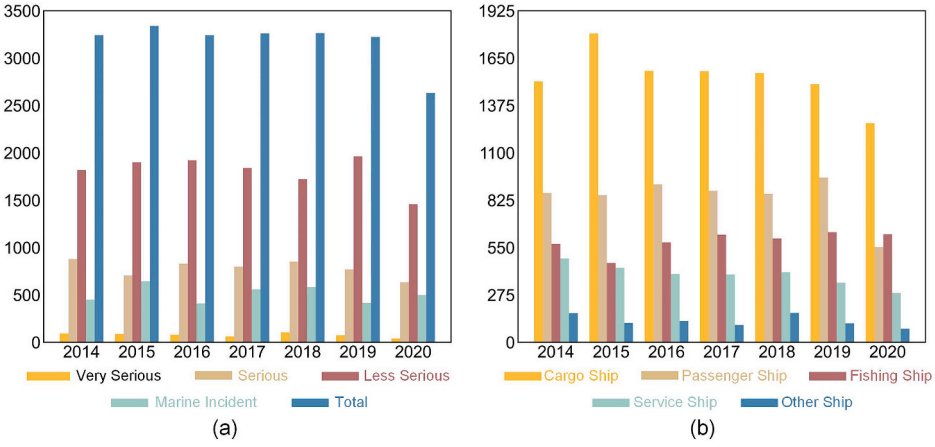


Fig. 1. The result of global maritime accidents from 2014 to 2020 by the EMSA, (a) the statistics of maritime traffic accidents and deaths, and (b) traffic accident data based on different ship types.

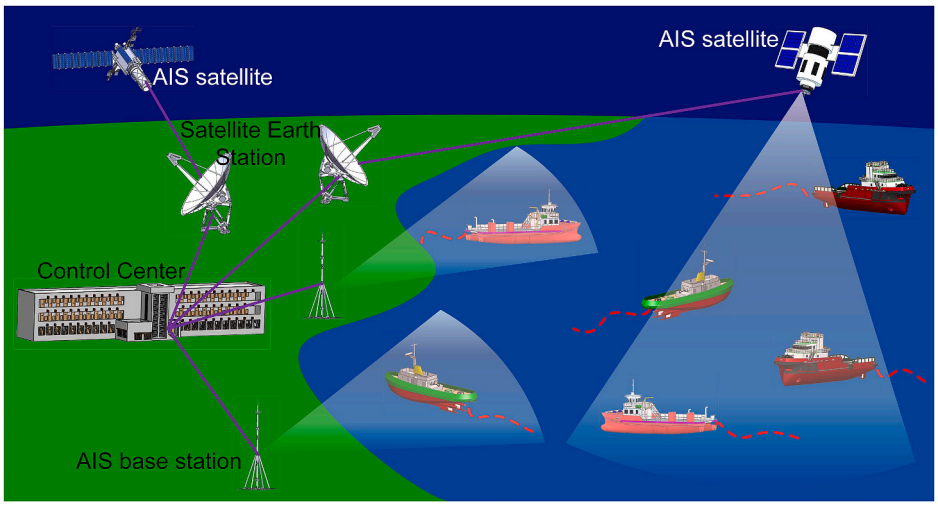


Fig. 2. Overview of terrestrial and satellite AIS networks by collecting AIS data.

data. Finally, the novel contributions of this study are revealed through a critical analysis of the current state of the art in the field.

2.1. Traditional ML methods in VTF prediction

In the early stages of VTF prediction research, scholars used traditional ML methods, also known as statistical methods, to conduct research (Li and Yang, 2023). These methods assume a specific correlation between historical data and forecast data. Meanwhile, VTF data generally changes periodically over time. Thus, mathematical methods can be developed through parameter estimation and curve fitting to predict VTF data.

The simple and commonly used ML methods have been applied to predict research data, for instance, linear regression (Sousa et al., 2007) and variance analysis models (Makowski et al., 2006). However, they are generally not effective in non-stationary time series prediction. Because of the complex features of the VTF data, there have been many studies in the literature on other traditional ML methods, such as Kalman Filtering (KF) (He et al., 2019; Muruganantham et al., 2016; Okutani and Stephanedes, 1984; Xu et al., 2017), Grey theory-based Models (GM) (Kayacan et al., 2010; J. Wang et al., 2021), Markov Model (MM) (Lin et al., 2023; Zou et al., 2022), Kernel Density Estimation (KDE) (Jiang et al., 2019; Zhang et al., 2022), AutoRegressive Integrated Moving Average (ARIMA) (Liu et al., 2022; Weng et al., 2019), Support Vector Machine (SVM) (Gao et al., 2023; Mokhtarimousavi et al., 2019; Zhang and Wu, 2022), and Bayesian forecasting model (Bürkner et al., 2020; Du et al., 2022; Chen and Sun, 2022).

The principles of the above methods in solving the VTF prediction problem are different. The KF can estimate the future trend of VTF under the uncertainty of its historical development characteristics and influential factors. However, if the historical data has poor regularity, the prediction accuracy based on the KF will be significantly reduced. The GM identifies the different degrees of development trends between VTF data at different historical time points. Based on the overall change rule of the original VTF data, the sequence data with strong regularity is regenerated. Finally, the future development trend of VTF can be predicted by establishing the corresponding differential equation model. However, if the VTF data has frequent mutations, the prediction accuracy of the data may be reduced due to an error accumulation problem. MM is a method to predict the probability of events, which is a type of random process. However, the method has a disadvantage in solving the VTF prediction problem. When the information of the stochastic process at time t_0 is known, the changing trend of t ($t > t_0$) is only related to the state of t_0, and it has had no relationship with the states at any other time. Therefore, the MM-based VTF prediction results may be inaccurate. The main principle of KDE is to infer the overall distribution using limited samples. The distribution characteristics of the VTF data can be captured by calculating the probability density function of the sample data through the KDE method, enabling the prediction of future data based on these characteristics. However, in cases where the VTF data exhibits non-periodic changes or the change characteristics are not evident, the KDE method cannot effectively estimate the global characteristics, leading to reduced prediction accuracy. The essence of the ARIMA is to predict future traffic through a linear combination of VTF historical and current data. SVM is a binary classification approach. When applied to VTF prediction, SVM employs historical data as training samples to construct control functions that capture the patterns of historical data changes. The Bayesian forecasting model relies on Bayesian statistics for prediction. This model differs from conventional ML methods as it leverages both model and data information and effectively exploits prior knowledge during the data prediction process.

The aforementioned traditional ML methods can make relatively accurate predictions when the fluctuation of VTF data is significant. Furthermore, several extensional versions of these methods have been extensively employed to improve the accuracy of time series prediction even further. For instance, the conventional ARIMA model has undergone significant improvements, such as Kohonen-ARIMA (KARIMA) (Van Der Voort et al., 1996), Seasonal ARIMA (SARIMA) (Williams and Hoel, 2003), ARIMA with Generalized Autoregressive Conditional Heteroscedasticity (ARIMA-GARCH) (Chen et al., 2011), and Online Change-Point-Based model (OCPB) (Comert and Bezuglov, 2013). SVM, being a classic and widely adopted ML method, has undergone various upgrades, including SVM with Grey Wolf Algorithm (GWO-SVM) (Zhou et al., 2022), SVM with Slime Mold Algorithm (SMA-SVM) (Zhao et al., 2023), and SVM with Whale Optimization Algorithm (WOA-SVM) (Kong et al., 2020). Scholars have made significant efforts to optimise traditional ML methods, improving their prediction accuracy to some extent.

Scholars have also begun to use the NN and DL methods for VTF prediction, motivated by the continuous development of NN and DL concepts and technologies. Although NN and DL also belong to ML methods, they leverage the characteristics of VTF data by constructing training networks. In network training, a loss function continuously optimises the weight value to infinitely reduce a prediction error. Therefore, compared with traditional ML methods, NN and DL approaches can better predict VTF data with high volatility. The development and application of the ML methods in VTF prediction are introduced and discussed in detail in Section 2.2.

2.2. NN and DL methods in VTF prediction

In recent years, NN and DL methods are playing an increasingly important role in VTF data prediction. Most NN structures consist of an input, hidden, and output layer and have strong self-learning abilities, making them well-suited for non-stationary time series data and complex VTF prediction. Recent studies reveal that NN methods have been applied to the prediction of VTF data, including Back Propagation Neural Network (BPNN) (Yi et al., 2021), Wavelet Neural Network (WNN) (Chen et al., 2021), Grey Neural Network (GNN) (Pang et al., 2020), Fuzzy Neural Network (FNN) (Chan and Dillon, 2013), Elman Neural Network (ENN) (Wei, 2019), Generalized Regression Neural Network (GRNN) (Celikoglu and Cigizoglu, 2007), and Probabilistic Neural Network (PNN) (Nguyen et al., 2022). Meanwhile, Li et al. (2019) applied the widely-used NN methods (i.e., WNN, ENN, FNN, BPNN, and GRNN) to predict the Port Cargo Throughput (PCT) and VTF time series. Their network structure is shown in Fig. 3. In addition, they improved the prediction framework of traditional NN methods and proposed similarity grouping-guided NN methods, which could improve the prediction accuracy of complex PCT and VTF. Scholars often enhance the prediction accuracy of time series data by optimising classic and commonly used networks. Chen et al. (2021) integrated the particle swarm optimization algorithm into WNN, which solves the problems of slow convergence and local optimization. Their improved network reduced the prediction error by 14.99% compared to the original WNN. Z. Zhang et al. (2019) used Self-Adaptive Particle Swarm Optimization (SAPSO) algorithm to adjust the structural parameters of BPNN, resulting in the SAPSO-BP network, which accurately and stably predicted the VTF change trend of the Los Angeles Port Area. Sadeghi-Niaraki et al. (2020) developed Elman Recurrent Neural Network called GA-MENN for short-term traffic flow prediction. The core concept involved optimising the Elman Recurrent Neural Network using Genetic Algorithm and considering qualitative factors like weather conditions and workdays during training. Kaffash Charandabi et al. (2022) proposed a GRNN adjusted by a self-organizing map to predict traffic accident risk. Their method comprehensively integrated 22 indicators (e.g., geographical characteristics, weather conditions, time conditions, driver characteristics, and driving conditions) affecting prediction accuracy into the network. Through 30 different application scenarios, the proposed method demonstrated excellent performance in experimental analysis. Tang et al. (2021) combined a fuzzy rough set and FNN to address the missing data. This method utilised FNN for data classification and predicted missing values based on fuzzy rough sets. The effectiveness of the new method was quantitatively evaluated using metrics such as Root Mean Square Error (RMSE), correlation coefficient, and relative accuracy.

Compared to traditional ML methods, NN methods and their upgraded versions have shown improved accuracy in time series data prediction. However, the increasing complexity of time series data due to external environmental influences poses challenges for simple training networks to accurately learn these irregular features. Fortunately, the advent of deep training networks has partially addressed the limitations of NN methods, providing opportunities to handle more complex time series data.

DL was first proposed by Hinton et al. (Hinton and Salakhutdinov, 2006), which originated from the research of NN, allowing for the development of networks with many hidden layers and massive training data to learn more valuable and complex features to improve the accuracy of prediction and classification. Therefore, several mainstream DL network frameworks, such as Convolutional Neural Networks (CNN) (W. Zhang et al., 2019), Recurrent Neural Networks (RNN) (Suo et al., 2020) and Attention Mechanisms (AM) (Sun et al., 2020), have attracted significant attention in VTF prediction, as displayed in Fig. 4 (Lim and Zohren, 2021). AM has been shown to improve long-term dependent learning compared with CNN and RNN. The attention layer uses dynamically generated weights (i.e., denoted by the red rows) to aggregate time characteristics, as shown in Fig. 4 (c), which allows the network node at the current time to focus directly on the network node at any time in the past, even if the time span is very long (Lim and Zohren, 2021). Overall, the development of NN and DL methods has significantly contributed to the accurate prediction of VTF data.

RNN models have become popular for time series data prediction due to their capability to combine historical and current data through hidden states. Belhadi et al. (2020) introduced the application of RNN in long-term traffic flow prediction and proposed RNN-LF for long-term traffic flow prediction from multiple data sources. Furthermore, they implemented a new network model called GRNN-LF in the GPU parallel computing framework, significantly improving the operating efficiency. The issue of gradient vanishing (or explosion) hinders the performance of RNNs, which affects their prediction accuracy. To address this, LSTM (Yang et al., 2019) and GRU (Hussain et al., 2021) networks have been put forward, showing better performance in predicting VTF data. Dong (2022) used wavelet analysis to decompose the original VTF data into trend and interference terms. LSTM was then employed to predict these sequences separately, and the final result was obtained by summing the predicted values. Hussain et al. (2021) proposed a GRU-based network for hyperparameter adjustment and sliding window step size optimization, effectively addressing the need for continuous manual adjustments of hyperparameters during network training. However, the LSTM and GRU methods belong to unidirectional networks that retain only past information. Bi-directional networks like Bi-LSTM (Li et al., 2021) and Bi-GRU (Huang et al., 2021) have better accuracy and stability since they retain both past knowledge and future information. Ma et al. (2022) combined the structures of LSTM and Bi-LSTM in a new network model, while Yu et al. (2021) proposed a short-term traffic flow prediction method based on Bi-GRU to capture the periodic change attributes of traffic flow data.

A limitation in the practical training of RNN and its optimised networks is the requirement for consistent data sizes in the input layer. To overcome this issue, Sequence to Sequence (Seq2Seq) was proposed in 2014, employing an Encoder-Decoder structure. The encoder converts the input sequence into a fixed-length vector, and the decoder generates an output sequence from this vector (Cho et al., 2014; Sutskever et al., 2014). Seq2Seq has found wide applications in intelligent

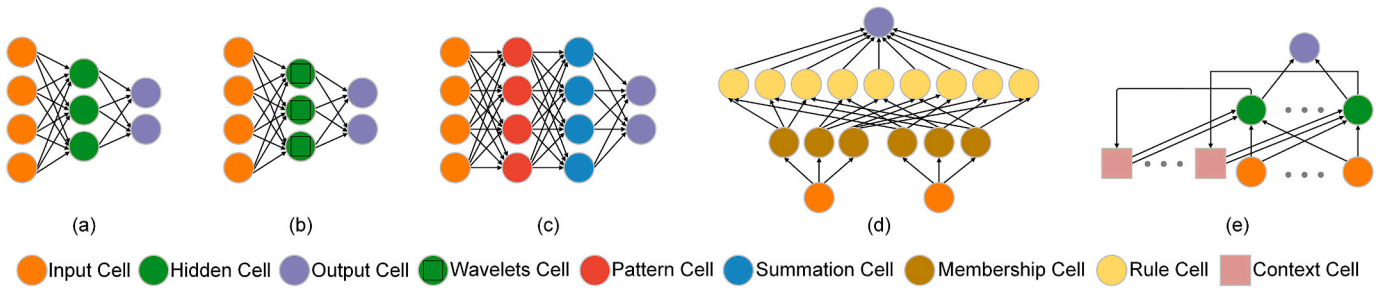


Fig. 3. The network structure diagram of five different NN methods. (a) BPNN, (b) WNN, (c) GRNN, (d) FNN, and (e) ENN.

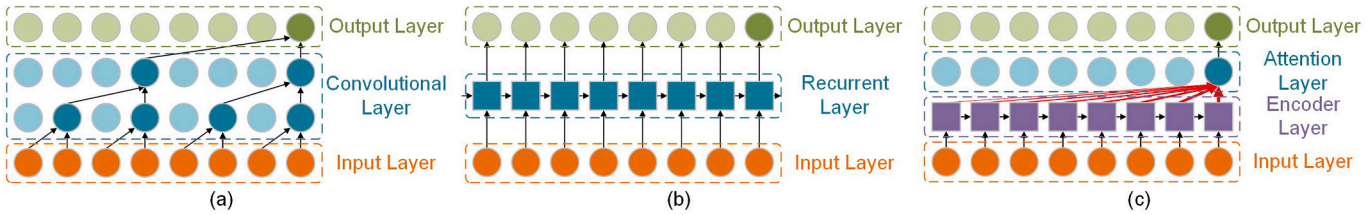


Fig. 4. The architectural diagrams of the three main DL methods mentioned. From left to right: (a) CNN method, (b) RNN method, (c) Attention-based method (red arrow indicates the attention weight).

recommendation, question answering, machine translation, information retrieval, and data prediction research.

Researchers have introduced various variations of Seq2Seq, such as the spatio-temporal Seq2Seq network (STSSN) (Cao et al., 2022), which uses an Enhanced Diffusion Convolutional Network (EDCN) and a Temporal Convolutional Network (TCN) in its encoder and decoder. Hao et al. (2019) embedded an attention mechanism in Seq2Seq to conduct short-term passenger traffic flow prediction, enhancing the training network's ability to capture long-term dependencies. Additionally, hybrid DL methods like CNN-LSTM networks (Narmadha and Vijayakumar, 2021) effectively capture the spatial and temporal characteristics of VTF time series data. The combinations of CNN and other RNNs (e.g., iGRU, Bi-LSTM, and Bi-GRU) are also widely used in the prediction research of time series data (Ma et al., 2023; Méndez et al., 2023; Wang et al., 2022).

NN and ML can learn the historical characteristics of VTF based on the network training model and then predict future data. However, existing training networks have one drawback, that is, they cannot effectively reflect the internal relationship between specific historical and current data, leading to a significant deviation in predicting VTF data. To address this problem, new solutions are needed by exploring high-dimensional data prediction, such as using tensor factorisation to impute and predict three-dimensional tensor data. Section 2.3 describes the development and applications of the tensor factorisation methods in VTF prediction.

2.3. Tensor factorisation methods in VTF prediction

The emergence of big data has contributed to the applications of large-scale and multi-dimensional spatiotemporal datasets due to the high relevance and shared potential patterns of spatiotemporal time series data (i.e., VTF time series with repeating temporal peaks) (Chen and Sun, 2022). This idea has inspired numerous studies to utilize tensor factorisation methods in analysing large-scale VTF data, which involves the projection of the original 1D data onto a three-dimensional tensor (i.e., hour × day × week). Compared to traditional ML, NN, and DL methods, the most significant advantage of tensor factorisation methods is their ability to treat all historical data as a whole and explore the development characteristics of the data from a global perspective.

The initial applications focused on tensor data imputation. For example, Liu et al. (2013) proposed three classic decomposition methods to compensate for missing values in tensors, namely Simple Low-Rank Tensor Completion (SiLRTC), Fast Low-Rank Tensor Completion (FaLRTC), and High accuracy Low-Rank Tensor Completion (HaLRTC). SiLRTC employed a straightforward implementation principle and utilised block coordinate descent to obtain the global optimal solution. FaLRTC transformed the non-smoothing problem into a smoothing problem using a smoothing scheme for processing. HaLRTC effectively applied the alternating direction method of multipliers to achieve high-precision interpolation results. SiLRTC exhibited the least ideal interpolation accuracy compared to the other two techniques. Several studies have reorganized multi-variable traffic flow data into four-dimensional tensors to complete missing values in the original data. (Tan et al., 2016). Chen et al. (2019a) proposed the Bayesian Augmented Tensor Factorisation (BATF) method, utilising variational Bayes to learn model parameters automatically. A new tensor factorisation method is utilised to complete the missing traffic flow data, called high-order singular value decomposition with soft thresholding core (Gong and Zhang, 2020). Baggag et al. (2021) proposed a new Temporal Regularized Tensor Factorisation (TRTF) method, which fully takes into account the temporal dependence and spatial attributes between traffic tensor data. Tensor factorisation using Alternating Least Square (TF-ALS) is a new decomposition method that provides a theoretical analysis to limit approximation errors and improve interpolation accuracy (Ma and Solomonik, 2022). The fundamental concept behind this approach involves decomposing the initial traffic flow data into its main trend and residual components, followed by individual processing of these two components using the tensor decomposition technique. Tensor factorisation methods are also used in Spatial-Temporal Bi-directional Residual Optimization (ST-BiRT) to address the issue of the missing value in traffic flow data. Compared to other methods, the innovation of the ST-BiRT focuses on a well-designed two-way residual structure, significantly reducing the model error. The ST-BiRT method can be implemented with better accuracy and robustness in dealing with rate problems and missing scenarios (Li et al., 2022).

The issue of traffic flow tensor prediction is similar to the one of completing missing values to a certain extent. While data missing value completion processes any time node of an entire time series data, the prediction handles the data on the future time nodes based on the whole time series. Prediction can be considered as an interpolation problem of missing values. Therefore, the tensor factorisation methods can be effectively applied to traffic flow tensor prediction. For instance, a compact tensor factorisation method is used to cluster and predict the temporal evolution of global congestion allocation in large-scale urban transport networks (Han and Moutarde, 2016). In another study, the urban traffic flow is reorganized into a four-dimensional (i.e., 4D) tensor, incorporating location, time of day, day of the week, and week of the month. Meanwhile, the proposed method estimates the correlation of each tensor dimension and uses a gradient descent strategy to optimise the prediction method, which can address the problem of data sparsity from the perspective of spatial and temporal traffic modes (Yang et al., 2020). Another advantage of using the tensor factorisation methods for VTF tensor data prediction is that they can handle missing values in the historical VTF data, which is not possible in conventional methods. Therefore, many scholars use this approach to predict high-dimensional data tensors.

Most aforementioned tensor factorisation methods often suffer from issues related to data interpolation and prediction. Specially, these methods only offer point estimation for interpolation or prediction tasks, which may cause interpolation errors or low prediction accuracy (Chen and Sun, 2022). To address the above problems and avoid over-fitting, an effective combination of Bayesian treatment and the tensor factorisation methods is proposed. For example, the Bayesian Tensor Factorisation (BTF) method could set the CANDECOMP/PARAFAC (CP) rank (Zhao et al., 2015) automatically. As an optimised version of the BTF method, the BGCP method is proposed to infer the posterior distribution of missing values (Chen et al., 2019b). Compared with other tensor factorisation methods, it could effectively complete the tasks of

three-dimensional tensor data imputation and prediction. Hence, this paper chooses the BGCP method as the basic framework to solve the 3D VTF prediction problem. In Section 5.4, quantitative experiments will be conducted to validate that BGCP outperforms other tensor factorisation methods of the same type in terms of accuracy and stability when carrying out VTF prediction tasks. However, the original 1D VTF data may be affected by some unexpected factors (e.g., climatic factors, collision accidents, and economic factors) in the development over time, making the data highly volatile. High-quality and stable prediction results cannot be obtained if the BGCP method is directly used to predict the original VTF tensor. Therefore, the paper develops the original BGCP method and introduces a novel prediction framework that can be used to effectively address the non-stationary and complex VTF prediction problem. The contributions of our proposed hierarchical method in VTF prediction are summarised in Section 2.4.

2.4. Contributions of our study

The proposed hierarchical method, BEMD-DTW-BGCP, decomposes the VTF tensor into high- and low-frequency tensors, and takes advantage of the inherent self-similarity between the VTF matrices of each week within the high-frequency tensor. To create a more appropriate high-frequency tensor for the BGCP model, a DTW method is employed to select the segments with significant similarities. The final prediction results are obtained by combining the predicted high- and low-frequency tensors. The proposed method shows superior prediction accuracy and robustness performance compared to other methods, as demonstrated in experiments on spatiotemporal VTF datasets.

The main contributions of the proposed method are listed below.

(1) A hierarchical methodology is proposed to predict non-stationary time series with satisfied performance, involving matrix decomposition, similarity grouping, and feature combination.
(2) A new matrix decomposition method based on BEMD is applied to transfer the original VTF tensor into high- and low-frequency tensors, while the inherent self-similarity of the high-frequency tensor is extracted by DTW to improve prediction accuracy.
(3) A new comprehensive comparative analysis with 17 other methods is implemented to evaluate the accuracy and robustness of the proposed hierarchical methodology in VTF prediction.
(4) A new evaluation method is proposed to evaluate the prediction performance via coarse-grained (i.e., RMSE and Mean Absolute Percentage Error (MAPE)) and fine-grained (i.e., the Mean Relative Error (MRE) and the Standard Deviation of Relative Error (SDRE)) features.

The primary advantage of the proposed method is its capacity to take full advantage of the BEMD-based property decomposition for matrices and DTW-based similarity grouping, while effectively accomplishing the task of 3D tensor data prediction using the BGCP method. As a result, the new hierarchical prediction framework can enhance accuracy and robustness in various applications.

3. Preliminary

A hierarchical prediction methodology is proposed based on the BEMD, DTW, and BGCP algorithms. Therefore, this section mainly introduces the principles of these three algorithms as a preliminary part, which will pave the way for the optimised hierarchical prediction method presented in Section 4.

3.1. Overview of BEMD-based matrix decomposition

Huang et al. (1998) first proposed the Empirical Mode Decomposition (EMD) algorithm in 1998 to adaptively analyse the non-stationary signals (i.e., time series data). A signal is decomposed by the algorithm into a restricted set of Intrinsic Mode Functions (IMFs) spanning from high to low frequencies, along with a residue. This process could be defined as follows:

$$x(t) = \sum_{p=1}^{N} c_p(t) + r_N(t), \tag{1}$$

where $x(t)$ denotes the time series data, N is the number of IMFs, $c_p(t)$ expresses the pth IMF. $r_N(t)$ represents the residual signal, which is the mean trend of the original data $x(t)$.

Each IMF has to meet two conditions.

(1) The total count of zero crossing and extreme points in the signal dataset needs to be equivalent or at least one difference.
(2) The average value of the upper and lower envelopes, represented by the local minimum and maximum, at any given point in the signal must be zero.

The EMD algorithm is an effective method for decomposing signals, particularly non-stationary time series data. The BEMD algorithm is created to decompose a matrix into high-frequency and low-frequency components, along with a few residues, leveraging the success of EMD in 1D signal analysis. The BEMD process works in a comparable manner to EMD, serving as a logical extension of EMD to a two-dimensional (2D) space. Specifically, the detailed operation of sifting the Bidimensional Intrinsic Mode Functions (BIMFs) of a given matrix could be summarised as follows.

Step 1. The positions of the extrema (local maxima and minima) are located in a decomposed matrix $MD_0^1 = M(x,y)$.

Step 2. The upper and lower 2D envelopes are generated by 2D interpolation based on the maximum and minimum point sets, denoting as $e_u(x,y)$ and $e_l(x,y)$, respectively.

Step 3. The local mean $e_m(x,y)$ is calculated by averaging the two envelopes above, which could be defined as

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

Step 4. Subtract the local average from the original matrix

$$MD_k^l(x,y) = MD_{k-1}^l(x,y) - e_m(x,y), \tag{3}$$

Step 5. Repeat the process from Steps 1 - 4 until the stop condition is met. The $MD_k^l(x,y)$ is output as a BIMF. In addition, the criterion adopted to stop the process depends on the normalized Standard Deviation (SD) between $MD_k^l(x,y)$ and $MD_{k-1}^l(x,y)$. To be more specific, the SD could be defined as

$$SD = \sum_{x=0}^{m} \sum_{y=0}^{n} \left(\frac{\left| MD_{k-1}^l(x,y) - MD_k^l(x,y) \right|}{MD_{k-1}^l(x,y)} \right)^2 < \eta, \tag{4}$$

where m and n represent the number of rows and columns in the matrix, k is the iteration number, and l is the index corresponding to the lth BIMF. The numerous tests show that the best range η is between 0.3 and 0.5.

Step 6. When meeting the stop criterion, the $BIMF(B_1(x,y))$ could be defined from Step 4 as.

$$B_1(x,y) = MD_k^l(x,y), \tag{5}$$

Step 7. According to Eq. (5), the BIMF is obtained. The

$residue(R_1(x,y))$ is then defined as

$$R_1(x,y) = MD_0^l(x,y) - B_1(x,y), \tag{6}$$

Step 8. The next BIMF is calculated from Step 1. In addition, the residue will be used as the input matrix to obtain the next BIMF, which could be defined as

$$MD_0^{l-1}(x,y) = R_1(x,y), \tag{7}$$

The BEMD algorithm can generate multiple BIMFs from the input matrix by repeating Steps 1 - 8. The filtering process is considered to be completed when there are no more extreme points in the residual matrix. Additionally, the BEMD algorithm requires two conditions to decompose a matrix, which is the same as the conditions for the EMD algorithm to decompose 1D time series. The original matrix $M(x,y)$ is decomposed into a series of BIMFs and a residue that could be defined as.

$$M(x,y) = \sum_{q=1}^{L} B_q(x,y) + R_L(x,y). \tag{8}$$

where $B_q(x,y)$ is the qth BIMF and $R_L(x,y)$ represents the residue. The BEMD algorithm is selected to reconstruct 1D VTF data into a 2D matrix.

3.2. Overview of DTW-based similarity grouping

Similarity measure has played an essential role in time series analysis in recent years, which significantly affects the quality of data mining (Liang et al., 2021). The Euclidean distance is a widely employed approach for describing the similarity between two distinct time series. However, it is sensitive to the outliers of time series. DTW is an optimization distance measure to compute the similarity between two different time series by dynamic programming to find the minimum distance. DTW can better identify corresponding points with similar geometric shapes than the Euclidean distance (Li et al., 2020), making the distance (similarity) measurements more accurate. Meanwhile, DTW can measure the similarity distance between two time series with different lengths. However, the Euclidean distance can only handle two-time series data with the same length. Therefore, the DTW method has extensive applications in various fields, including data clustering, feature extraction, voice identification, etc.

Let $X = \{x_1, x_2, \cdots, x_M\}$ and $Y = \{y_1, y_2, y_3, \cdots, y_N\}$ denote two different time sequences with the lengths M and N, respectively. Given a $M \times N$ matrix ω with the $\omega(s,t)$ is the distance $d(x_s, y_t)$ between the two points x_s and y_t. In addition, the $d(x_s, y_t)$ corresponds to the weighted Euclidean distance. To align two time sequences, the shortest path corresponds to the best match. More specifically, the warping path could be calculated by

$$DTW(X,Y) = \gamma(s,t), \tag{9}$$

where the minimum cumulative distance could be given by

$$\gamma(s,t) = d(x_s, y_t) + \min\{\gamma(s-1,t-1), \gamma(s-1,t), \gamma(s,t-1)\}. \tag{10}$$

The smaller the distance, the more similar these two time sequences. In this paper, the matrix of each week is transformed from a high-frequency tensor into a 1D series. Subsequently, the grouped segments with high DTW-based similarities are selected to form more suitable high-frequency tensors for the BGCP, thereby improving prediction performance.

3.3. Overview of BGCP-based tensor prediction

In this subsection, a multi-dimensional data tensor decomposition called the BGCP method is introduced to solve traffic data imputation problems. VTF data prediction can be regarded as a particular data imputation problem. Hence, this study selects the BGCP method as a primary research framework. The previous applications of tensor decomposition have mainly focused on data imputation, with no prior use of this model in VTF data prediction. In contrast, 1D VTF time series data is rearranged as a 3D tensor to be predicted based on the BGCP method.

The BGCP model can be transferred into a sum of r rank-one component tensors, and r is defined as the original tensor's CANDECOMP/PARAFAC (CP) rank. The BGCP model is summarised in the following two parts (Chen et al., 2019b).

Gaussian assumption: Given a matrix $Z \in \mathbb{R}^{m \times n \times f}$ with missing (or predicted) entries with m is the number of divided time intervals, n indicates the days of the week, and f represents the number of weeks in the dataset. Then, factorisation can be applied to reconstruct the missing (or predicted) values within Z, which could be defined as follows.

$$z_{ijt} \sim N\left(\sum\nolimits_{s=1}^{r} g_{is} h_{js} x_{ts}, \delta^{-1}\right), \forall i,j,t,s \in \{1, \cdots, r\}, \tag{11}$$

where $N(\cdot)$ is a Gaussian distribution, and the vector $g_s \in \mathbb{R}^m$, $h_s \in \mathbb{R}^n$, $x_s \in \mathbb{R}^f$ are columns of latent factor matrices. In addition, g_{is}, h_{js}, x_{ts} are their elements, and δ denotes precision.

Bayesian framework: To introduce the Bayesian framework, the BGCP method sets conjugate Gaussian-Wishart priors for hyper-parameters $\eta_x \in \mathbb{R}^r$ and $\Theta_x \in \mathbb{R}^{r \times r}$, which could be defined as follows.

$$\eta_x \sim N\left(\eta_0, (\lambda_0 \Theta_x)^{-1}\right), \tag{12}$$

$$\Theta_x \sim Wishart(W_0, h_0), \tag{13}$$

where η_x is normally distributed with a mean η_0 and a variance $(\lambda_0 \Theta_x)^{-1}$. Θ_x follows Wishart distribution. W_0 denotes the degrees of freedom parameter, which determines the number of random variables that are being squared and summed to generate the Wishart random matrix. h_0 represents the scale matrix parameter, which is a positive definite matrix. It influences the dispersion and spread of the generated Wishart random matrix.

The $X \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ is given a partially observed tensor. Gibbs sampling is applied to implement Bayesian inference.

Concerning hyper-priors, the initialization values are set as $\eta_0 = 0$, $\lambda_0 = 1$, $W_0 = I$, and $h_0 = r$. In addition, I and r are the identity matrix and low rank, respectively.

The $\widehat{W}_0^{(k)}$ and $\widehat{h}_0^{(k)}$ are calculated as shown.

$$\widehat{W}_0^{(1)} = \left(n_1 S^{(1)} + W_0^{-1}\right)^{-1} + \left(\frac{n_1 \lambda_0}{n_1 + \lambda_0}\left(\overline{g}^{(1)} - \eta_0\right)\left(\overline{g}^{(1)} - \eta_0\right)^T\right)^{-1}, \tag{14}$$

$$\widehat{h}_0^{(1)} = n_1 + h_0, \tag{15}$$

where $\overline{g}^{(1)}$ and $S^{(1)}$ denote two statistics, defined as

$$\overline{g}^{(1)} = \sum_{i_1=1}^{n_1} g_{i_1}^{(1)}, \tag{16}$$

$$S^{(1)} = \frac{1}{n_1} \sum_{i_1=1}^{n_1} \left(g_{i_1}^{(1)} - \overline{g}^{(1)}\right)\left(g_{i_1}^{(1)} - \overline{g}^{(1)}\right)^T, \tag{17}$$

Samples $\Theta^{(k)}$ follows Wishart distribution. The $\widehat{\eta}^{(k)}$ and $\widehat{\Theta}^{(k)}$ are calculated as shown.

$$\widehat{\eta}^{(1)} = \frac{1}{n_1 + \lambda_0} n_1 \left(\overline{g}^{(1)} + \lambda_0 \eta_0\right), \tag{18}$$

$$\widehat{\Theta}^{(1)} = (n_1 + \lambda_0)\Theta^{(1)}, \tag{19}$$

Samples $\eta_{(k)}$ follows Gaussian distribution. The $\widehat{\eta}_{i_k}^{(k)}$ (the value at the row i_k in the kth $\widehat{\eta}^{(k)}$) and $\widehat{\Theta}_{i_k}^{(k)}$ (the value at the row i_k in the kth $\widehat{\Theta}^{(k)}$) are calculated as shown.

$$\widehat{\Theta}_{i_1}^{(1)} = \delta_\varepsilon \sum_{i_2=1}^{n_2} \sum_{i_3=1}^{n_3} q_i \left(g_{i_2}^{(2)} \otimes g_{i_3}^{(3)}\right) \left(g_{i_2}^{(2)} \otimes g_{i_3}^{(3)}\right)^T + \Theta^{(1)}, \tag{20}$$

$$\widehat{\eta}_{i_1}^{(1)} = \left(\Theta_{i_1}^{(1)}\right)^{-1} \left(\delta_\varepsilon \sum_{i_2=1}^{n_2} \sum_{i_3=1}^{n_3} x_i \left(g_{i_2}^{(2)} \otimes g_{i_3}^{(3)}\right)\right) + \left(\widehat{\Theta}_{i_1}^{(1)}\right)^{-1} \Theta^{(1)} \eta^{(1)}, \tag{21}$$

Samples $g_{i_k}^{(k)}$ follows Gaussian distribution. Based on placing a flexible conjugate Gamma prior over δ_ε.

$$\delta \sim Gamma(p_0, q_0), \tag{22}$$

where δ represents precision, p_0 and q_0 are the shape parameter and rate parameter. In addition, $\widehat{p}_0$ and $\widehat{q}_0$ are calculated by

$$\widehat{p}_0 = \frac{1}{2} \sum_{i \in \Phi} q_i + p_0, \tag{23}$$

$$\widehat{q}_0 = \frac{1}{2} \sum_{i \in \Phi} (x_i - \widehat{x}_i)^2 + q_0, \tag{24}$$

where Φ is the index set of the observed entries.

4. Methodology

This section describes the newly proposed hierarchical methodology for VTF data prediction. Section 4.1 presents the hierarchical steps and overall framework from a macro perspective, while Section 4.2 delves into the detailed application of each step in the hierarchical methodology for VTF prediction.

4.1. The whole framework

The proposed hierarchical prediction methodology for VTF prediction is illustrated in Fig. 5. This paper calculates the VTF time series based on AIS data, including time stamps and vessel trajectories. The original 1D VTF dataset is then rearranged as a 3D tensor, and the proposed methodology consists of three steps.

The first step is to decompose each matrix (hour of the day and day) in the VTF tensor using the BEMD method to obtain the $BIMF_1$ and residues components, which represent high- and low-frequency components, respectively. These components are then used to construct high- and low-frequency tensors.

The low-frequency tensor is composed of stable data, which can be predicted directly by the BGCP method to achieve high-precision prediction results. However, the data in the high-frequency tensor is irregular and random, which means that the use of the BGCP method to predict it would not yield high-precision results. Therefore, this paper measures the similarity of each VTF matrix in the high-frequency tensor based on the DTW method. The VTF tensor is converted into 1D

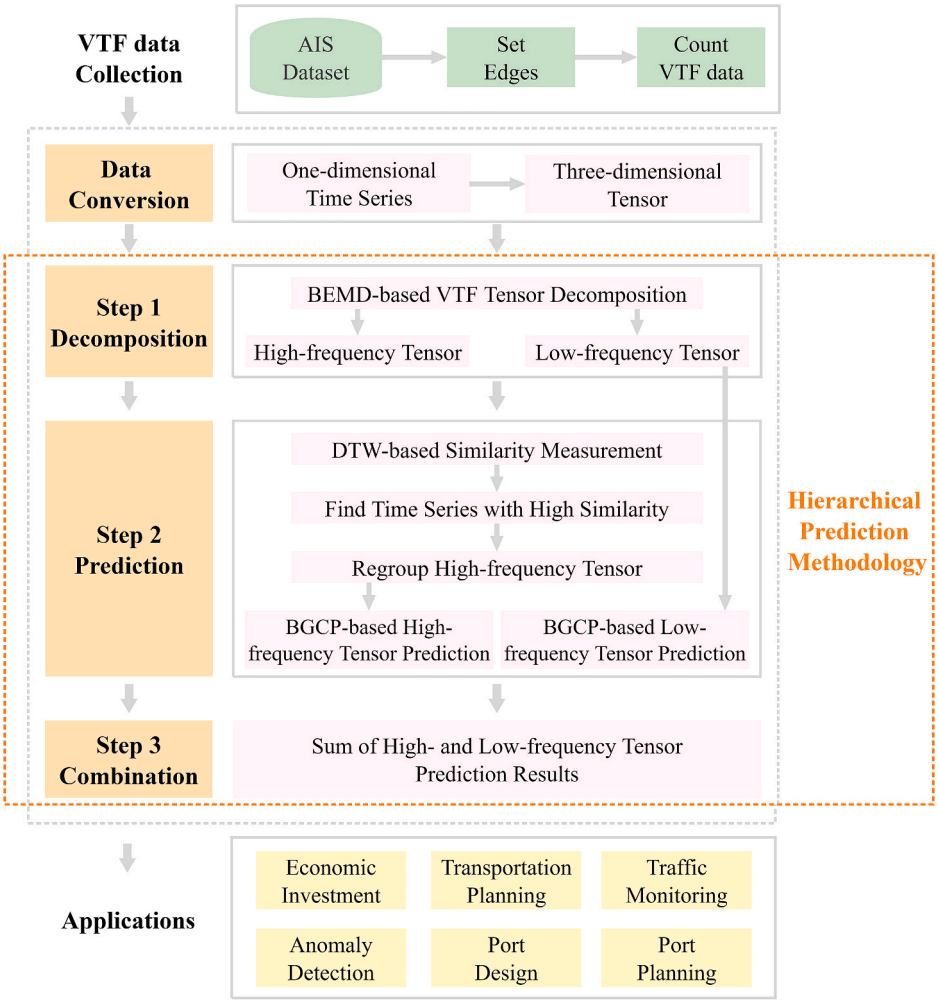


Fig. 5. The flowchart of our proposed hierarchical prediction methodology.

sequence data, and the similarity distance is calculated between the last sequence data and other sequence data, with the last data sequence representing the prediction data node. The new high-frequency tensor is regrouped based on the similarity measurement results, resulting in higher self-similarity and decreasing data irregularity and volatility. The BGCP method performs better prediction accuracy and stability when used to predict the new high-frequency tensor.

Finally, the prediction results of high- and low-frequency tensors are summed to obtain the prediction result. The proposed hierarchical methodology takes into account the spatial and temporal characteristics of VTF while addressing the impact of data volatility on prediction accuracy and stability. Thus, the BEMD-DTW-BGCP method is suitable for complex VTF prediction tasks.

4.2. The proposed hierarchical prediction method

Following the overview of the hierarchical framework in Section 4.1, this section describes the three steps (i.e., N1–N3) within the context of VTF prediction.

N1. *VTF matrix decomposition based on BEMD.*

The original 1D spatiotemporal VTF time series is rearranged as a 3D tensor (hour × day × week). Hence, the resulting tensor consists of several VTF matrices (hour × day). Using the BEMD method, each VTF matrix in tensor is decomposed into high-frequency component (i.e., C_h) and low-frequency component (i.e., C_l). The decomposed high- and low-frequency matrices form high- and low-frequency tensors, respectively. Fig. 6 shows the high- and low-frequency components obtained when the VTF matrix is decomposed into two layers (i.e., $BIMF_1$ and residue).

N2. Similarity grouping and tensor prediction.

To address the prediction difficulty in high-frequency tensors, the inherent self-similarities between VTF matrices of each week within the high-frequency tensor are extracted and fully taken into account. Specifically, each week of the VTF matrix $C_h(i)$ from the high-frequency tensor is rearranged as 1D time series $T_h(i)$, respectively. In particular, i represents ith (week) VTF matrix. The DTW distance between the last time series (to be predicted) and other time series is calculated by the following content, $DTW_{vec} = \{DTW(T_h(1), T_h(N),), DTW(T_h(2), T_h(N)), \cdots, DTW(T_h(N-1), T_h(N))\}$, and N denotes the number of VTF matrices from the high-frequency tensor.

It is crucial to reconstruct the reconstruction of a new and appropriate high-frequency tensor to be used with the BGCP method for achieving accurate prediction results. A suitable threshold is needed to be set to measure the similarity of the other matrices to the $C_h(N)$ in this paper. This threshold determines which matrices are similar enough to $C_h(N)$ to form a more suitable high-frequency tensor for the BGCP method to predict. Within the proposed framework, only high-similarity segments are extracted to guarantee accuracy and robustness.

N3. Combination.

The high-frequency tensor is predicted using the newly proposed BEMD-DTW-BGCP hierarchical methodology to improve prediction accuracy and robustness. The prediction results are obtained by integrating the predicted high- and low-frequency tensors. The proposed hierarchical prediction method can generate a more accurate result by benchmarking with 17 state-of-the-art VTF prediction models, as demonstrated in Section 5.

5. Experimental results and discussion

To measure the prediction accuracy and robustness of the proposed hierarchical methodology, a comparative analysis with 17 established prediction methods is conducted based on the realistic VTF time series, including tensor factorisation methods (i.e., HaLRTC, TF-ALS, TRTF, BATF, BGCP, and BEMD-BGCP), traditional mathematical methods (i.e., GM(1,1) and ARIMA), ML methods (i.e., SVM), NN methods (i.e., BPNN and WNN), and DL methods (i.e., RNN, LSTM, GRU, Bi-LSTM, Bi-GRU, and CNN-LSTM).

5.1. Experimental dataset description

The VTF data is derived from AIS data, including the time stamp and vessel trajectory, obtained from two water areas: Wuhan Yangtze River Bridge (WYRB) and Caofeidian District (CD). When a vessel's trajectory passes through a specific water area, the VTF of that water area is increased by one. The WYRB and CD areas represent inland and coastal waters, respectively. The effectiveness of the proposed hierarchical prediction model is validated through these two experiments, which involve different geographical features. In practical application scenarios, AIS devices are only responsible for transmitting and receiving data signals and do not correct the accuracy of data information. Consequently, there may be errors in the data collected by AIS-receiving devices. If original AIS data is used directly to calculate VTF, it will cause significant errors. Therefore, it is necessary to preprocess the data to obtain high-quality data. Detailed information on AIS data preprocessing and VTF calculation processes will be presented in Sections 5.1.1 and 5.1.2, respectively.

5.1.1. AIS data preprocessing

This paper focuses on the preprocessing process of AIS data, mainly

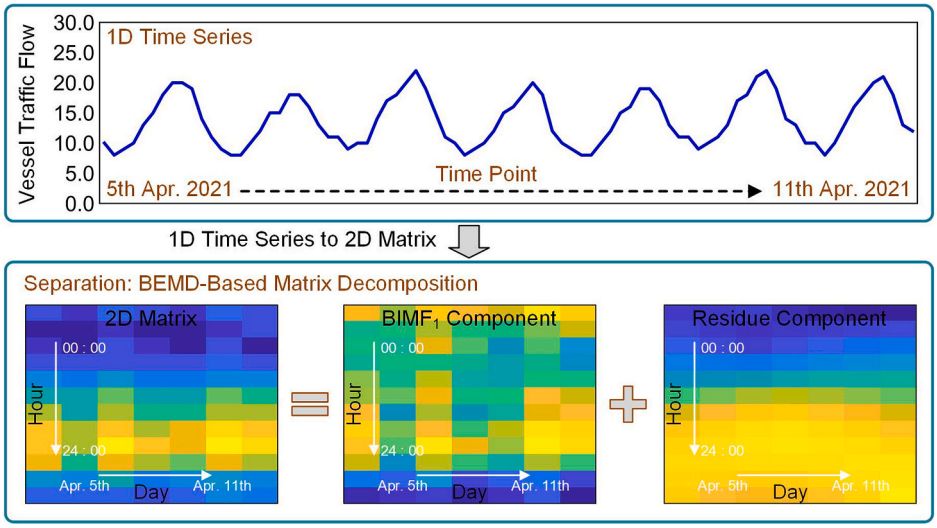


Fig. 6. Visual effect of VTF matrix decomposition based on the BEMD method. $BIMF_1$ and Residue represent high- and low-frequency components, respectively.

addressing the issue of noise in the original trajectory data and preparing for calculating VTF data. As shown in Fig. 7 (a), the noise data is distributed among other channels, adversely impacting the accuracy of VTF data statistics in those channels. Therefore, before conducting VTF counting, the data preprocessing work involves denoising the original trajectory and obtaining a high-quality trajectory, visually illustrated in Fig. 7 (b).

The denoising process of the original vessel trajectory data consists of two aspects: identifying noisy data and reconstructing noisy data. The clustering algorithm (Yang et al., 2022) can effectively solve this problem. The essence is to divide all trajectory points into two categories: normal points and noise points. Density-Based Spatial Clustering of Applications with Noise (DBSCAN) (Yu et al., 2022) is used in this study as it is an effective and widely used clustering algorithm. The DBSCAN algorithm classifies trajectory points into three categories, including core points, boundary points, and noise points. Both core and boundary points belong to normal data, and their classification depends on two super parameters, *Eps* and *Minpts*. A trajectory point is considered a core point if it has more than *Minpts* data points within the radius *Eps*. Conversely, if there are fewer than *Minpts* data points, the trajectory point is defined as a boundary point, indicating that it falls within the area of a core point. If a trajectory point is neither a core point nor a boundary point, it is classified as a noise point.

The DBSCAN algorithm can effectively identify noise points in the original vessel trajectory data, as shown in Fig. 8 and Fig. 9. If noise data is directly removed, it will affect the distribution structure of vessel trajectories. Hence, this paper will use the time information of noise points and the trajectory points of adjacent times to repair the noise data, which essentially involves a linear interpolation process. It can be expressed as follow,

$$l_t = l_{t-1} + (t-(t-1)) \times \left(\frac{l_{t+1} - l_{t-1}}{(t+1)-(t-1)} \right) \tag{25}$$

where t is the current time of the noise point. $L_{t\text{-}1}$ and l_{t+1} represent the longitude (or latitude) data of two adjacent time points t-1 and t+1 of the noise data, respectively. L_t is the result (longitude or latitude) of linear interpolation calculation at time t. Fig. 8 (b) and Fig. 9 (b) illustrate the restoration results of noise data in the two experimental water areas. In conclusion, DBSCAN and linear interpolation algorithms effectively identify and repair vessel trajectory data with noise, thereby facilitating accurate VTF data statistics.

5.1.2. VTF data generation

High-quality AIS data can be used to calculate VTF time series within the WYRB and CD study areas, with a time span from April 5, 2021 to June 27, 2021. This paper has set two distinct edges to calculate VTF data based on the vessel navigation directions of each research area. The edges, labelled E_1, E_2, E_3, and E_4, are to count VTF data in these two water areas, as illustrated in Fig. 10 (a) and (b). Consequently, there are two sets of VTF data for each water area, resulting in four datasets for the prediction experiment (i.e., dataset E_1, dataset E_2, dataset E_3, and dataset E_4). Table 1 displays the longitude and latitude of all nodes (i.e., N_1, N_2, N_3, N_4, N_5, and N_6) in Fig. 10, with each edge determined by two nodes.

AIS data is analysed to track the movement of vessels passing through different cross-sections (i.e., E_1, E_2, E_3, and E_4) in specific research areas. The analysis is based on the time stamp, trajectory, and heading information provided in the AIS data. A day is divided into 12 2-h periods for traffic flow statistics. The trajectory of a vessel passing through a cross-section is used to determine the time node to which the flow data belongs. This process results in the collection of VTF data for 12 time nodes per day. The statistical period for VTF data is 84 days, starting from April 5, 2021 to June 27, 2021. Each VTF dataset comprises 1008 flow data points. The trends in the VTF time series are presented in Fig. 11.

The VTF time series from E_1 to E_4, which reflects actual vessel traffic, is based on a 12-week spatiotemporal dataset. The original 1D time series data is transformed into a three-dimensional tensor with dimensions of 12 × 7 × 12. The first dimension indicates 12 2-h periods in a day, while the second dimension represents seven days a week. Therefore, the flow data for each week form a VTF matrix of 12 × 7. The final tensor data consists of 12 VTF matrices and is presented visually in Fig. 12.

Fig. 12 depicts the visualisation of VTF tensors from the four distinct research datasets, revealing that the distribution of VTF data is irregular due to the highly volatile and random nature of flow data over time. To enhance prediction accuracy, this study utilises the BEMD method to decompose each matrix in the VTF tensor into high- and low-frequency

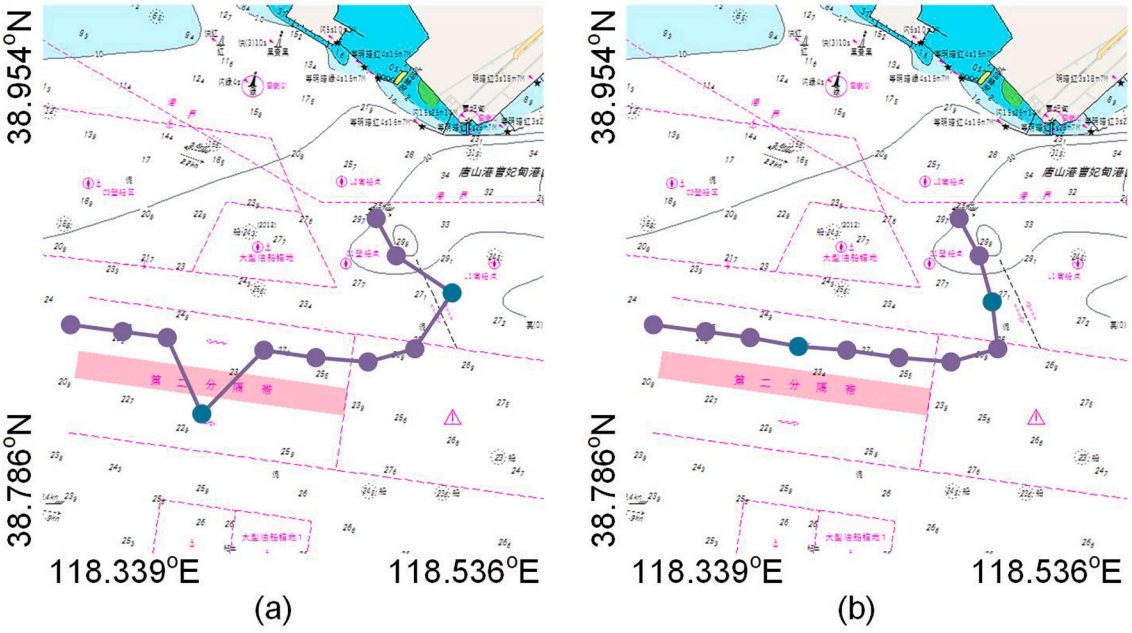


Fig. 7. The visual illustration of trajectory data denoising, (a) vessel trajectory data with noise and (b) trajectory data after removing noise. In particular, the cyan-blue dots denote noisy data.

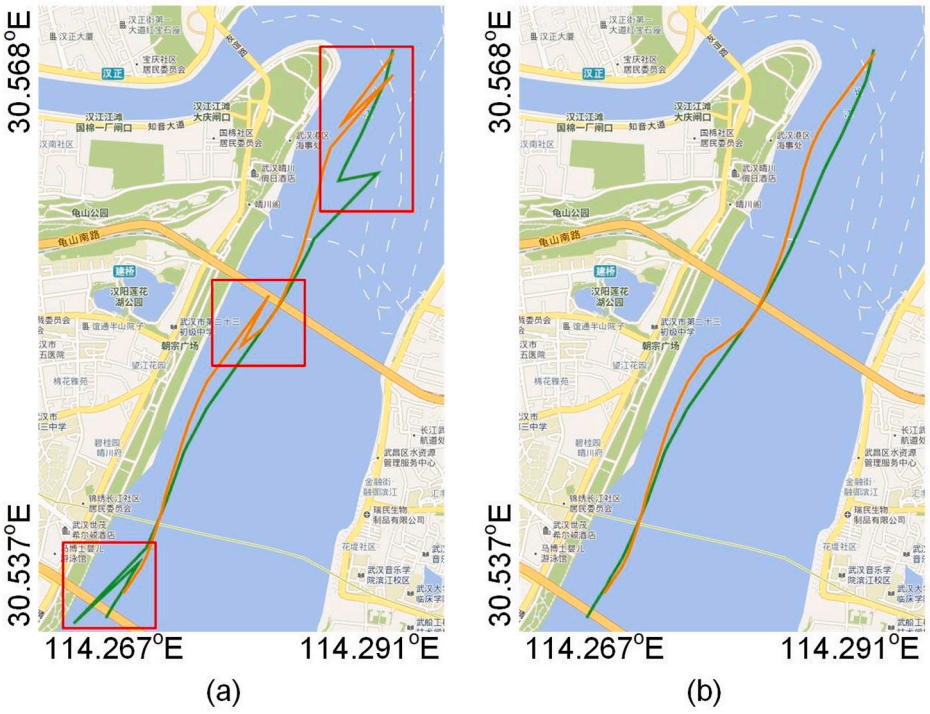


Fig. 8. The visual illustration of vessel trajectory data denoising in the WYRB area, (a) trajectory data with noise and (b) trajectory data after removing noise. In particular, the red boxes represent noisy data.

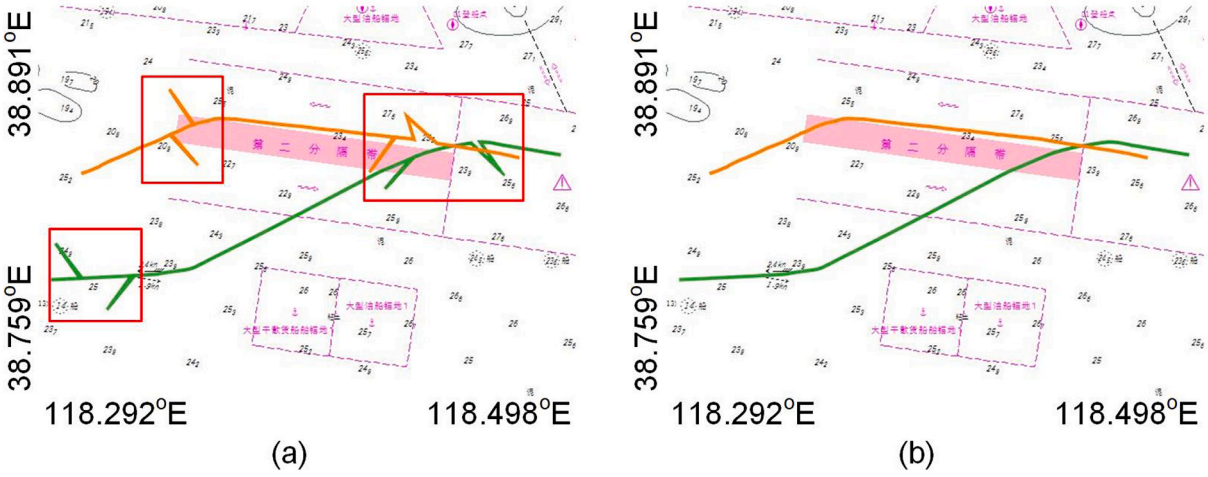


Fig. 9. The visual illustration of vessel trajectory data denoising in the CD area.

matrices. These decomposed matrices are then recombined into tensors, resulting in separate high-frequency and low-frequency tensors, as shown in Fig. 13.

5.2. Performance indexes on vessel traffic flow prediction

The fine-grained and coarse-grained results are combined to assess the overall prediction performance. For quantitatively evaluating the fine-grained result of 12 different time periods, the MRE and the SDRE are selected in this study. These two indexes can measure the prediction performance of the proposed prediction framework for VTF data at different time nodes. The mathematical formulations are defined as follows.

To quantitatively analyse the prediction performance,

$$MRE = \frac{1}{n}\sum_{i=1}^{n}\frac{|x-\hat{x}_i|}{x}, \tag{26}$$

$$SDRE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\frac{|x-\hat{x}_i|}{x} - MRE\right)^2}. \tag{27}$$

where n is the number of running times of each method. To ensure prediction robustness, each prediction model will run ten times in our experiment. x and $\hat{x}_i$ are the actual and ith predicted VTF data, respectively. The prediction performance can be represented by the MRE and SDRE values, where smaller values indicate better performance. To comprehensively evaluate the prediction performance of the proposed method under multiple time nodes by the coarse-grained result, the RMSE and MAPE indexes are chosen, defined below.

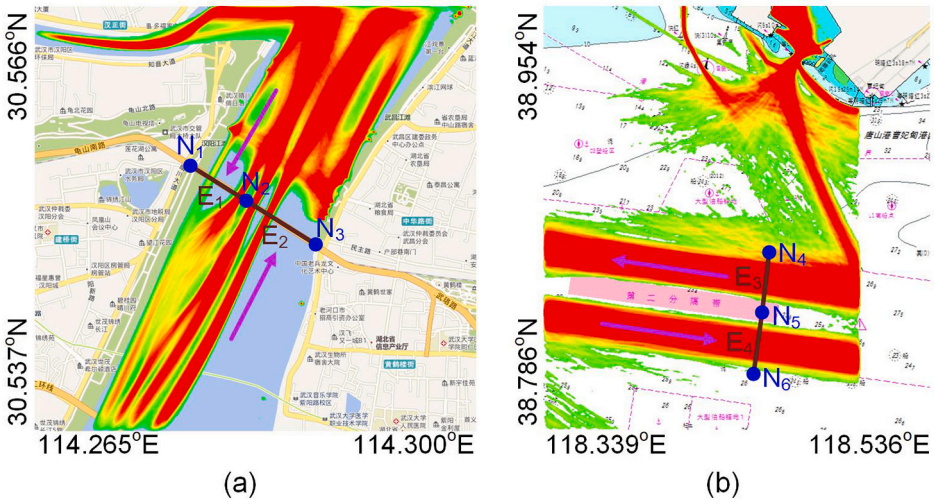


Fig. 10. Visualisation of realistic AIS-based vessel trajectories and different nodes, (a) visualisation in Wuhan Yangtze River Bridge and (b) visualisation in Caofeidian District. In particular, the pink arrow indicates the vessel's course in the channel. The brown line segments are two edges for counting VTF data, which is determined by the blue nodes.

Table 1
Statistical and geometrical information of six nodes in WYRB and CD water areas.

Water Areas	Node	Longitude(°)	Latitude(°)
WYRB	N_1	114.2772	30.5552
	N_2	114.2818	30.5527
	N_3	114.2877	30.5494
CD	N_4	118.4290	38.8567
	N_5	118.4259	38.8361
	N_6	118.4227	38.8154

$$RMSE = \sqrt{\frac{1}{m}\sum_{j=1}^{m}\left(y_j - \overline{y}_j\right)^2}, \tag{28}$$

$$MAPE = \frac{1}{m}\sum_{j=1}^{m}\left|\frac{y_j - \overline{y}_j}{y_j}\right|. \tag{29}$$

where y_j and $\overline{y}_j$ are jth actual VTF data and the average prediction result, respectively. m is the overall number of prediction data. The lower RMSE and MAPE values indicate a better prediction result.

5.3. Comparison with other prediction methods

This paper conducts comparative experiments by selecting 17 established methods to verify the superior performance of the proposed hierarchical method in VTF prediction tasks. According to the classification principles of VTF prediction methods described in Section 2, these comparison methods are divided into four categories. Specifically, it includes three traditional ML methods (i.e., GM(1,1), ARIMA, and SVM), two NN methods (i.e., BPNN and WNN), six DL methods (i.e., RNN, LSTM, GRU, Bi-LSTM, Bi-GRU, and CNN-LSTM), and six tensor factorisation methods (i.e., HaLRTC, TF-ALS, TRTF, BATF, BGCP, and BEMD-BGCP). To facilitate understanding, the design logic of this comparative experiment is presented in Fig. 14.

The following are the details of these 17 comparative methods.

(a) GM(1,1) (Wang et al., 2010): GM is widely used due to its ease of implementation (Chen and Huang, 2013). It uses the formula $y(t) = p(t) \times q(t)$ to perform data prediction tasks. In particular, $p(t)$ is a nonlinear function representing the factors affecting the time series data, while $q(t)$ is an accumulative function used to accumulate the influence of historical data to fit the data better.

(b) ARIMA (Liu et al., 2022): A time series prediction method combining the autoregressive and sliding average models containing three hyperparameters (i.e., p, d, and q). p is an autoregressive term representing how much historical data is used to predict future values. d is the difference item, indicating how many times the difference is used to make the time series data tend to a stable state. q is the moving average term, which reflects how many historical errors are used to predict future data.

(c) SVM (Zhang and Wu, 2022): It is a two-classification method which takes historical data as training samples and builds control functions to fit the change characteristics of historical data.

(d) BPNN (Yi et al., 2021) and WNN (Chen et al., 2021): BPNN is the most fundamental network model in the NN methods. This method iterates continuously through the training network to search for the changing characteristics of historical data, thereby completing the data prediction task.

WNN is an upgraded version of BPNN that replaces the sigmoid activation function with a wavelet function.

(e) RNN (Suo et al., 2020), LSTM (Yang et al., 2019) and GRU (Hussain et al., 2021): RNN can transmit the changing characteristics of historical data over time during training to the current moment, known as a feedforward network with memory ability. LSTM is an optimised version of RNN that utilises three gating structures (i.e., forget, input, and output gates). GRU is a simplified version of LSTM, having only two gating structures: the update gate and the reset gate, as opposed to the three gating structures in LSTM.

(f) Bi-LSTM (Li et al., 2021) and Bi-GRU (Huang et al., 2021): To further optimise network structure for LSTM and GRU, both forward and backward time-varying features are exploited in training the network to facilitate the development of bidirectional deep networks.

(g) CNN-LSTM (Narmadha and Vijayakumar, 2021): The hybrid structure, known as the CNN-LSTM or multi-view network, is a prominent deep learning model. It combines CNN and LSTM networks, where CNN captures spatial views, and LSTM captures temporal views. This innovative approach has been widely adopted by researchers to create a novel network architecture.

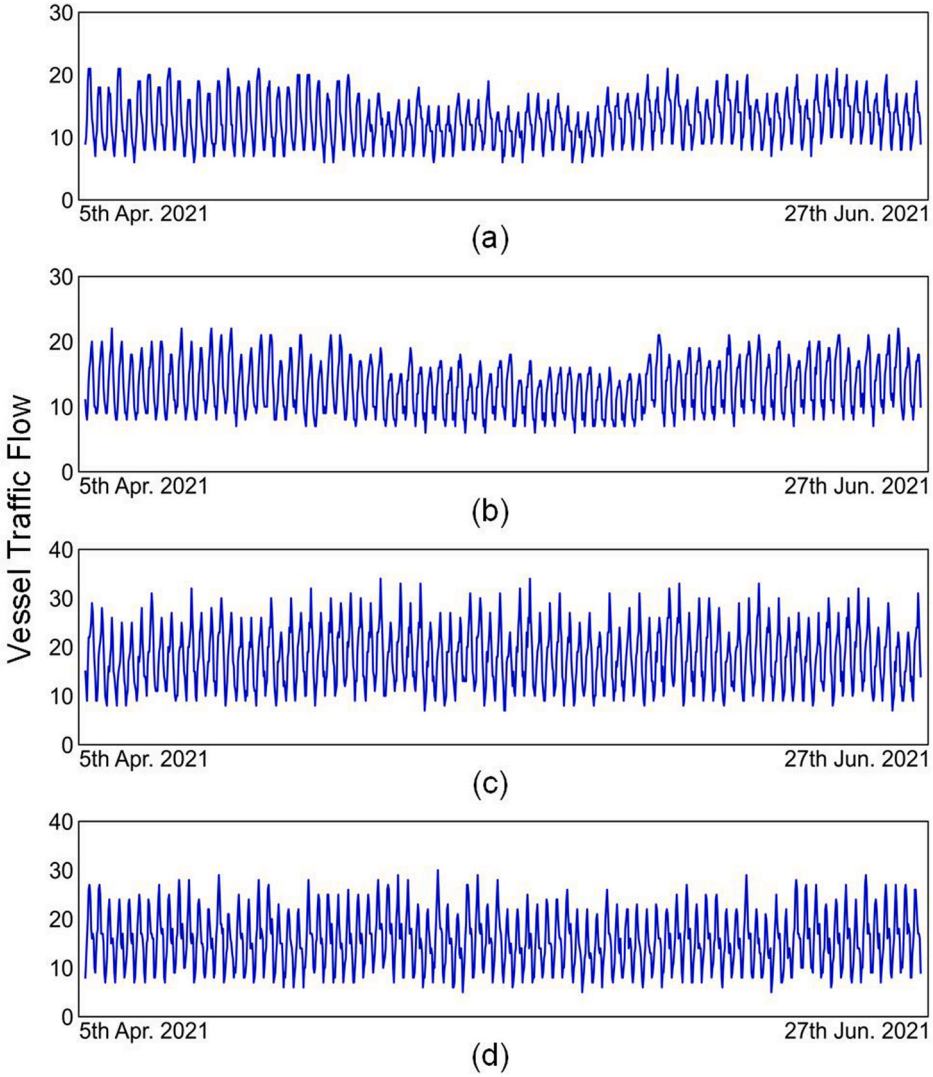


Fig. 11. The growing trends of VTF time series data in four different datasets from April 5, 2021 to June 27, 2021, (a) dataset E_1, (b) dataset E_2, (c) dataset E_3, and (d) dataset E_4.

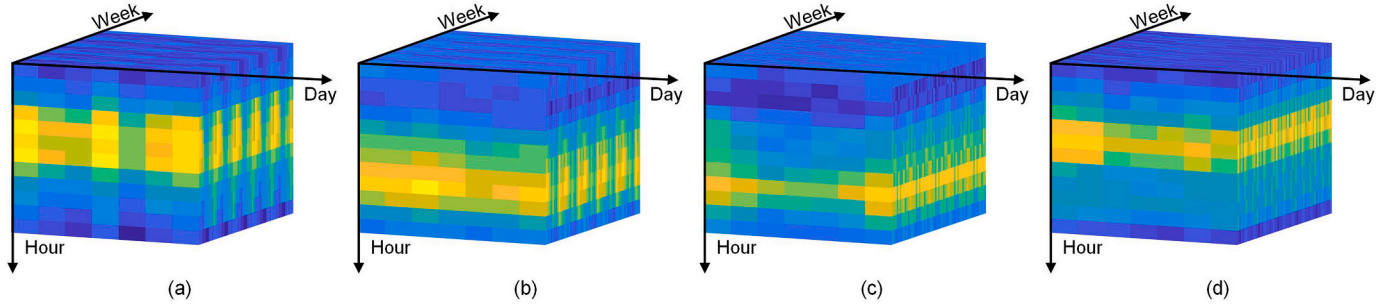


Fig. 12. Visualisation of VTF tensors (i.e., three-dimension) from four different datasets. From left to right: (a) dataset E_1, (b) dataset E_2, (c) dataset E_3, and (d) dataset E_4.

(h) HaLRTC (Liu et al., 2013), TF-ALS (Ma and Solomonik, 2022), TRTF (Baggag et al., 2021), BATF (Chen et al., 2019a) and BGCP (Chen et al., 2019b): They are five classic and commonly used tensor factorisation methods. During the execution of prediction tasks, the original 1D VTF time series dataset is rearranged into a 3D tensor (hour × day × week) as input data for these tensor factorisation methods. Based on the BGCP, BEMD-BGCP and BEMD-DTW-BGCP prediction frameworks are further developed to improve the prediction performance.

(i) BEMD-BGCP: It is proposed by combining the BEMD algorithm with the BGCP model as one of the comparative models. The BEMD algorithm is adopted to decompose the VTF matrix (hour × day) of each week from the VTF tensor into high- and low-matrices. Its essence is to decompose the original tensor into high- and low-frequency tensors. The BGCP model is applied to

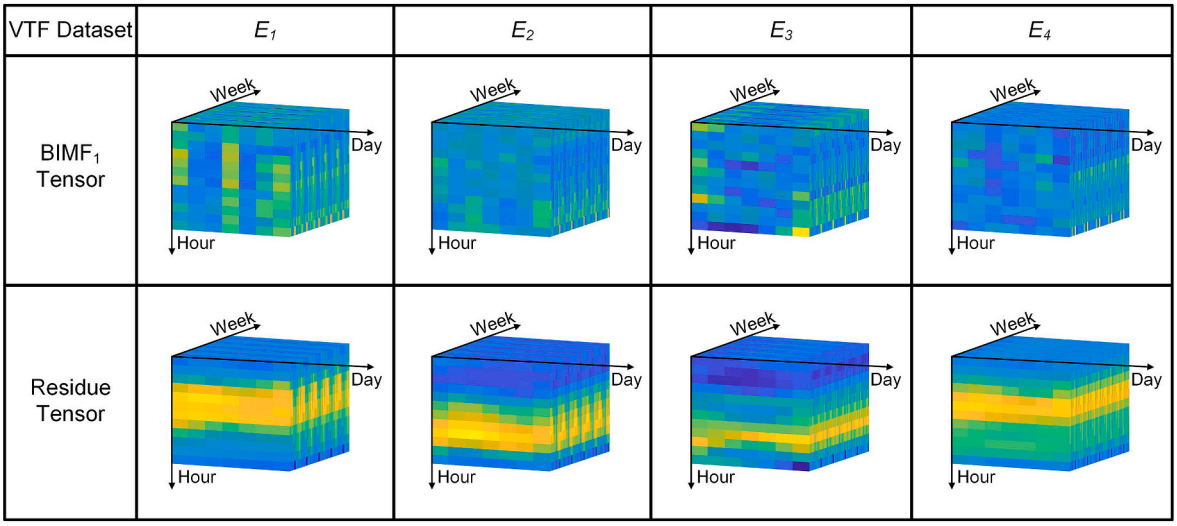


Fig. 13. The VTF tensor decomposition visualized result based on the BEMD method with $BIMF_1$ (a high-frequency tensor) and residue (a low-frequency tensor).

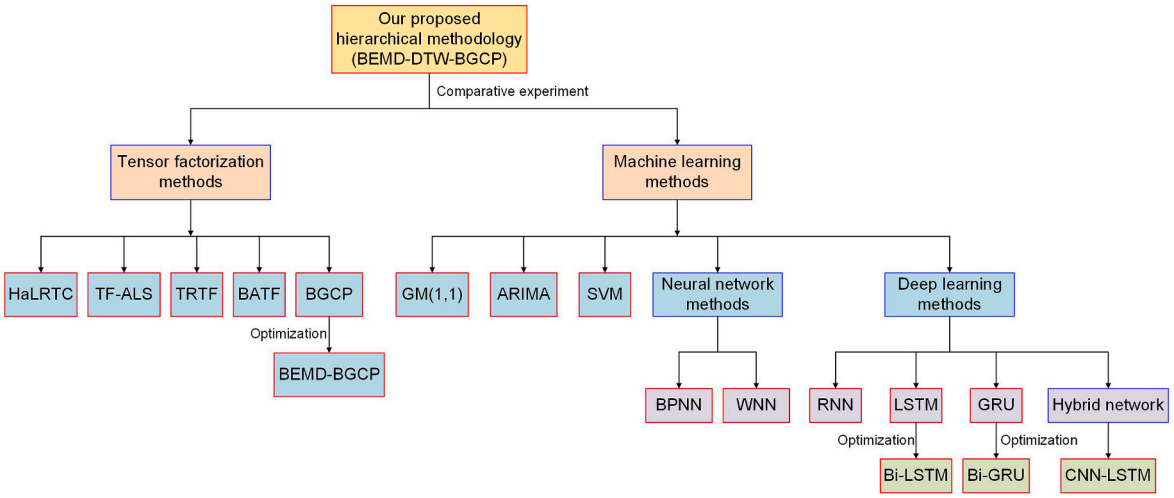


Fig. 14. The scheme of the comparative methods. The blue box indicates the method category, while the red box expresses the 18 models in the comparison experiment.

predict both high- and low-tensor directly. The prediction results are obtained by integrating the predicted high- and low-frequency tensors. Consequently, BEMD-BGCP can perform more efficiently than the single BGCP method. In particular, it is a simplified version of the hierarchical prediction method, and its role in comparative experiments is primarily to confirm the significance of the hierarchical method in calculating the self-similarity between VTF matrices using DTW with high-frequency tensors.

(j) BEMD-DTW-BGCP: The proposed hierarchical prediction framework extends the two-step BEMD-BGCP prediction method by leveraging the inherent self-similarities between VTF matrices of each week within high-frequency tensor. This three-step framework predicts the low-frequency tensor using the BGCP model while rearranging each week's matrix from the high-frequency tensor as a 1D time series. Additionally, the grouped segments with high DTW-based similarities are selected to generate more proper high-frequency tensors for the BGCP method to enhance prediction accuracy. Finally, the predicted high- and low-frequency tensors are integrated to obtain the prediction results.

5.4. Prediction results of different methods

The VTF time series data on the April 27, 2021 with 12 VTF data points (i.e., from time points 997 (00:00–02:00) to 1008 (22:00–24:00)) are used as the predicted datasets to access the prediction results.

To minimise the effect of randomness, 18 prediction models are executed ten times, and the MRE and SDRE values are determined based on the results of these ten results. The fine-grained quantitative results of 18 different prediction models for 12 time nodes are shown and compared in Figs. 15–18. The results measure the prediction accuracy and robustness for 12 different VTF data nodes on April 27, 2021. Specifically, the experiments in Figs. 15–18 are based on datasets E_1 to E_4, respectively. Additionally, to comprehensively evaluate the prediction accuracy and robustness of the newly proposed method across multiple time nodes, the coarse-grained indexes RMSE and MAPE are applied as quality metrics, and the results of different indicators are presented in Table 2.

According to the experimental results in Figs. 15–18 and Table 2 and it is evident that the GM(1,1) and ARIMA methods perform poorly in terms of prediction accuracy under almost all conditions. In particular, the GM(1,1) method has the worst effect in all prediction experiments. On the other hand, the SVM method generally yields more precise

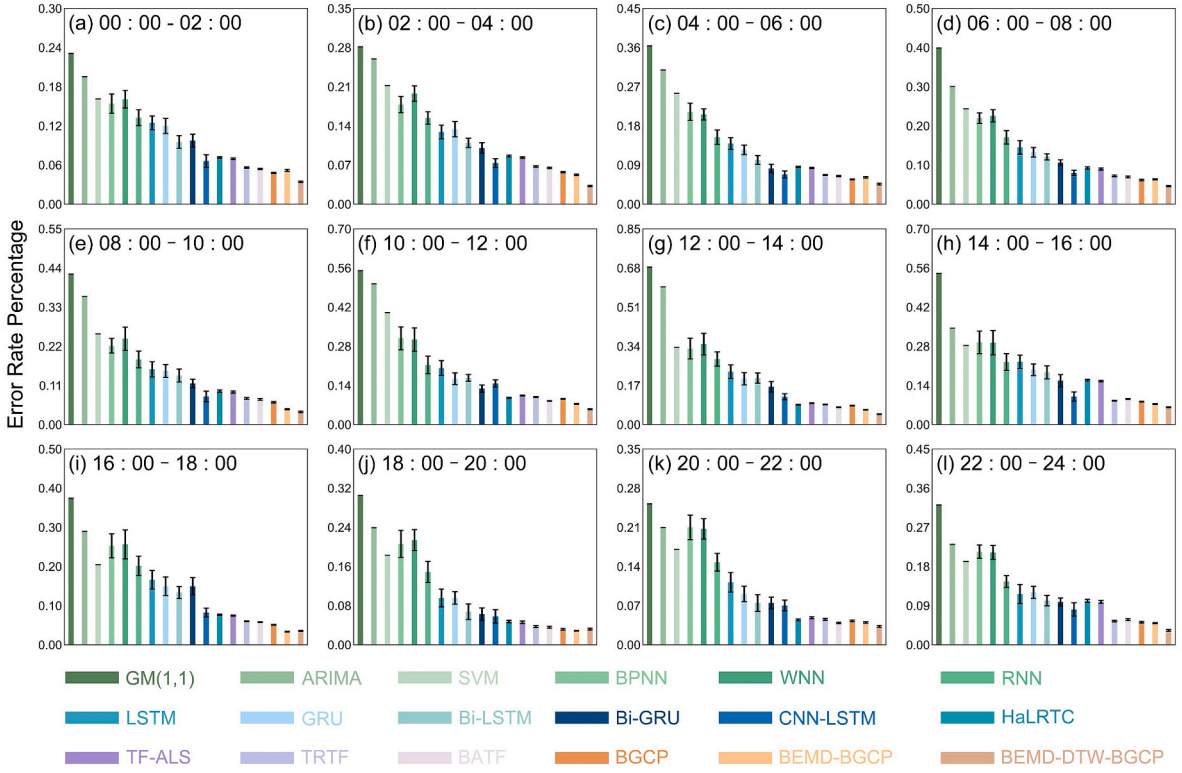


Fig. 15. Prediction results (MRE ± SDRE) of VTF data based on different methods (i.e., GM(1,1), ARIMA, SVM, BPNN, WNN, RNN, LSTM, GRU, Bi-LSTM, Bi-GRU, CNN-LSTM, HaLRTC, TF-ALS, TRTF, BATF, BGCP, BEMD-BGCP and BEMD-DTW-BGCP) in dataset E_1.

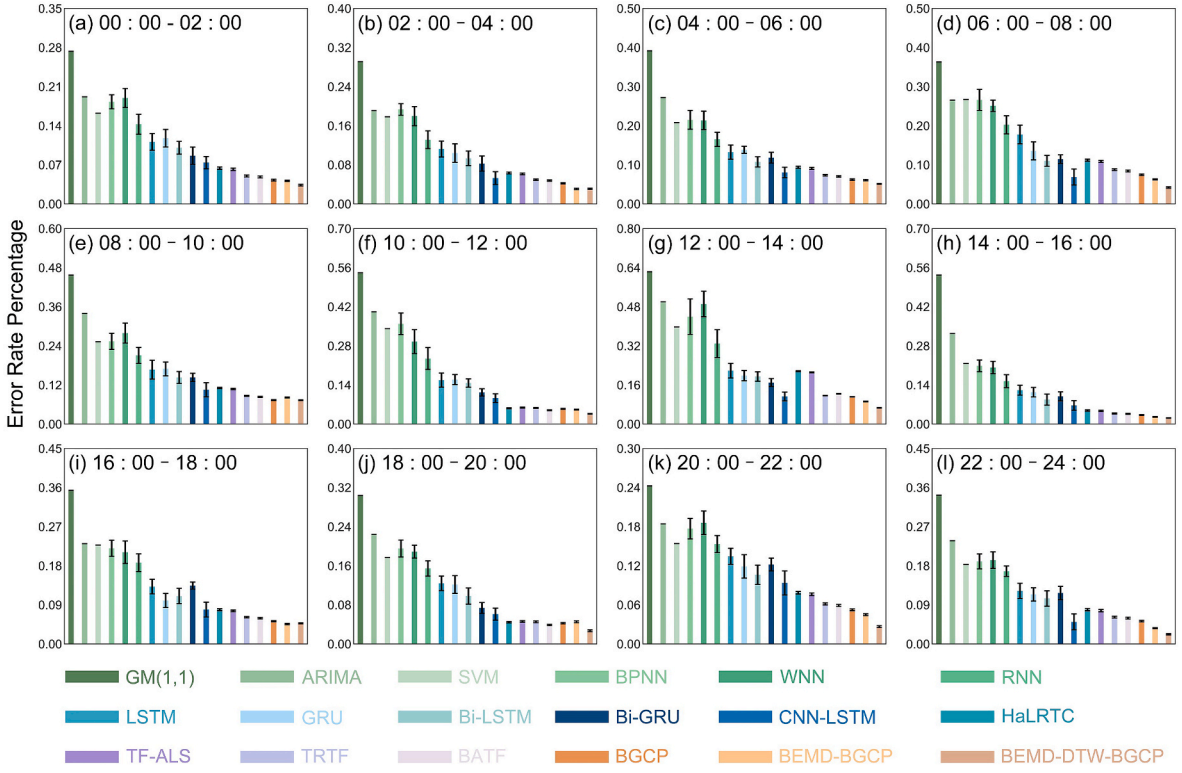


Fig. 16. Prediction results (MRE ± SDRE) of VTF data based on different methods in dataset E_2.

prediction results than the GM(1,1) and ARIMA methods, although it performs worse than ARIMA in some specific time nodes (e.g., Fig. 17 (a), (c), (e), (f), and Fig. 18 (f), (g), and (k)). The NN methods (i.e., BPNN and WNN) can predict the future trend based on the development characteristics of historical VTF data. However, the prediction results are worse than those of the SVM method in most cases due to the

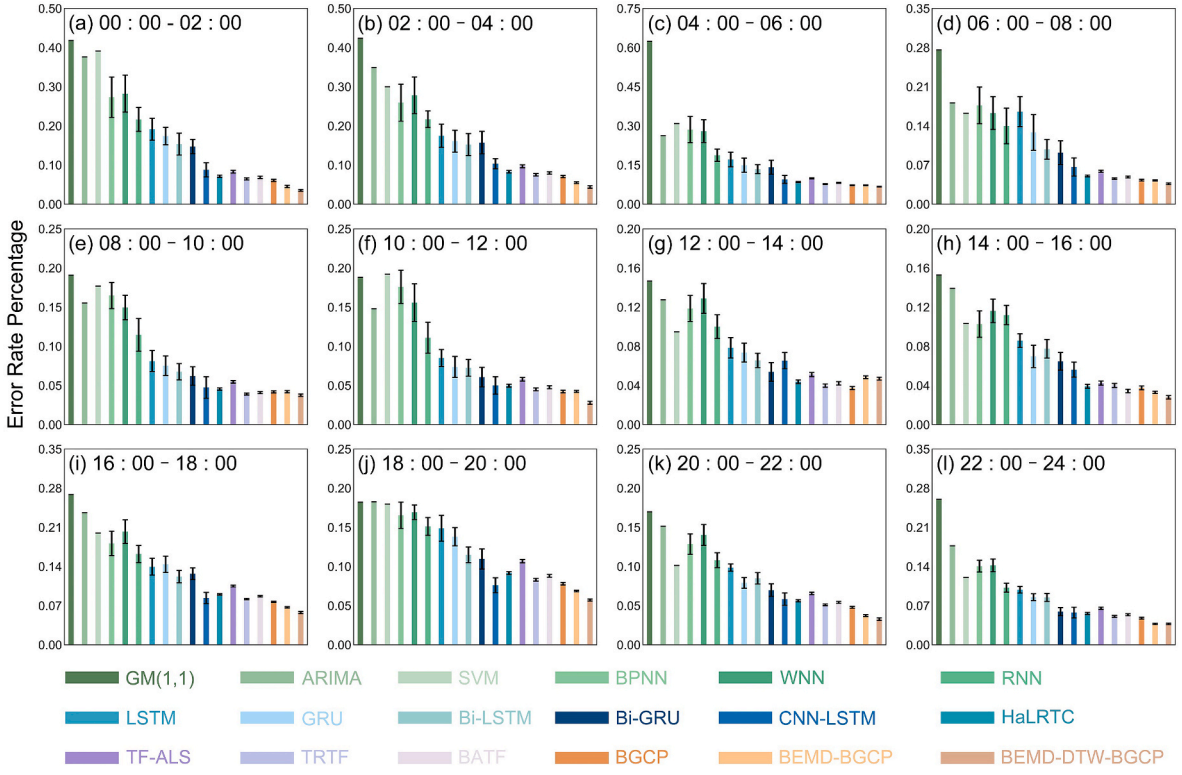


Fig. 17. Prediction results (MRE ± SDRE) of VTF data based on different methods in dataset E_3.

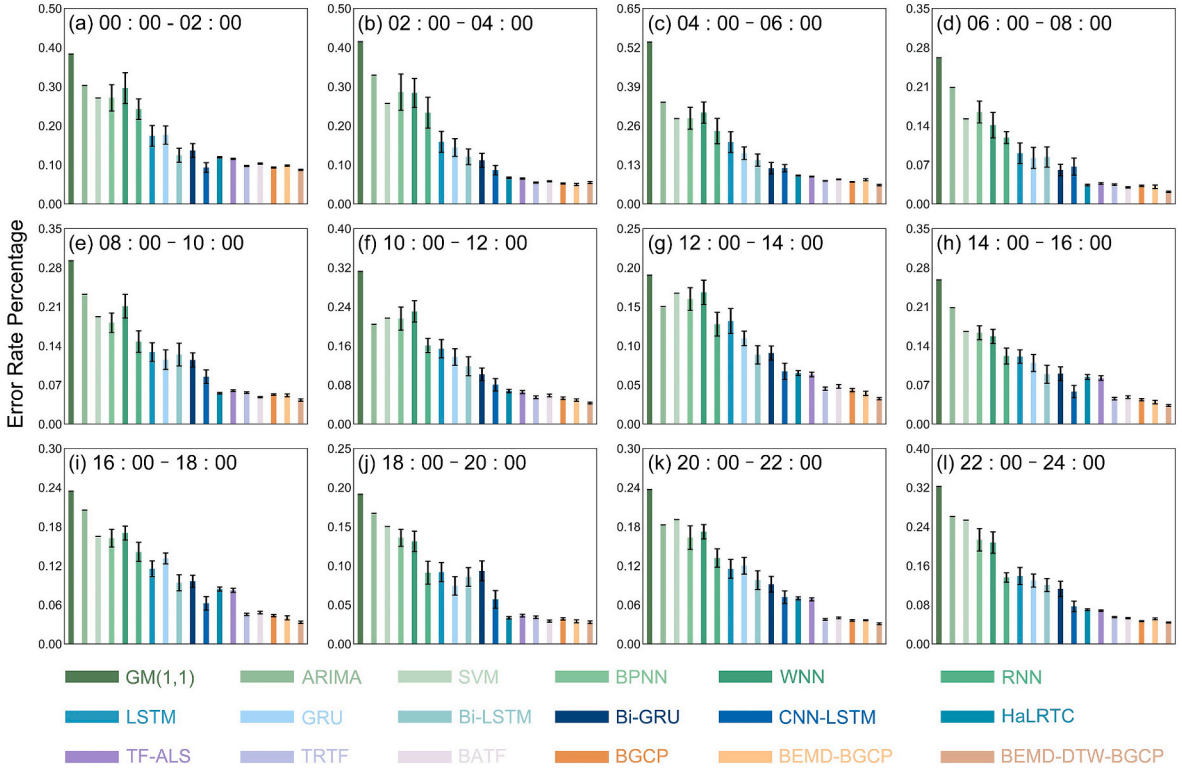


Fig. 18. Prediction results (MRE ± SDRE) of VTF data based on different methods in dataset E_4.

substantial volatility of VTF data. The WNN and BPNN methods exhibit similar prediction accuracy.

DL methods (RNN) are more effective in learning the characteristics of highly volatile data, especially since the RNN method can add historical data features to the current time point for learning based on the hidden state, thereby significantly improving prediction accuracy for time series data. In comparison to NN methods, RNN has better prediction accuracy. LSTM and GRU networks can mitigate the gradient

Table 2
Comparison of 18 methods against the RMSE and MAPE indexes for VTF data on April 27, 2021 in WYRB and CD.

Dataset		E_1		E_2		E_3		E_4	
Evaluation Metrics		RMSE	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE	MAPE
Model	GM(1, 1)	6.5340	0.3940	5.9901	0.3932	4.8124	0.2751	5.3388	0.3030
	ARIMA	5.4132	0.3200	4.3274	0.2803	3.9393	0.2071	3.9868	0.2325
	SVM	3.8375	0.2412	3.5088	0.2305	3.6890	0.1942	3.5599	0.2056
	BPNN	3.6147	0.2335	3.7000	0.2418	3.3803	0.1811	3.4877	0.2001
	WNN	3.7100	0.2391	3.7425	0.2402	3.4805	0.1840	3.6622	0.2061
	RNN	2.8289	0.1801	2.8120	0.1860	2.7928	0.1435	2.7554	0.1579
	LSTM	2.4386	0.1535	2.0699	0.1426	2.4657	0.1266	2.4408	0.1356
	GRU	2.1629	0.1399	1.9101	0.1327	2.2749	0.1126	2.2020	0.1249
	Bi-LSTM	2.0406	0.1250	1.7341	0.1170	2.0130	0.1023	1.9545	0.1078
	Bi-GRU	1.7518	0.1115	1.6731	0.1146	1.9196	0.0955	1.8096	0.1016
	CNN-LSTM	1.3634	0.0858	1.1555	0.0783	1.3987	0.0706	1.3910	0.0770
	HaLRTC	1.3293	0.0869	1.4243	0.0872	1.3621	0.0635	1.2272	0.0704
	TF-ALS	1.3355	0.0866	1.3940	0.0854	1.5866	0.0739	1.2141	0.0695
	TRTF	1.0440	0.0672	0.9633	0.0659	1.2399	0.0578	0.9314	0.0531
	BATF	0.9752	0.0638	0.9507	0.0635	1.3036	0.0606	0.9384	0.0538
	BGCP	0.9499	0.0602	0.8691	0.0579	1.1703	0.0547	0.8853	0.0502
	BEMD-BGCP	0.8026	0.0531	0.7771	0.0517	1.0615	0.0493	0.8596	0.0496
	BEMD-DTW-BGCP	**0.6246**	**0.0410**	**0.6051**	**0.0401**	**0.9103**	**0.0424**	**0.7326**	**0.0428**

vanishing problem to ensure robust prediction, with the latter being a simplified version of the former. Consequently, LSTM and GRU outperform the original RNN regarding VTF time series data prediction accuracy. LSTM and GRU belong to unidirectional networks that only retain past information. Bi-directional networks can preserve past and future information at any given moment by combining two hidden states. Overall, the prediction accuracy and stability of Bi-LSTM and Bi-GRU surpass those of the original LSTM and GRU. CNN-LSTM is a hybrid DL network structure that can simultaneously capture the spatial and temporal characteristics of VTF time series data, leading to less prediction error than NN methods (i.e., BPNN and WNN) and other DL methods (i.e., RNN, LSTM, GRU, Bi-LSTM, and Bi-GRU).

This paper's research focuses on tensor factorisation prediction methods. These methods transform 1D time series data into 3D tensors, allowing for the analysis of historical data from a global perspective. It is a significant advantage compared to other prediction methods. According to the coarse-grained quantitative evaluation results presented in Table 2, tensor factorisation methods outperform traditional ML, NN, and single DL methods across all four VTF datasets in terms of predictive performance. However, the HaLRTC method has higher RMSE and MAPE than the hybrid DL method (CNN-LSTM) in datasets E_1 and E_2. On the other hand, the TF-ALS method only has the smallest RMSE and MAPE compared to the CNN-LSTM in dataset E_4. It is worth noting that the TRTF, BATF, and BGCP methods consistently demonstrate superior predictive effects than CNN-LSTM in any situation. While ML, NN, and single DL methods only explore the temporal characteristics of historical VTF data, CNN-LSTM can capture the temporal attributes and simultaneously extract spatial features during data training, resulting in superior predictive performance compared to general DL and some tensor factorisation methods. Regarding comparison with HaLRTC, TF-ALS, TRTF and BATF methods, BGCP consistently achieves the smallest RMSE and MAPE in all datasets. Additionally, Figs. 15–18 reflect that BGCP obtains the minimum MRE and SDRE prediction results for most time nodes in each dataset, indicating its high accuracy and stable results in VTF prediction tasks. Through comparative experiments between these methods, it has been verified that the BGCP method outperforms the others in terms of higher and more consistent accuracy in dealing with the VTF prediction problem. This result justifies the adoption of the BGCP method as the foundation for optimising the hierarchical prediction approach proposed in this paper. Therefore, the hierarchical prediction method proposed in this paper improves and optimises the BGCP-based approach to achieve more accurate predictions in VTF data analysis.

The BEMD-BGCP method takes advantage of the BEMD method's features, which can decrease the uncertainty in the original data by decomposing non-stationary tensors into high- and low-frequency tensors. By using DTW to group segments with high similarities, the proposed hierarchical prediction methodology generates more suitable high-frequency tensors. These high-frequency tensors are then factorised using the BGCP method to obtain accurate and stable prediction results, making the prediction performance of the BEMD-DTW-BGCP method better than the BGCP and BEMD-BGCP methods. In summary, the experimental results show that the newly proposed hierarchical methodology is effective in predicting non-stationary VTF datasets in the maritime industry.

6. Conclusions and future research

To improve the VTF prediction accuracy for intelligent traffic management, a hierarchical prediction framework called BEMD-DTW-BGCP is proposed in this paper. Since the same dataset can be represented in different structures (i.e., 1D time series, 2D matrix, and 3D tensor), the initial 1D VTF time series is transformed into a 3D tensor (hour × day × week). Then, the original VTF matrix (hour × day) of each week within tensor is decomposed into high- and low-frequency matrices by the BEMD model. The original VTF tensor is similarly decomposed into high- and low-frequency tensors. To leverage the inherent self-similarities between VTF matrices of each week within high-frequency tensor, these matrices are further rearranged into 1D time series and select grouped segments with high similarities by DTW to form a more proper high-frequency tensor for the BGCP model. It can improve the prediction accuracy of tensors. The low-frequency tensor can be directly predicted using the BGCP model due to its solid mathematical regularity. The results of predicted high- and low-frequency tensors are combined to generate the prediction outcomes. The proposed BEMD-DTW-BGCP model is compared with 17 established prediction methods by analysing VTF data in the WYRB inland waters and the CD coastal waters to demonstrate its effectiveness and robustness by the higher accuracy and lower error. The proposed hierarchical prediction methodology holds tremendous potential for producing satisfactory performance in predicting maritime traffic flow.

Future research content could focus on developing a GPU-accelerated ensemble method, extending the applicability of the proposed hierarchical methodology to supervise VTF in a broader context. To enhance the method's efficiency in handling big data, it is imperative to implement the prediction method using a GPU parallel framework. The GPU-based parallel computing framework can map each data in the VTF tensor to different threads. The GPU-based parallel computing

framework can map each data in the VTF tensor to different threads. The hierarchical methodology can be encapsulated within multiple kernel functions, executed by each thread to predict future VTF data in the research waters. Such parallel computing can significantly improve prediction speed and facilitate real-time forecasting in large-scale datasets. By leveraging the power of GPU parallelism, the hierarchical method can efficiently handle vast amounts of data, making it more practical for real-world applications in VTF supervision. Additionally, the development trend of VTF in two adjacent geographical spaces may have a mutual influence. Therefore, it is crucial to explore further optimisations for the prediction method that will enable the collaborative prediction framework of VTF in adjacent geographical areas.

CRediT authorship contribution statement

Wenbin Xing: Investigation, Software, Validation, Formal analysis, Visualization, Writing – original draft. **Jingbo Wang:** Investigation, Software, Validation, Formal analysis, Visualization, Writing – original draft. **Kaiwen Zhou:** Investigation, Software, Validation, Formal analysis, Visualization, Writing – original draft. **Huanhuan Li:** Conceptualization, Methodology, Data curation, Software, Investigation, Supervision, Writing – original draft, Writing – review & editing. **Yan Li:** Conceptualization, Methodology, Data curation, Software, Investigation, Validation, Formal analysis, Visualization, Writing – original draft, Writing – review & editing. **Zaili Yang:** Conceptualization, Methodology, Resources, Supervision, Project administration, Funding acquisition, Writing – original draft, Writing – review & editing.

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.

Data availability

Data will be made available on request.

Acknowledgements

This work is supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (TRUST CoG 2019 864724) and Royal Society International Exchanges 2021 Cost Share (NSFC) (IEC\NSFC\211211).

References

Baggag, A., Abbar, S., Sharma, A., Zanouda, T., Al-Homaid, A., Mohan, A., Srivastava, J., 2021. Learning spatiotemporal latent factors of traffic via regularized tensor factorization: imputing missing values and forecasting. IEEE Trans. Knowl. Data Eng. 33, 2573–2587. https://doi.org/10.1109/TKDE.2019.2954868.

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. https://doi.org/10.1007/s10489-020-01716-1.

Bürkner, P.-C., Gabry, J., Vehtari, A., 2020. Approximate leave-future-out cross-validation for Bayesian time series models. J. Stat. Comput. Simulat. 90, 2499–2523. https://doi.org/10.1080/00949655.2020.1783262.

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. https://doi.org/10.1016/j.ins.2022.07.125.

Celikoglu, H.B., Cigizoglu, H.K., 2007. Public transportation trip flow modeling with generalized regression neural networks. Adv. Eng. Software 38, 71–79. https://doi.org/10.1016/j.advengsoft.2006.08.003.

Chan, K.Y., Dillon, T.S., 2013. On-Road sensor configuration design for traffic flow prediction using fuzzy neural networks and taguchi method. IEEE Trans. Instrum. Meas. 62, 50–59. https://doi.org/10.1109/TIM.2012.2212506.

Chen, C., Hu, J., Meng, Q., Zhang, Y., 2011. Short-time traffic flow prediction with ARIMA-GARCH model. In: 2011 Ieee Intelligent Vehicles Symposium (Iv). Ieee, New York, pp. 607–612. https://doi.org/10.1109/ICCE.2011.5722766.

Chen, C.-I., Huang, S.-J., 2013. The necessary and sufficient condition for GM(1,1) grey prediction model. Appl. Math. Comput. 219, 6152–6162. https://doi.org/10.1016/j.amc.2012.12.015.

Chen, J., Chen, H., Zhao, Y., Li, X., 2022. FB-BiGRU: a Deep Learning model for AIS-based vessel trajectory curve fitting and analysis. Ocean Eng. 266, 112898 https://doi.org/10.1016/j.oceaneng.2022.112898.

Chen, Q., Song, Y., Zhao, J., 2021. Short-term traffic flow prediction based on improved wavelet neural network. Neural Comput. Appl. 33, 8181–8190. https://doi.org/10.1007/s00521-020-04932-5.

Chen, W.-K., Lee, J.-C., Han, W.-Y., Shih, C.-K., Chang, K.-C., 2013. Iris recognition based on bidimensional empirical mode decomposition and fractal dimension. Inf. Sci. 221, 439–451. https://doi.org/10.1016/j.ins.2012.09.021.

Chen, X., He, Z., Chen, Y., Lu, Y., Wang, J., 2019a. Missing traffic data imputation and pattern discovery with a Bayesian augmented tensor factorization model. Transport. Res. C Emerg. Technol. 104, 66–77. https://doi.org/10.1016/j.trc.2019.03.003.

Chen, X., He, Z., Sun, L., 2019b. A Bayesian tensor decomposition approach for spatiotemporal traffic data imputation. Transport. Res. C Emerg. Technol. 98, 73–84. https://doi.org/10.1016/j.trc.2018.11.003.

Chen, X., Sun, L., 2022. Bayesian temporal factorization for multidimensional time series prediction. IEEE Trans. Pattern Anal. Mach. Intell. 44, 4659–4673. https://doi.org/10.1109/TPAMI.2021.3066551.

Cho, K., van Merrienboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., Bengio, Y., 2014. Learning Phrase Representations Using RNN Encoder-Decoder for Statistical Machine Translation. https://doi.org/10.48550/arXiv.1406.1078.

Comert, G., Bezuglov, A., 2013. An online change-point-based model for traffic parameter prediction. IEEE Trans. Intell. Transport. Syst. 14, 1360–1369. https://doi.org/10.1109/TITS.2013.2260540.

Dikshit, A., Pradhan, B., Santosh, M., 2022. Artificial neural networks in drought prediction in the 21st century-A scientometric analysis. Appl. Soft Comput. 114, 108080 https://doi.org/10.1016/j.asoc.2021.108080.

Do, L.N.N., Vu, H.L., Vo, B.Q., Liu, Z., Phung, D., 2019. An effective spatial-temporal attention based neural network for traffic flow prediction. Transport. Res. C Emerg. Technol. 108, 12–28. https://doi.org/10.1016/j.trc.2019.09.008.

Dong, Z., 2022. Prediction of ship traffic flow based on wavelet decomposition and LSTM. In: 2022 7th International Conference on Cloud Computing and Big Data Analytics (ICCCBDA). Presented at the 2022 7th International Conference on Cloud Computing and Big Data Analytics. ICCCBDA), pp. 88–93. https://doi.org/10.1109/ICCCBDA55098.2022.9778881.

Du, L., Gao, R., Suganthan, P.N., Wang, D.Z.W., 2022. Bayesian optimization based dynamic ensemble for time series forecasting. Inf. Sci. 591, 155–175. https://doi.org/10.1016/j.ins.2022.01.010.

EMSA, 2022. Preliminary Annual Overview of Marine Casualties and Incidents 2014-2020 [WWW Document]. URL. https://www.emsa.europa.eu/publications/item/4378-preliminary-annual-overview-of-marine-casualties-and-incidents-2014-2020.html (accessed May.30.22).

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. https://doi.org/10.1016/j.apenergy.2022.120261.

Gao, M., Shi, G.-Y., 2019. Ship spatiotemporal key feature point online extraction based on AIS multi-sensor data using an improved sliding window algorithm. Sensors 19, 2706. https://doi.org/10.3390/s19122706.

Goerlandt, F., Kujala, P., 2011. Traffic simulation based ship collision probability modeling. Reliab. Eng. Syst. Saf. 96, 91–107. https://doi.org/10.1016/j.ress.2010.09.003.

Gong, C., Zhang, Y., 2020. Urban traffic data imputation with detrending and tensor decomposition. IEEE Access 8, 11124–11137. https://doi.org/10.1109/ACCESS.2020.2964299.

Han, Y., Moutarde, F., 2016. Analysis of large-scale traffic dynamics in an urban transportation network using non-negative tensor factorization. Int. J. ITS Res. 14, 36–49. https://doi.org/10.1007/s13177-014-0099-7.

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. https://doi.org/10.1016/j.trc.2019.08.005.

He, W., Zhong, C., Sotelo, M.A., Chu, X., Liu, X., Li, Z., 2019. Short-term vessel traffic flow forecasting by using an improved Kalman model. Cluster Comput. 22, S7907–S7916. https://doi.org/10.1007/s10586-017-1491-2.

Hinton, G.E., Salakhutdinov, R.R., 2006. Reducing the dimensionality of data with neural networks. Science 313, 504–507. https://doi.org/10.1126/science.1127647.

Hou, W.-L., Jia, R.-S., Sun, H.-M., Zhang, X.-L., Deng, M.-D., Tian, Y., 2019. Random noise reduction in seismic data by using bidimensional empirical mode decomposition and shearlet transform. IEEE Access 7, 71374–71386. https://doi.org/10.1109/ACCESS.2019.2920021.

Huang, M., Zhu, M., Xiao, Y., Liu, Y., 2021. Bayonet-corpus: a trajectory prediction method based on bayonet context and bidirectional GRU. Digit. Commun. Netw. 7, 72–81. https://doi.org/10.1016/j.dcan.2020.03.002.

Huang, N.E., Shen, Z., Long, S.R., Wu, M.C., Shih, H.H., Zheng, Q., Yen, N.-C., Tung, C.C., Liu, H.H., 1998. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proceedings of the Royal Society of London. Series A: Math. Phys. Eng. Sci. 454, 903–995. https://doi.org/10.1098/rspa.1998.0193.

Hussain, B., Afzal, M.K., Ahmad, S., Mostafa, A.M., 2021. Intelligent traffic flow prediction using optimized GRU model. IEEE Access 9, 100736–100746. https://doi.org/10.1109/ACCESS.2021.3097141.

Jiang, P., Fan, Z., Pan, M., Hu, W., 2022. Research of traffic flow saturation on waters of the coastal ship routing system. Ocean Eng. 263, 112417 https://doi.org/10.1016/j.oceaneng.2022.112417.

Jiang, Y., Huang, G., Yang, Q., Yan, Z., Zhang, C., 2019. A novel probabilistic wind speed prediction approach using real time refined variational model decomposition and

conditional kernel density estimation. Energy Convers. Manag. 185, 758–773. https://doi.org/10.1016/j.enconman.2019.02.028.
Jin, H., Gu, Z.-M., Tao, T.-M., Li, E., 2021. Hierarchical attention-based machine learning model for radiation prediction of WB-bga package. IEEE Trans. Electromagn C. 63, 1972–1980. https://doi.org/10.1109/TEMC.2021.3075020.
Kaffash Charandabi, N., Gholami, A., Abdollahzadeh Bina, A., 2022. Road accident risk prediction using generalized regression neural network optimized with self-organizing map. Neural Comput. Appl. 34, 8511–8524. https://doi.org/10.1007/s00521-021-06549-8.
Kayacan, E., Ulutas, B., Kaynak, O., 2010. Grey system theory-based models in time series prediction. Expert Syst. Appl. 37, 1784–1789. https://doi.org/10.1016/j.eswa.2009.07.064.
Kong, D., Chen, Y., Li, N., Duan, C., Lu, L., Chen, D., 2020. Tool wear estimation in end milling of titanium alloy using NPE and a novel WOA-SVM model. IEEE Trans. Instrum. Meas. 69, 5219–5232. https://doi.org/10.1109/TIM.2019.2952476.
Li, H., Liu, J., Yang, Z., Liu, R.W., Wu, K., Wan, Y., 2020. Adaptively constrained dynamic time warping for time series classification and clustering. Inf. Sci. 534, 97–116. https://doi.org/10.1016/j.ins.2020.04.009.
Li, H., Ren, X., Yang, Z., 2023. Data-driven Bayesian network for risk analysis of global maritime accidents. Reliab. Eng. Syst. Saf. 230, 108938. https://doi.org/10.1016/j.ress.2022.108938.
Li, J., Xu, L., Li, R., Wu, P., Huang, Z., 2022. Deep spatial-temporal bi-directional residual optimisation based on tensor decomposition for traffic data imputation on urban road network. Appl. Intell. 52, 11363–11381. https://doi.org/10.1007/s10489-021-03060-4.
Li, H., Yang, Z., 2023. Incorporation of AIS data-based machine learning into unsupervised route planning for maritime autonomous surface ships. Transp. Res. Part E Logist. Transp. Rev. 176, 103171. https://doi.org/10.1016/j.tre.2023.103171.
Li, H., Yang, Z., 2023. Ship trajectory prediction based on machine learning and deep learning: A systematic review and methods analysis. Eng. Appl. Artif. Intell. Accepted.
Li, H., Yang, Z., 2023. Towards safe navigation environment: The imminent role of spatio-temporal pattern mining in maritime piracy incidents analysis. Reliab. Eng. Syst. Saf. 238, 109422. https://doi.org/10.1016/j.ress.2023.109422.
Li, L., Yang, Y., Yuan, Z., Chen, Z., 2021. A spatial-temporal approach for traffic status analysis and prediction based on Bi-LSTM structure. Mod. Phys. Lett. B 35, 2150481. https://doi.org/10.1142/S0217984921504819.
Li, M., Mou, J., Chen, P., Chen, L., van Gelder, P.H.a.J.M., 2023. Real-time collision risk based safety management for vessel traffic in busy ports and waterways. Ocean Coast Manag. 234, 106471 https://doi.org/10.1016/j.ocecoaman.2022.106471.
Li, H., Jiao, H., Yang, Z., 2023a. AIS data-driven ship trajectory prediction modelling and analysis based on machine learning and deep learning methods. Transp. Res. Part E Logist. Transp. Rev. 175, 103152. https://doi.org/10.1016/j.tre.2023.103152.
Li, H., Lam, J.S.L., Yang, Z., Liu, J., Liu, R.W., Liang, M., Li, Y., 2022. Unsupervised hierarchical methodology of maritime traffic pattern extraction for knowledge discovery. Transp. Res. Part C Emerg. Technol. 143, 103856. https://doi.org/10.1016/j.trc.2022.103856.
Li, Y., Liang, M., Li, H., Yang, Z., Du, L., Chen, Z., 2023. Deep learning-powered vessel traffic flow prediction with spatial-temporal attributes and similarity grouping. Eng. Appl. Artif. Intell. 126, 107012. https://doi.org/10.1016/j.engappai.2023.107012.
Li, Y., Liu, R.W., Liu, Z., Liu, J., 2019. Similarity grouping-guided neural network modeling for maritime time series prediction. IEEE Access 7, 72647–72659. https://doi.org/10.1109/ACCESS.2019.2920436.
Liang, M., Liu, R.W., Li, S., Xiao, Z., Liu, X., Lu, F., 2021. An unsupervised learning method with convolutional auto-encoder for vessel trajectory similarity computation. Ocean Eng. 225, 108803. https://doi.org/10.1016/j.oceaneng.2021.108803.
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. Transp. Syst. 23, 23694–23707. https://doi.org/10.1109/TITS.2022.3199160.
Lim, B., Zohren, S., 2021. Time-series forecasting with deep learning: a survey. Philos. Trans. R. Soc. A-Math. Phys. Eng. Sci. 379, 20200209 https://doi.org/10.1098/rsta.2020.0209.
Lin, M., You, Y., Wang, W., Wu, J., 2023. Battery health prognosis with gated recurrent unit neural networks and hidden Markov model considering uncertainty quantification. Reliab. Eng. Syst. Saf. 230, 108978. https://doi.org/10.1016/j.ress.2022.108978.
Liu, D., Chen, H., Tang, Y., Liu, C., Cao, M., Gong, C., Jiang, S., 2022. Slope micrometeorological analysis and prediction based on an ARIMA model and data-fitting system. Sensors 22, 1214. https://doi.org/10.3390/s22031214.
Liu, J., Musialski, P., Wonka, P., Ye, J., 2013. Tensor completion for estimating missing values in visual data. IEEE Trans. Pattern Anal. Mach. Intell. 35, 208–220. https://doi.org/10.1109/TPAMI.2012.39.
Liu, L., Shibasaki, R., Zhang, Y., Kosuge, N., Zhang, M., Hu, Y., 2023. Data-driven framework for extracting global maritime shipping networks by machine learning. Ocean Eng. 269, 113494 https://doi.org/10.1016/j.oceaneng.2022.113494.
Ma, C., Dai, G., Zhou, J., 2022. Short-term traffic flow prediction for urban road sections based on time series analysis and LSTM_BILSTM method. IEEE Trans. Intell. Transport. Syst. 23, 5615–5624. https://doi.org/10.1109/TITS.2021.3055258.
Ma, C., Zhao, Y., Dai, G., Xu, X., Wong, S.-C., 2023. A novel STFSA-CNN-GRU hybrid model for short-term traffic speed prediction. IEEE Trans. Intell. Transport. Syst. 24, 3728–3737. https://doi.org/10.1109/TITS.2021.3117835.
Ma, L., Solomonik, E., 2022. Accelerating alternating least squares for tensor decomposition by pairwise perturbation. Numer. Lin. Algebra Appl. 29, e2431 https://doi.org/10.1002/nla.2431.
Makowski, D., Naud, C., Jeuffroy, M.-H., Barbottin, A., Monod, H., 2006. Global sensitivity analysis for calculating the contribution of genetic parameters to the variance of crop model prediction. Reliab. Eng. Syst. Saf. 91, 1142–1147. https://doi.org/10.1016/j.ress.2005.11.015.
Méndez, M., Merayo, M.G., Núñez, M., 2023. Long-term traffic flow forecasting using a hybrid CNN-BiLSTM model. Eng. Appl. Artif. Intell. 121, 106041 https://doi.org/10.1016/j.engappai.2023.106041.
Mokhtarimousavi, S., Anderson, J.C., Azizinamini, A., Hadi, M., 2019. Improved support vector machine models for work zone crash injury severity prediction and analysis. Transport. Res. Rec. 2673, 680–692. https://doi.org/10.1177/0361198119845899.
Muruganantham, A., Tan, K.C., Vadakkepat, P., 2016. Evolutionary dynamic multiobjective optimization via kalman filter prediction. IEEE Trans. Cybern. 46, 2862–2873. https://doi.org/10.1109/TCYB.2015.2490738.
Narmadha, S., Vijayakumar, V., 2021. Spatio-Temporal vehicle traffic flow prediction using multivariate CNN and LSTM model. Mater. Today 81, 826–833. https://doi.org/10.1016/j.matpr.2021.04.249.
Nguyen, D., Vadaine, R., Hajduch, G., Garello, R., Fablet, R., 2022. GeoTrackNet–A maritime anomaly detector using probabilistic neural network representation of AIS tracks and A contrario detection. IEEE Trans. Intell. Transport. Syst. 23, 5655–5667. https://doi.org/10.1109/TITS.2021.3055614.
Okutani, I., Stephanedes, Y.J., 1984. Dynamic prediction of traffic volume through Kalman filtering theory. Transp. Res. Part B Methodol. 18, 1–11. https://doi.org/10.1016/0191-2615(84)90002-X.
Pang, X., Li, Z., Tseng, M.-L., Liu, K., Tan, K., Li, H., 2020. Electric vehicle relay lifetime prediction model using the improving fireworks algorithm-grey neural network model. Appl. Sci.-Basel 10, 1940. https://doi.org/10.3390/app10061940.
Park, Y.-S., Jong, J.-Y., Inoue, K., 2002. A study on assessment of vessel traffic safety management by marine traffic flow simulation. Journal of the Korea Society for Simulation 11, 43–55.
Robards, M.D., Silber, G.K., Adams, J.D., Arroyo, J., Lorenzini, D., Schwehr, K., Amos, J., 2016. Conservation science and policy applications of the marine vessel Automatic Identification System (AIS)-a review. Bull. Mar. Sci. 92, 75–103. https://doi.org/10.5343/bms.2015.1034.
Rong, H., Teixeira, A.P., Guedes Soares, C., 2022. Maritime traffic probabilistic prediction based on ship motion pattern extraction. Reliab. Eng. Syst. Saf. 217, 108061 https://doi.org/10.1016/j.ress.2021.108061.
Rong, H., Teixeira, A.P., Guedes Soares, C., 2020. Data mining approach to shipping route characterization and anomaly detection based on AIS data. Ocean Eng. 198, 106936 https://doi.org/10.1016/j.oceaneng.2020.106936.
Sadeghi-Niaraki, A., Mirshafiei, P., Shakeri, M., Choi, S.-M., 2020. Short-term traffic flow prediction using the modified elman recurrent neural network optimized through a genetic algorithm. IEEE Access 8, 217526–217540. https://doi.org/10.1109/ACCESS.2020.3039410.
Salakhutdinov, R., Mnih, A.. Bayesian probabilistic matrix factorization using Markov chain Monte Carlo. https://doi.org/10.1145/1390156.1390267.
Sousa, S.I.V., Martins, F.G., Alvim-Ferraz, M.C.M., Pereira, M.C., 2007. Multiple linear regression and artificial neural networks based on principal components to predict ozone concentrations. Environ. Model. Software 22, 97–103. https://doi.org/10.1016/j.envsoft.2005.12.002.
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 https://doi.org/10.1016/j.oceaneng.2022.111954.
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, 1249–1258. https://doi.org/10.1049/iet-its.2020.0004.
Suo, Y., Chen, W., Claramunt, C., Yang, S., 2020. A ship trajectory prediction framework based on a recurrent neural network. Sensors 20, 5133. https://doi.org/10.3390/s20185133.
Sutskever, I., Vinyals, O., Le, Q.V., 2014. Sequence to sequence learning with neural networks. In: Advances in Neural Information Processing Systems. 27.
Tan, H., Wu, Y., Shen, B., Jin, P.J., Ran, B., 2016. Short-term traffic prediction based on dynamic tensor completion. IEEE Trans. Intell. Transport. Syst. 17, 2123–2133. https://doi.org/10.1109/TITS.2015.2513411.
Tang, J., Zhang, X., Yin, W., Zou, Y., Wang, Y., 2021. Missing data imputation for traffic flow based on combination of fuzzy neural network and rough set theory. Journal of Intelligent Transportation Systems 25, 439–454. https://doi.org/10.1080/15472450.2020.1713772.
Tang, Y., Cheng, N., Wu, W., Wang, M., Dai, Y., Shen, X., 2019. Delay-minimization routing for heterogeneous VANETs with machine learning based mobility prediction. IEEE Trans. Veh. Technol. 68, 3967–3979. https://doi.org/10.1109/TVT.2019.2899627.
Van Der Voort, M., Dougherty, M., Watson, S., 1996. Combining Kohonen maps with ARIMA time series models to forecast traffic flow. Transport. Res. C Emerg. Technol. 4, 307–318. https://doi.org/10.1016/S0968-090X(97)82903-8.
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, 1231. https://doi.org/10.3390/jmse9111231.
Wang, J., Chen, R., He, Z., 2019. Traffic speed prediction for urban transportation network: a path based deep learning approach. Transport. Res. C Emerg. Technol. 100, 372–385. https://doi.org/10.1016/j.trc.2019.02.002.
Wang, J., Liu, S., Wang, S., Liu, Q., Liu, H., Zhou, H., Tang, J., 2021. Multiple indicators-based health diagnostics and prognostics for energy storage technologies using fuzzy

comprehensive evaluation and improved multivariate grey model. IEEE Trans. Power Electron. 36, 12309–12320. https://doi.org/10.1109/TPEL.2021.3075517.
Wang, T., Fu, L., Zhou, Y., Gao, S., 2022. Service price forecasting of urban charging infrastructure by using deep stacked CNN-BiGRU network. Eng. Appl. Artif. Intell. 116, 105445 https://doi.org/10.1016/j.engappai.2022.105445.
Wang, Y., Dang, Y., Li, Y., Liu, S., 2010. An approach to increase prediction precision of GM(1,1) model based on optimization of the initial condition. Expert Syst. Appl. 37, 5640–5644. https://doi.org/10.1016/j.eswa.2010.02.048.
Wei, G., 2019. Traffic prediction and attack detection approach based on PSO optimized elman neural network. In: 2019 11th International Conference on Measuring Technology and Mechatronics Automation (Icmtma 2019). Ieee, New York, pp. 504–508. https://doi.org/10.1109/ICMTMA.2019.00117.
Weng, Y., Wang, X., Hua, J., Wang, H., Kang, M., Wang, F.-Y., 2019. Forecasting horticultural products price using ARIMA model and neural network based on a large-scale data set collected by web crawler. IEEE Trans. Comput. Soc. Syst. 6, 547–553. https://doi.org/10.1109/TCSS.2019.2914499.
Williams, B.M., Hoel, L.A., 2003. Modeling and forecasting vehicular traffic flow as a seasonal ARIMA process: theoretical basis and empirical results. J. Transport. Eng. 129, 664–672. https://doi.org/10.1061/(ASCE)0733-947X(2003)129:6(664).
Xiao, Z., Fu, X., Zhang, L., Zhang, W., Liu, R.W., Liu, Z., Goh, R.S.M., 2020. Big data driven vessel trajectory and navigating state prediction with adaptive learning, motion modeling and particle filtering techniques. IEEE Trans. Intell. Transport. Syst. 23 (4), 3696–3709. https://doi.org/10.1109/TITS.2020.3040268.
Xiao, Z., Fu, X., Zhao, L., Zhang, L., Teo, T.K., Li, N., Zhang, W., Qin, Z., 2022. Next-generation vessel traffic services systems—From "passive" to "proactive. IEEE Intell. Transport. Syst. Mag. 15 (1), 363–377. https://doi.org/10.1109/MITS.2022.3144411.
Xiao, H., Zhao, Y., Zhang, H., 2023. Predict vessel traffic with weather conditions based on multimodal deep learning. J. Mar. Sci. Eng. 11, 39. https://doi.org/10.3390/jmse11010039.
Xin, X., Liu, K., Loughney, S., Wang, J., Li, H., Ekere, N., Yang, Z., 2023. Multi-scale collision risk estimation for maritime traffic in complex port waters. Reliab. Eng. Syst. Saf. 240, 109554. https://doi.org/10.1016/j.ress.2023.109554.
Xin, X., Liu, K., Loughney, S., Wang, J., Li, H., Yang, Z., 2023. Graph-based ship traffic partitioning for intelligent maritime surveillance in complex port waters. Expert Syst. Appl. 231, 120825. https://doi.org/10.1016/j.eswa.2023.120825.
Xu, D., Wang, Y., Jia, L., Qin, Y., Dong, H., 2017. Real-time road traffic state prediction based on ARIMA and Kalman filter. Front. Inform. Technol. Elect. Eng. 18, 287–302. https://doi.org/10.1631/FITEE.1500381.
Xu, T., Zhang, Q., 2022. Ship traffic flow prediction in wind farms water area based on spatiotemporal dependence. J. Mar. Sci. Eng. 10, 295. https://doi.org/10.3390/jmse10020295.
Yang, B., Sun, S., Li, J., Lin, X., Tian, Y., 2019. Traffic flow prediction using LSTM with feature enhancement. Neurocomputing 332, 320–327. https://doi.org/10.1016/j.neucom.2018.12.016.
Yang, F., Liu, G., Huang, L., Chin, C.S., 2020. Tensor decomposition for spatial-temporal traffic flow prediction with sparse data. Sensors 20, 6046. https://doi.org/10.3390/s20216046.
Yang, J., Liu, Y., Ma, L., Ji, C., 2022. Maritime traffic flow clustering analysis by density based trajectory clustering with noise. Ocean Eng. 249, 111001 https://doi.org/10.1016/j.oceaneng.2022.111001.
Yi, Q., Zuo, Y., Li, T., Mao, Y., Xiao, Y., 2021. Forecasting of vessel traffic flow using BPNN based on genetic algorithm optimization. In: Iwcmc 2021: 2021 17th International Wireless Communications & Mobile Computing Conference (Iwcmc). Ieee, New York, pp. 1059–1063. https://doi.org/10.1109/IWCMC51323.2021.9498607.
Yin, X., Wu, G., Wei, J., Shen, Y., Qi, H., Yin, B., 2022. Deep learning on traffic prediction: methods, analysis and future directions. IEEE Trans. Intell. Transport. Syst. 23, 4927–4943. https://doi.org/10.1109/TITS.2021.3054840.
Yu, X., Long, W., Li, Y., Gao, L., Shi, X., 2022. Trajectory dimensionality reduction and hyperparameter settings of DBSCAN for trajectory clustering. IET Intell. Transp. Syst. 16, 691–710. https://doi.org/10.1049/itr2.12166.
Yu, X., Sun, L., Yan, Y., Liu, G., 2021. A short-term traffic flow prediction method based on spatial–temporal correlation using edge computing. Comput. Electr. Eng. 93, 107219 https://doi.org/10.1016/j.compeleceng.2021.107219.
Zhang, K., Yu, X., Liu, S., Dong, X., Li, D., Zang, H., Xu, R., 2022. Wind power interval prediction based on hybrid semi-cloud model and nonparametric kernel density estimation. Energy Rep. 8, 1068–1078. https://doi.org/10.1016/j.egyr.2022.02.094.
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 https://doi.org/10.1016/j.ress.2021.108317.
Zhang, M., Kujala, P., Musharraf, M., Zhang, J., Hirdaris, S., 2023. A machine learning method for the prediction of ship motion trajectories in real operational conditions. Ocean Eng. 283, 114905. https://doi.org/10.1016/j.oceaneng.2023.114905.
Zhang, W., Wu, Z., 2022. Optimal hybrid framework for carbon price forecasting using time series analysis and least squares support vector machine. J. Forecast. 41, 615–632. https://doi.org/10.1002/for.2831.
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 15, 1688–1711. https://doi.org/10.1080/23249935.2019.1637966.
Zhang, Z., Yin, J., Wang, N., Hui, Z., 2019. Vessel traffic flow analysis and prediction by an improved PSO-BP mechanism based on AIS data. Evolving Systems 10, 397–407. https://doi.org/10.1007/s12530-018-9243-y.
Zhao, C., Li, X., Zuo, M., Mo, L., Yang, C., 2022. Spatiotemporal dynamic network for regional maritime vessel flow prediction amid COVID-19. Transport Pol. 129, 78–89. https://doi.org/10.1016/j.tranpol.2022.09.029.
Zhao, Q., Li, L., Zhang, L., Zhao, M., 2023. Recognition of corrosion state of water pipe inner wall based on SMA-SVM under RF feature selection. Coatings 13, 26. https://doi.org/10.3390/coatings13010026.
Zhao, Q., Zhang, L., Cichocki, A., 2015. Bayesian CP factorization of incomplete tensors with automatic rank determination. IEEE Trans. Pattern Anal. Mach. Intell. 37, 1751–1763. https://doi.org/10.1109/TPAMI.2015.2392756.
Zheng, H., Zhang, L., Zhao, X., 2022. How does environmental regulation moderate the relationship between foreign direct investment and marine green economy efficiency: an empirical evidence from China's coastal areas. Ocean Coast Manag. 219, 106077 https://doi.org/10.1016/j.ocecoaman.2022.106077.
Zhou, J., Huang, S., Wang, M., Qiu, Y., 2022. Performance evaluation of hybrid GA–SVM and GWO–SVM models to predict earthquake-induced liquefaction potential of soil: a multi-dataset investigation. Eng. Comput. 38, 4197–4215. https://doi.org/10.1007/s00366-021-01418-3.
Zou, Y., Zhu, T., Xie, Y., Zhang, Yunlong, Zhang, Yue, 2022. Multivariate analysis of car-following behavior data using a coupled hidden Markov model. Transport. Res. C Emerg. Technol. 144, 103914 https://doi.org/10.1016/j.trc.2022.103914.