Mixed-frequency data-driven forecasting port throughput: a novel attention-DeepAR-MIDAS model

Bo Shao

Department of Industrial and Systems Engineering,
University of Wisconsin-Madison,
Madison, USA
Email: bo.shao@wisc.edu

Xiaoli Su

School of Economics and Management,
Fuzhou University,
Fuzhou, China
Email: 210710012@fzu.edu.cn

Xin Li

School of Transportation,
Fuzhou University of Technology,
Fuzhou, China
Email: lx1592421854@163.com

Xingxuan Zhuo*

School of Economics and Management,
Fuzhou University,
Fuzhou, China
Email: zhuoxx1991@fzu.edu.cn
*Corresponding author

Abstract: Forecasting cargo throughput is an essential albeit challenging task for national and port optimisation decision-making, resource allocation, and control planning. To this end, a novel forecasting model is developed for mixed-frequency data called attention-DeepAR-MIDAS (ADM) by introducing the mixed data sampling (MIDAS) technique and attention mechanism into the DeepAR forecasting algorithm in this study. The proposed ADM model is specifically designed with an attention mechanism to accurately identify and prioritise the most influential variables, both endogenous and exogenous, over time. Hence, it can effectively use the nonlinear information of mixed-frequency data, which is conducive to port throughput forecasting. Furthermore, the ADM model possesses both long-term and short-term high-precision forecasting capabilities, enabling multi-step probability forecasting and better tracking of abnormal changes in endogenous and exogenous variables of port throughput, fitting their fluctuation trends. By analysing the differences in model performance before and after improvement based on forecast accuracy metrics, probability interval testing, and DM testing methods, the ADM model achieves accurate forecasting. Finally, China's monthly port throughput forecast results also illustrate the superiority of the ADM model, which provides decision-makers with more timely, accurate, and comprehensive forecasts.

Keywords: machine learning; mixed-frequency data; nonlinearity; time-series; port throughput forecasting.

Reference to this paper should be made as follows: Shao, B., Su, X., Li, X. and Zhuo, X. (xxxx) 'Mixed-frequency data-driven forecasting port throughput: a novel attention-DeepAR-MIDAS model', *Int. J. Shipping and Transport Logistics*, Vol. X, No. Y, pp.xxx–xxx.

Biographical notes: Bo Shao is currently pursuing a Master's degree in the Department of Industrial and Systems Engineering at the University of Wisconsin-Madison. His research is primarily concentrated in the fields of industrial and systems engineering, focusing on operations research, operations management, and machine learning.

Xiaoli Su is a PhD student at the School of Economics and Management, Fuzhou University, China. Her research primarily delves into the management systems engineering, focusing on operations management, supply chain management, machine learning, and data analysis.

Xin Li holds an MSc in Traffic and Transportation Engineering from the Fuzhou University of Technology, China. She works as a research assistant at Fuzhou University currently. Her research interests are focused on machine learning, data mining, and transportation optimisation.

Xingxuan Zhuo is an Assistant Professor of Economics and Management at Fuzhou University of China. She holds a PhD in Management from the Hefei University of Technology of China in 2019. Her research interests are in the fields of time series analysis, forecasting models, and deep learning. Some of her recent research focused on mixed frequency data analysis and high dimensional variable selection.

1 Introduction

Maritime transport serves as a critical link in global supply chains. Over 80% of global trade by volume is carried by sea (UNCTAD, 2022). Port throughput is a key indicator in evaluating maritime transport performance of ports and countries. It not only directly reflects the production capacity of the port, but also indirectly represents a new barometer of global supply. In addition, the fluctuation of port throughput can effectively reflect the fluctuation of the global maritime trade, the balance between supply and demand, and the competitive situation (Du et al., 2019). It can provide a basis for national and port optimisation decision-making, resource allocation, and control planning. Therefore, the alignment of navigational capacity and infrastructure investment of the port with the volatile market demand can be enhanced by conducting reasonable and accurate port throughput forecasting. Additionally, it is possible to effectively avoid issues (e.g., capacity vacancy, resource wastage, and investment mistakes) by exploring the fluctuation trend of port throughput and allocating port resources reasonably based on the forecasting results. This has significant importance in improving the layout of port infrastructure and enhancing operational efficiency.

Forecasting port throughput is an essential albeit challenging task in ensuring efficient maritime transport. In the time dimension, the throughput of a future period is closely related to the throughput of history nodes with similar characteristics. Moreover, the time series data of port throughput show volatile, non-stationary, unstructured characteristics and nonlinear trends, exhibiting highly complex behaviour. It is a complex time series with long-term memory and nonlinear dependence in the volatility process, making it vulnerable to exogenous and endogenous factors. Currently, global trade is experiencing high uncertainties, with external factors (e.g., high inflation, protectionism, and restructuring of global industrial chains) exacerbating the fluctuation and complexity of port throughput. However, existing mathematical and statistical methods cannot capture the uncertainty and nonlinearity of port throughput well, leading to poor forecasting performance. Hence, it is very essential to develop an effective forecasting approach to capture the uncertainty and nonlinearity of port throughput against exogenous and endogenous disruptions.

The application of big data and machine learning enables the sampling or analysis of multi-source time series data. Data-driven machine learning algorithms have been widely used in time series forecasting issues. Existing studies have shown that forecasting models based on machine learning have superiority when applied to throughput time series data. It can not only effectively fit complex nonlinear relationships, but also avoid the overfitting problem of shallow structures (Justice et al., 2016; Shankar et al., 2020). However, port throughput forecasting is not only influenced by the autocorrelation of historical data but also affected by various uncertainties, resulting in strong randomness and irregularity in trend. Moreover, the relationships among collected port throughput data are often complex due to the varying sampling frequencies. Hence, the exploration of complex mixed-frequency time series, the nonlinearity of the sequences, and the timely detection of abnormal fluctuations pose a significant challenge.

To improve the forecasting accuracy of time series models in handling mixed-frequency data, this study introduces an attention-DeepAR-MIDAS (ADM) model. The ADM model introduces the DeepAR algorithm with the mixed data sampling (MIDAS) technique and an attention mechanism, preserving the original data structure to safeguard the authenticity of the information. It can identify hidden nonlinear relationships within mixed-frequency data, discern the significance of time sequences, quickly react to external disruptions, and perform multi-step probabilistic forecasting. This model significantly forecasts uncertainty, accurately tracking port throughput fluctuations, thus enabling ports to swiftly adapt to market trends, revise operational strategies, and diminish market risks. Furthermore, the ADM model facilitates the precise assessment of economic and trade situations by the country, as well as the formulation and implementation of policies and plans tailored for maritime transport.

The remainder of our work is organised as follows. In Section 2, we provide an overview of relevant study work. In Section 3, we develop a novel forecasting model based on mixed-frequency time series data analytics, and Section 4 presents the practice-based validation results, as well as the statistical analysis. Finally, Section 5 draws the main conclusions of the study, highlighting the most significant findings.

2 Literature review

Growth and global trade could slow in a deglobalising world economy. International market demand remains weak, international trade and investment growth are sluggish. In particular, the rise of trade protectionism and unilateralism has severely impacted global trade liberalisation and added uncertainties to the global economy. Consequently, global maritime transport and port activity were severely affected as well. Various uncertainties have intensified the fluctuations in port throughput, making it difficult to forecast port throughput. Therefore, how to make accurate and effective port throughput forecasting has become a focus of attention and exploration in the academic community.

Existing methods for port throughput forecasting can be primarily categorised into mathematical and statistical methods and machine-learning-based methods. The study of forecasting models based on mathematical and statistics has a long history and has led to various forecasting models, including error correction model (ECM) (Fung, 2002), exponential smoothing (ES) (Zhang et al., 2019), autoregressive integrated moving average (ARIMA) (Rashed et al., 2017), seasonal autoregressive integrated moving average (SARIMA) (Min and Ha, 2014), and vector autoregressive (VAR) model (Tian et al., 2010). Schulze and Prinz (2009) used SARIMA and ES models to predict Germany's quarterly container throughput, and the results showed that the SARIMA forecasting model outperformed the ES forecasting model for quarterly container throughput. Xiao et al. (2014) employed causal analysis to construct a model for predicting port cargo throughput, and experimental results demonstrated that the combination model based on causal analysis had higher accuracy than time series forecasting model. Patil and Sahu (2016) utilised regression models to predict cargo volume at the Mumbai port and conducted a simple elasticity analysis of the factors influencing cargo volume. This forecasting model not only considered the time series of container throughput but also investigated the impact of relevant factors on container throughput, thus effectively improving the forecast accuracy (Gosasang et al., 2011). However, these mathematical and statistical models are based on linear assumptions and cannot capture nonlinear patterns hidden in the original data, leading to poor predictive performance, especially for certain nonlinear time series data (García et al., 2014). Moreover, researchers have gradually realised their limitations, such as handling multi-source heterogeneous complex data, fitting relationships between exogenous variables, capturing port throughput fluctuations accurately, and dealing with uncertainty.

In recent years, with the integration of statistics and computer science, the application of machine learning has further enhanced the accuracy of port throughput forecasting. To overcome the above limitations, some scholars have introduced neural networks with strong adaptability and the ability to fit nonlinear relationships into port throughput forecasting (Gosasang et al., 2011; Xiao et al., 2023). García et al. (2014) also employed artificial neural networks (ANN) to study port planning parameters, and the results showed that ANN can accurately predict port traffic flow and equipment demand. Ruiz-Aguilar et al. (2020a) considered machine learning methods as powerful tools for detecting peaks and congestion in port cargo facility workloads. Through a study on the cargo volume of the Algeciras Bay port, the results showed that neural networks performed better than any other independent autoregressive models and exhibited more stability in short-term forecasting. Furthermore, Moscoso-López et al. (2021) investigated the application of machine learning methods in predicting waterborne cargo volume, and their results also indicated lower prediction errors of machine learning compared to traditional regression models. Xu et al. (2022b) used nine methods to forecast the historical throughput of the world's top 20 container ports and compared the results between methods. The study found that the average accuracy of machine learning forecasting methods was higher than that of traditional methods and more suitable for predicting port-related cargo volumes. To further enhance the accuracy and reliability of our port throughput forecasting, we have integrated the DeepAR model into our study. This model, known for its probabilistic output, represents a significant advancement over traditional models that primarily focus on point predictions. DeepAR's approach of providing distributions of future values is vital in addressing the inherent uncertainties in port throughput forecasting. This feature allows for a more nuanced understanding of potential future scenarios, thereby mitigating risks associated with reliance on single-point forecasts. The efficacy of the DeepAR model in handling complex forecasting tasks is well-documented in various fields. For instance, its application in deformation prediction of unstable slopes as explored by Dong et al. (2020) and in the domain of stock price forecasting as demonstrated by Xie et al. (2023), underlines its versatility and robustness. These instances of application across different sectors show how DeepAR effectively manages data patterns that are complex and influenced by multiple external factors, reflecting the similar challenges we face in forecasting port throughput. In summary, machine learning-based methods can more accurately capture data variations, and demonstrate superior accuracy, stability, and significance in throughput forecasting.

However, despite the ability of machine learning-based methods to accurately forecast both linear and nonlinear stationary data, they also have their weaknesses, such as sensitivity to parameters, data consistency, and computational complexity. Most of the commonly used literature on the port throughput forecasting mentioned above conducts research by constructing models using same-frequency data, ignoring the diversity information and volatility of high frequency endogenous and exogenous factors that affect port throughput. This leads to the inability to fully utilise the valuable information brought by mixed-frequency variables (however, the diversity information and volatility of high frequency data in the data have been ignored, which will cause information loss). Tang et al. (2014) and Yu et al. (2015) mentioned that port throughput forecasting in the maritime field is a highly complex and dynamic process, requiring further analysis of the data characteristics' complexity, including various nonlinear features such as chaos, fractality, irregularity, and long-term memory.

Consequently, how to conduct models and analyses based on mixed-frequency data has become a research hotspot in the academic community. Regarding the processing of mixed-frequency data, there are two approaches. The first method is to transform high frequency data to lower frequencies by calculating averages or using algorithms such as replacing discrete points. The second method is to transform low frequency data to higher frequencies using algorithms such as fitting, interpolation, or bridging model algorithms. However, both approaches have problems, such as the loss of valuable information in high frequency data or the introduction of ineffective information due to human manipulation, leading to increased errors. These issues may have implications for model estimation and strategy selection. To address this, Ghysels et al. (2006) proposed the MIDAS regression models, which does not alter the data structure. It utilises weighting schemes based on distribution lag models and constructs bridging equations, thereby enhancing the utilisation of high frequency information, avoiding parameter diffusion, and ensuring timely model forecasting.

Recent advancements in time-series forecasting, exemplified by studies such as those by Li et al. (2023), have made significant strides in incorporating traditional attention mechanisms into the DeepAR model, a pivotal development for complex tasks like port throughput forecasting. Despite these advancements, a notable shortcoming in current research is the inadequate consideration of mixed frequency data, an essential element for accurately capturing the nuances of complex environments. This gap is particularly evident in the context of port throughput forecasting, which has recently witnessed a shift towards integrating machine learning methods with traditional forecasting techniques. A key instance of this trend is the GARCH-MIDAS model, developed by Xu et al. (2022a), which applies threshold generalised autoregressive conditional heteroscedasticity and exponential generalised autoregressive heteroscedasticity to analyse the asymmetric impact of carbon emissions control variables on international shipping. Subsequently, Xu et al. (2022a) extended this approach by examining the influence of new infection rates on market volatility, factoring in diverse elements like freight rates, Brent crude oil prices, container idle rates, and port congestion levels, thereby highlighting the variables' substantial impact on the Baltic Dry Index (BDI). This exploration laid a foundational theoretical framework for incorporating multi-source mixed-frequency data into machine learning models tailored for port throughput forecasting. Parallel to these developments, methodologies such as QRNN (Xu et al., 2021) and SVR (Xu et al., 2020) have made significant contributions to frequency alignment in machine learning predictions. However, their focus predominantly on point prediction reveals a limitation in addressing the broader spectrum of uncertainty in port throughput forecasting.

Addressing these challenges, we present a novel forecasting model termed the ADM, by embedding the MIDAS technique and an attention mechanism into the DeepAR framework to refine port throughput forecasting. The ADM model uniquely processes mixed-frequency data without the need for preliminary preprocessing, inputting it directly into the DeepAR network layer. This innovation exploits the DeepAR's data-driven and adaptive learning strengths to autonomously identify and explore potential nonlinear patterns within the endogenous and exogenous variables of port throughput. The model can output the probability distribution of port throughput directly, thereby quantifying uncertainties and related forecasting risks. With its integrated attention mechanism, the ADM model is adept at extracting significant temporal patterns, learning not only the interplay among endogenous and exogenous variables within the same timeframe but also their dependencies over all antecedent time steps. This ability ensures that the model selects crucial variables and time points, thereby capturing short-term fluctuations and improving long-term forecast accuracy.

3 Methodology

We present a novel ADM model by introducing the attention mechanism and the MIDAS technique into the framework of DeepAR in this section. The MIDAS method is appropriate in situations where the explanatory variable is high frequency data while the response variable is sampled at a lower frequency. It handles series sampled at different frequencies, such as forecasting low frequency information like China's port cargo throughput (monthly) from high frequency information like the Shanghai Stock Exchange Composite Index (daily) and the China Coastal Bulk Freight Index (weekly). The combination of attention mechanism, DeepAR, and MIDAS provides a new approach to forecasting port throughput. To address dynamic correlations and time sensitivity among the factors influencing port throughput and port throughput itself, filters are introduced in the attention mechanism.

3.1 ADM model

The goal of the ADM model is to establish a conditional probability distribution of mixed-frequency data, which can be viewed as the conjunctive of the likelihood function $l(\cdot)$at a series of times to reduce the uncertainty of port throughput. Among them, the frequency mismatch between the low frequency response variable $y_t(\theta)(t = 1,2,\ldots,T)$ and the high frequency explanatory variable $x_{\tau_i}(i = 1,2,\ldots,I;\ \tau_i = 1,2,\ldots,m_iT)$ that affects port throughput is represented by $\{m_i\}_{i=1}^{I}$. To reflect the multi-step ahead forecasting, the low frequency response variable $y_t(\theta)$ of port throughput is divided by t_0 time point, that is, the prediction range $h^{(t_0)} = t_0{:}T$ related to the low frequency variable of port throughput. The low frequency forecasting range $h^{(t_0)}$ and the high frequency forecasting range s_i included in the multi-step ahead forecast have a specific connection, which is $h^{(t_0)} = \lceil s_1 \rceil = \lceil s_2 \rceil = \cdots = \lceil s_I \rceil$, where $\lceil \cdot \rceil$ is the ceiling function.

Given the data set $\Omega_t = \{y_t(\theta), x_{\tau_1}, \ldots, x_{\tau_I}\}$, the high frequency variables that impact port throughput are first transformed into low frequency data with the same frequency as port throughput using the frequency alignment approach and parameter function restrictions. This avoids parameter dispersion and guarantees data frequency consistency. To effectively perceive the information of endogenous and exogenous data at the port, the relevance and temporal sensitivity of high frequency factors affecting port throughput are then identified. The output of the multi-step ahead projections follows the application of the weighting.

$$P\left(y_t(\theta)\middle|\Omega_t\right) = \prod_{t=t_0}^{T} l\left(y_t(\theta)\middle|\theta\left(h_i', \Theta\right)\right) \tag{1}$$

This study assumes that the low frequency variable $y_t(\theta)$ of each port throughput obeys a Gaussian likelihood function satisfying a certain parameter in equation (3), and finds the best parameter Θ by optimising the loss function of equation (2). Our model improved the network structure of the DeepAR model, enhanced the degree of temporal attention to the hidden state h_i of the original DeepAR model network layer through the attention mechanism, and then changed it into the updated hidden state h_i' (more detailed information on attention mechanism is found in Section 3.2.4). Different from the study on demand distribution assumed by Silver et al. (2016), the mean value and standard deviation $\Theta = \{\mu(\cdot), \sigma(\cdot)\}$were obtained by optimising ADM model fitting in equation (4) and equation (5), and *Softplus* activated function $\sigma(\cdot)$ to ensure the continuity and non-negative standard deviation.

$$Loss = \sum_{i=1}^{I}\sum_{t=t_0}^{T} -\log l\left(y_t(\theta)\middle|\theta\left(h_i', \Theta\right)\right) \tag{2}$$

$$l\left(\mu(\cdot), \sigma(\cdot)\right) = \left(2\pi\sigma^2(\cdot)\right)^{-\frac{1}{2}} \exp\left(-\left(y_t(\theta) - \mu(\cdot)\right)^2 \Big/ \left(2\sigma^2(\cdot)\right)\right) \tag{3}$$

$$\mu\left(h_i'\right) = \mathrm{W}_\mu^{\mathrm{T}} h_i' + b_\mu \tag{4}$$

$$\sigma\left(h_i'\right) = \log\left(1 + \exp\left(\mathrm{W}_\sigma^{\mathrm{T}} h_i' + b_\sigma\right)\right) \tag{5}$$

3.2 ADM model structure

3.2.1 Input layer

Frequency mismatch occurs because the endogenous and exogenous data sources for port throughput are many, varied, and gathered at various frequencies. However, neural network models cannot directly consider both the low frequency factors of port throughput in the input layer and the high frequency variables impacting port throughput with frequency mismatch. Therefore, it is required to enhance the input layer by directly feeding mixed-frequency data into the neural network input layer using the MIDAS technique. With this, preprocessing of mixed-frequency data before input into the model is intended to be avoided, and the useful information of high frequency variables is completely used.

3.2.2 MIDAS layer

The frequency alignment of each port throughput endogenous and exogenous variable $x_{\tau_i - s_i}$ is achieved based on the corresponding maximum lag order L_i, i.e., each high frequency explanatory variable $x_{\tau_i - s_i}$ is transformed into a low frequency vector $(x_{i,t-s_i}, x_{i,t-1/m_i-s_i}, \ldots, x_{i,t-L_i/m_i-s_i})^{\mathrm{T}}$. The frequency alignment vector is restricted to reduce the number of estimated parameters, and a low frequency variable $x_{i,t-s_i}$ with the same frequency as the output $y_t(\theta)$ is generated.

$$x_{i,t-s_i} = \sum\nolimits_{l=0}^{L_i} w_i(\delta; l) x_{i,t-l/m_i-s_i} \tag{6}$$

where $w_i(\delta; l)$ is the weighting scheme; $t = q, q,+1, \ldots, T$ and q is the smallest integer that satisfies $q - L_i / m_i - s_i \geq 0$.

Following Ghysels et al. (2007) and Gagliardini et al. (2017), we specify $w_i(\delta; l)$ as the two parameters of exponential Almon lag polynomial, namely

$$w_i\left(\delta; l\right) = \frac{\exp\left(\delta_1 l + \delta_2 l^2\right)}{\sum\nolimits_{l=0}^{L_i} \exp\left(\delta_1 l + \delta_2 l^2\right)}$$

with $\delta = (\delta_1, \delta_2)^{\mathrm{T}}$.

3.2.3 DeepAR network layer

The data from the previous time step (the input $x_{i,t-s_i}$, the previously hidden state $h_{i,t-s_i-1}$, the weight matrix $W_r(\theta)$, and the bias term $b_r(\theta)$ of the input gate) is converted into the control signal of the input gate based on the sigmoid function $\sigma(x) = 1 / (1 + e^{-x})$ when calculating the value of the input gate $Z_i^r(\theta)$. It is then entered into the DeepAR network layer together with the aligned value. DeepAR network layer learning is accomplished using the network structure (long short-term memory neural network) of the DeepAR model, rather than recurrent neural network (RNN) models, which suffer from gradient vanishing and explosion issues. The current time step's cell state $C_i'(\theta)$ is first calculated, followed by the control signal for the forget gate $f_i(\theta)$ and the i^{th} DeepAR network layer node $C_i(\theta)$. The DeepAR network layer's output $Z_i^o(\theta)$ is then acquired.

$$f_i(\theta) = \sigma\left(W_f(\theta)\left[x_{i,t-s_i}, h_{i,t-s_i-1}\right] + b_f(\theta)\right) \tag{7}$$

$$C_i'(\theta) = \tanh\left(W_c(\theta)\left[x_{i,t-s_i}, h_{i,t-s_i-1}\right] + b_c(\theta)\right) \tag{8}$$

$$C_i(\theta) = \sum_{i=1}^{I}\left(f_i(\theta) \circ C_{i-1}(\theta) + Z_i^r(\theta) \circ C_i'(\theta)\right) \tag{9}$$

$$Z_i^o(\theta) = \sigma\left(W_o(\theta)\left[x_{i,t-s_i}, h_{i,t-s_i-1}\right] + b_o(\theta)\right) \tag{10}$$

where $i = 1,2,\ldots,I$; the symbol '$\circ$' represents the Hadamard product; $\sigma(\cdot)$ and $\tanh(\cdot)$are sigmoid functions and hyperbolic tangent functions. In addition to the weight matrix and bias term of the forget gate, the current input cell state is represented by W_c and b_c, the current output is represented by W_o and b_o, and the weight matrix and bias term of the forget gate is represented by W_f and b_f.

The long-term memory $C_{i-1}(\theta)$ and the present memory $C_i'(\theta)$ are merged to create the new cell state $C_i(\theta)$ based on the previously mentioned variables. It can retain the knowledge of historical port throughput statistics by including the control function of the forget gate.

3.2.4 Attention mechanism layer

Based on the DeepAR network layer, it is necessary to dig deeply into the crucial time information of port throughput with the aid of attention mechanism to extract effective time information and learn the degree of influence of various variables affecting port throughput at different moments on port throughput.

The traditional attention mechanism selects only the information related to the current time step of the port throughput. This design is suitable for the task that each time step contains one port throughput information, but if the generated data inside and outside the port is introduced, that is, if there are multiple variables in each time step, the variables with noise in the prediction utility cannot be ignored. In addition, because the traditional attention mechanism averages information over multiple time steps, it is impossible to detect time patterns that are useful for prediction. To explore the temporal correlation pattern among the mixed sampling frequency variables of port throughput, Shih et al. (2019) proposed an attention mechanism containing information spanning multiple time steps. The proposed method involves several steps. Firstly, the result $h_i(\theta)$ of the attention mechanism layer is calculated based on the output of the DeepAR network layer. Secondly, feature extraction is performed using convolutional neural network (CNN) filters. Subsequently, the score function $f(\cdot)$ is used to evaluate the impact strength of the multiple variables that affect each port throughput, both endogenous and exogenous, on the forecasting results at different time intervals. This leads to determining the weights of each time sequence pattern that can contribute to more effective forecasting. Finally, the weight information is integrated, and updates are made using calculated weights for obtaining the processed result $h_i'(\theta)$. The pseudo-code of the attention mechanism is as follows:

Algorithm 1 The pseudo-code of the attention mechanism

Input: time series data $x_{i,t-s_i}$, DeepAR network layer hidden state dimension $h_{i,t-s_i-1}$

Parameters: number of CNN filter *d*, CNN feature dimension *j*

Out: $h_i'(\theta)$

For input $x_{i,t-s_i}$ and $h_{i,t-s_i-1}$, obtain target value $h_i(\theta)$ using the DeepAR network layer.

$$Z_i^o(\theta) = \sigma\left(W_o(\theta)\left[x_{i,t-s_i}, h_{i,t-s_i-1}\right] + b_o(\theta)\right)$$

$$C_i(\theta) = \sum_{i=1}^{I}\left(f_i(\theta) \circ C_{i-1}(\theta) + Z_i^r(\theta) \circ C_i'(\theta)\right)$$

$$h_i(\theta) = Z_i^o(\theta) \circ \tanh\left(C_i(\theta)\right)$$

Initialise the empty matrix H^C

$H^C \leftarrow$ InitialSolution( )

for $j \in \{1,2,\ldots,d\}$ and $k \in \{1,2,\ldots,n\}$ **do**

Calculate the convolutional value of the k^{th} row vector and the j^{th} filter

$$H_{k,j}^C(\theta) = \sum_{l=1}^{w} H_{k,t\text{-}w\text{-}1+l}(\theta) \times C_{j,w-w+l}(\theta)$$

We put all $H_{k,j}^C(\theta)$ into a pre-set matrix H^C

$H^C \leftarrow$ append $\left(H_{k,j}^C(\theta)\right)$

end

Substituting the k^{th} row of H^C $\left(H_{k,j}^C(\theta)\right)$ and $h_i(\theta)$ into score function $f(\cdot)$, the mechanism weight $A_k(\theta)$ can be obtained. Note that we use the sigmoid activation function.

$$f\left(H_k^C(\theta), h_i(\theta)\right) = \left(H_k^C(\theta)\right)^T W_a(\theta) h_i(\theta) \quad s.t. \quad W_a(\theta) \in \mathbb{R}^{d \times n}$$

$$A_k(\theta) = \sigma\left(f\left(H_k^C(\theta), h_i(\theta)\right)\right)$$

The row vectors of H^C are weighted by $A_k(\theta)$ to obtain the context vector $v_i(\theta)$, then we integrate $v_i(\theta)$ and $h_i(\theta)$ to update and obtain $h_i'(\theta)$

$$v_i(\theta) = \sum_{k=1}^{n} A_k(\theta) H_k^C(\theta) \quad s.t. \quad v_i(\theta) \in \mathbb{R}^d$$

$h_i'(\theta) = W_h(\theta)h_i(\theta) + W_v(\theta)v_i(\theta)$ $s.t.$ $W_h(\theta) \in \mathbb{R}^{n\times n}$ and $W_v(\theta) \in \mathbb{R}^{n\times d}$

Output: $h_i'(\theta)$

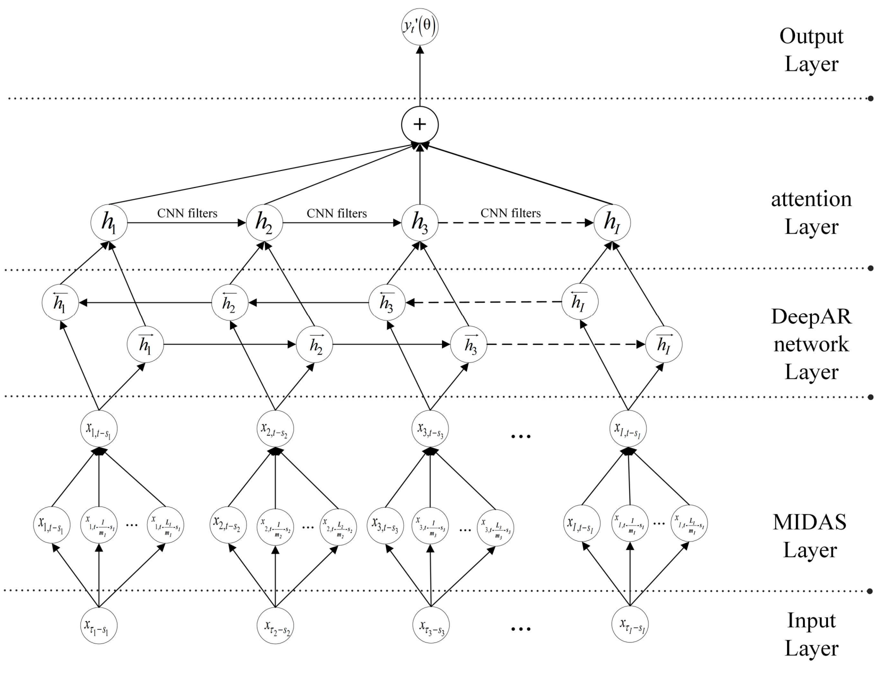
3.2.5 Output layer

Following the methods, port throughput mixed-frequency data mining is accomplished. The primary purpose of the output layer is to create the ADM model for the final port throughput forecast using the mixed-sampling frequency. In that example, the following equation provides the model's output $y_t'(\theta)$.

$$y_t'(\theta) \sim P\left(y_t'(\theta) \middle| \Omega_t\right) \tag{11}$$

where $t = t_0,\ldots,T$. The ADM model's schematic is seen in Figure 1.

Figure 1 Schematic diagram of the ADM model



4 Practice-based validation study

In this section, we analyse the actual forecast outcomes and estimation accuracy of the ADM model for China's port cargo throughput across various dimensions. A comparison between the ADM model and the attention-DeepAR (AD) model, which processes the same frequency, is conducted in this study. Additionally, we also present an analysis of the DeepAR model using real shipping data. It is worth noting that the converted high frequency predictors into low frequency data are used to ensure consistency with the predictor in the traditional DeepAR model and the AD model, which involves a sample averaging method. The study demonstrates the effectiveness of mixed-frequency alignment technology and attention mechanism. Furthermore, we contrast the ADM model with the MIDAS-DeepAR (MD) model to reveal the advantageous interaction between the optimised neural network parameters and attention mechanism. And we opted for the LSTM model as a representative of conventional machine learning techniques, primarily due to its widespread application in time-series forecasting.

4.1 The data

This work makes use of the ADM model to investigate potential high-dimensional, nonlinear temporal patterns to address the problem of interpreting mixed-frequency data when low frequency data are predicted using high frequency data. Six shipping-related data sets, covering the period from January 01, 2008 to January 31, 2023, make up the collection of data used in this study, which was gathered via a publicly accessible data platform. These data sets include indicators such as the daily high frequency Shanghai Stock Exchange Composite Index, the daily high frequency Brent Oil Futures and the daily high frequency Crude Oil WTI Futures, the monthly low frequency China's port cargo throughput, the weekly high frequency China Coastal Bulk Freight Index, and the weekly high frequency China Container Freight Index. Tables 1 and 2 include comprehensive information on shipping indicators.

Table 1 Description of variables for China's monthly port throughput forecast

Variable	*Definitions*	*Frequency*	*Data period*
CPCT	China's port cargo throughput (tons)	Monthly	2008M1-2023M1
CCBFI	China Coastal Bulk Freight Index	Weekly	2008W1-2023W4
CCFI	China Container Freight Index	Weekly	2008W1-2023W4
SSEC	Shanghai Stock Exchange Composite Index	Daily	2008D1-2023D31
BOF	Brent Oil Futures	Daily	2008D1-2023D31
COWF	Crude Oil WTI Futures	Daily	2008D1-2023D31

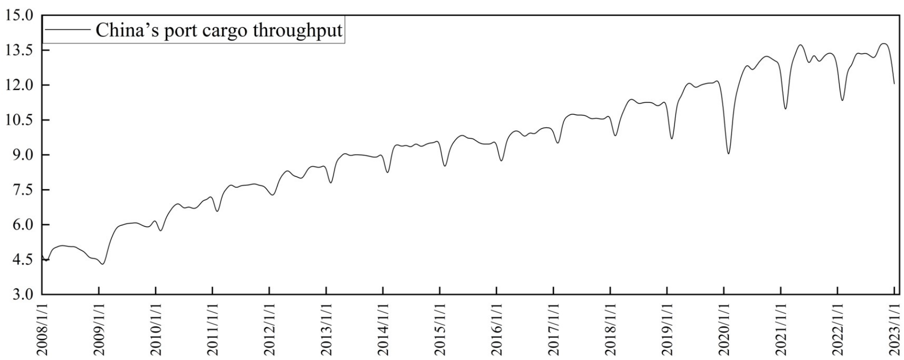
Notes: (1) We collect China's port cargo throughput data from the eastmoney website (https://data.eastmoney.com/cjsj/hyzs_list_EMI00108258.html).
(2) We download weekly China Coastal Bulk Freight Index data from SHANGHAI SHIPPING EXCHANGE website (https://www.sse.net.cn/index/singleIndex?indexType=cbfi).
(3) We download weekly China Coastal Freight Index data from MacroMicro website (https://sc.macromicro.me/collections/4356/freight/947/commodity-ccfi-scfi).
(4) The remaining data are available on the investing website (https://www.investing.com).

The evolution of China's port cargo throughput is also shown in Figure 2 from January 2008 to January 2023. China's port cargo throughput displays high uncertainty and nonlinear features, as seen in the image, which makes it difficult for conventional forecasting techniques to estimate with accuracy. In this study, we examine the forecast accuracy and distribution fitting impact of China's port cargo throughput under the ADM model and compare it to the MD, AD, and DeepAR models.

Table 2 Summary statistics of the transformed variables for China's monthly port throughput forecast

Variable	*Obs.*	*Mean*	*Max*	*Min*	*Std. dev.*
CPCT	181	9.4220	13.8500	4.1181	2.5537
CCBFI	724	1.1934	28.8692	0.7710	0.0341
CCFI	724	0.1194	35.8791	0.0632	0.0656
SSEC	3982	0.2914	0.5498	0.1707	0.0562
BOF	3982	0.0072	0.0145	-0.0038	0.0024
COWF	3982	0.0078	0.0146	0.0019	0.0026

Figure 2 The time series plot of the China's port cargo throughput



4.2 Evaluate performance

The ADM model is constructed in this study, using data from January 2008 to December 2019 as in-sample data, and data from Jan. 2020 to Jan. 2023 is employed as out-of-sample data to evaluate the forecasting performance. The robustness of the ADM model is frequently verified using the rolling window approach. We make pseudo forecasts on out-of-sample data by using a rolling forecasting scheme. The initial estimation period is from Jan. 2008 to May 2019 with the fixed window size of 137 months. Then estimation period moves along the whole data set and ends on the 1st month 2023. For each rolling, DeepAR, AD, MD and ADM are estimated, and the corresponding evaluation index are calculated with three different high frequency forecast horizons $s_i = \{1/22, 22/22, 44/22\}$ which correspond to low frequency forecast horizons $h = \{1, 1, 2\}$. For the conventional AD model and DeepAR model based on the same observed frequency, we produce monthly predictors by simply averaging the weekly and daily predictors respectively. All high frequency explanatory variables are simultaneously delayed by the same period and the maximum lag order $J = 11$ (i.e., $j = 1, 2, \ldots, 11$) to keep things simple.

In this study, three different performance evaluation indexes, forecast error evaluation index, probability forecasting evaluation index and model difference importance determination index, are used to evaluate the model's credibility.

1 Indicators for evaluating forecast inaccuracy. Mean absolute error (MAE), root mean square error (RMSE), and their improvement degree Δ are used to evaluate the model's forecast accuracy in comparison to the baseline model in this study. The following are the formulas:

$$MAE(\theta)=\frac{1}{T}\sum_{t=1}^{T}\left|y_t'(\theta)-y_t(\theta)\right| \tag{12}$$

$$RMSE(\theta)=\sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(y_t(\theta)-y_t'(\theta)\right)^2} \tag{13}$$

$$\Delta_{Model}^{MAE}=\frac{MAE_{Model}(\theta)-MAE_{DeepAR}(\theta)}{MAE_{DeepAR}(\theta)}\times 100\% \tag{14}$$

$$\Delta_{Model}^{RMSE}=\frac{RMSE_{Model}(\theta)-RMSE_{DeepAR}(\theta)}{RMSE_{DeepAR}(\theta)}\times 100\% \tag{15}$$

where Δ represents the percentage of improved performance of the model compared to the baseline model (DeepAR model), Model represents the ADM model, AD model, and MD model, $y_t'(\theta)$ is the estimated value of the true quantile $y_t(\theta)$. Based on the experiment and research of scholar Salinas et al. (2017), $\theta = 0.5$ is used as the estimated value of the quantile-type model. Generally, the smaller the values of MAE, MAPE, and RMSE, the better the performance of the model.

2 Performance evaluation of probability forecasting. The prediction interval coverage probability (PICP) and prediction interval normalised average width (PINAW) are used to disclose the efficacy of probability port throughput forecasting and quantify the impact of model interval forecasting to make up for the regret of Xu's study (Xu et al., 2021). As a reliability indicator, PICP shows the likelihood that the genuine port throughput number will fall within the expected range. The equation reads as follows:

$$PICP=\frac{1}{T}\sum_{t=1}^{T}I_t \tag{16}$$

$$I_t=\begin{cases}1, y_t\in\left[W_{t,1},W_{t,Z}\right]\\ 0, y_t\notin\left[W_{t,1},W_{t,Z}\right]\end{cases} \tag{17}$$

where $[W_{t,1}, W_{t,Z}]$ represents the predicted interval within the predetermined coverage range of port throughput, I_t is a Boolean variable. When the true value of port throughput at the t^{th} time falls within the predicted interval, it is assigned a value of 1, and vice versa, it is assigned a value of 0. The larger the PICP, the more port throughput forecasting is covered by the predicted interval, and the more convincing the forecasting results are. In addition to evaluating reliability, the PINAW is used to comprehensively evaluate the performance of interval forecasting. The formula is as follows:

$$PINAW = \sum_{t=1}^{T} \frac{W_{t,Z} - W_{t,1}}{TD} \tag{18}$$

The port throughput target values' maximum and lowest goal values are represented here by D, and the lower the PINAW, the more accurate the findings of the port throughput forecast are.

3 Metrics for evaluating the importance of model discrepancies. The statistical significance of the disparities between two sets of forecasting is investigated in this study using Diebold and Mariano's DM test, which has been enhanced by Franses (2016)

$$H_0 : E(d_t) = 0, \forall t \tag{19}$$

where $d_t = \rho_\theta(e_{1t}) - \rho_\theta(e_{2t})$ is the loss error between the true and predicted values of port throughput for two models, and $\rho_\theta(\cdot)$is a loss function for the forecast error e, which is assumed to be mean squared error (MSE) in this instance. The following formula yields the DM test statistic for multi-step ahead forecasts.

4.3 Empirical results

4.3.1 Deterministic forecast results and analysis

Deterministic forecast outcomes from all models are examined using markers for evaluating forecast error. For China's port cargo throughput data, Table 3 compares the forecast errors of each model in multi-step ahead.

Table 3 Results of RMSE and MAE for China's port cargo throughput forecast

s_i	*Model*	*RMSE*	*MAE*	$\Delta^{\mathrm{MAE}}_{Model}$	$\Delta^{\mathrm{RMSE}}_{Model}$
s_i = 1/22	LSTM	4.9646	4.8249	13.20%	17.53%
	DeepAR	4.3856	4.1053	-	-
	AD	4.1399	3.9344	–4.10%	–5.60%
	MD	4.2062	3.9959	–2.66%	–4.09%
	ADM	*3.3996*	*3.1229*	*–23.93%*	*–22.48%*
s_i = 22/22	LSTM	4.8995	4.7579	11.72%	15.90%
	DeepAR	4.3856	4.1053	-	-
	AD	4.1398	3.9343	–4.10%	–5.60%
	MD	4.2060	3.9957	–2.67%	–4.10%
	ADM	*3.4132*	*3.4132*	*–23.57%*	*–22.17%*
s_i = 44/22	LSTM	4.8661	4.7235	10.96%	15.06%
	DeepAR	4.3856	4.1053	-	-
	AD	4.1398	3.9343	–4.09%	–5.60%
	MD	4.2062	3.9960	–2.66%	–4.09%
	ADM	*3.3994*	*3.1227*	*–23.94%*	*–22.49%*

Table 3 illustrates that the LSTM model's performance exhibited a decline of over 10% relative to our baseline model, DeepAR. This outcome positions LSTM as the least effective model among those evaluated. Consequently, this paper will now pivot to a detailed examination of the baseline model, DeepAR, along with its enhanced iterations, to elucidate the factors contributing to their differing levels of accuracy. Given the observed limitations of LSTM in comparison to DeepAR, our focus shifts towards the ADM model, which represents a more advanced approach in this realm. Table 3 demonstrates that the ADM model outperforms the other three models in terms of RMSE and MSE accuracy. Furthermore, an analysis of the multi-step ahead errors reveals a consistent error pattern, underlining the ADM model's high degree of forecast accuracy and resilience. This consistent performance, especially in multi-step forecasting, highlights the potential of advanced modelling techniques in capturing complex temporal dynamics. The value of Δ is negative, it means that the accuracy of the AD, MD, and ADM models have all increased as compared to the benchmark model (DeepAR model), with a maximum increase of 23.93%. Further analysis shows that compared with the same frequency models (DeepAR model and AD model), the ADM model has lower MAE and RMSE values, proving that the introduction of mixed-frequency data sampling method can improve the prediction performance. The reason for this is that the introduction of mixed-frequency data sampling method can effectively retain and utilise the information of the predictive factors, avoiding the information loss caused by converting high frequency predictive factors into low frequency monthly data, such as China's port cargo throughput. Secondly, by comparing the forecast accuracy of the ADM model with the MD model, it is found that under the co-action of attention mechanism and neural network, not only the complex and nonlinear relationship of China's port cargo throughput with fluctuations is overcome, but also the supervision mechanism of data fluctuations and time sensitivity is considered. Moreover, compared with the DeepAR model and MD model without improved attention mechanism, the MAE and RMSE values of the AD model and ADM model are lower, proving that the introduction of a CNN filter attention mechanism based on time sensitivity and variable selection can optimise the forecast accuracy to a certain extent.

4.3.2 Probability prediction results and analysis

The DeepAR, AD, MD, and ADM models were evaluated using the PICP and PINAW evaluation standards to more thoroughly quantify the models' capacity to characterise China's port cargo throughput in uncertain contexts.

Table 4 Probability prediction performance evaluation of multi-step ahead forecast

s_i	*DeepAR*		*AD*		*MD*		*ADM*	
	PICP	*PINAW*	*PICP*	*PINAW*	*PICP*	*PINAW*	*PICP*	*PINAW*
1/22	100%	8.8147	100%	6.9498	100%	6.8410	100%	*4.6726*
22/22	100%	8.8147	100%	6.9498	100%	6.8412	100%	*4.6728*
44/22	100%	8.8147	100%	6.9798	100%	6.8411	100%	*4.6727*

Once external uncertainty increases, the forecast errors could amplify due to information loss, which requires validating the model's ability to avoid external risks. Based on the PICP (90% threshold) and PINAW indicators from Table 4, all four models meet the predetermined confidence level and remain relatively stable in multi-step ahead forecasts. The AD and MD models' smaller bandwidth than the benchmark model shows that the benchmark model's capacity for risk avoidance has been enhanced by their frequency and attention mechanism. This enhancement foresees the capacity of the ADM model to incorporate these two technologies. The experimental findings demonstrate that the ADM model has a narrower bandwidth than the other three models while ensuring coverage rates. This highlights the potential of the ADM model to significantly enhance forecasting performance by effectively reducing forecasting uncertainty.

To showcase the predictive ability of the ADM model, Figure 3 plots the multi-step ahead forecasts (s_i = 1/22) results of China's port cargo throughput from January 2020 to January 2023. As shown in Figure 3, the ADM model effectively captures the fluctuations within all forecasting ranges, and all China's port cargo throughput values from January 2020 to January 2023 are contained within the 90% level confidence interval.

Figure 3 China's port cargo throughput forecasts on test data with the ADM model (s_i = 1/22)

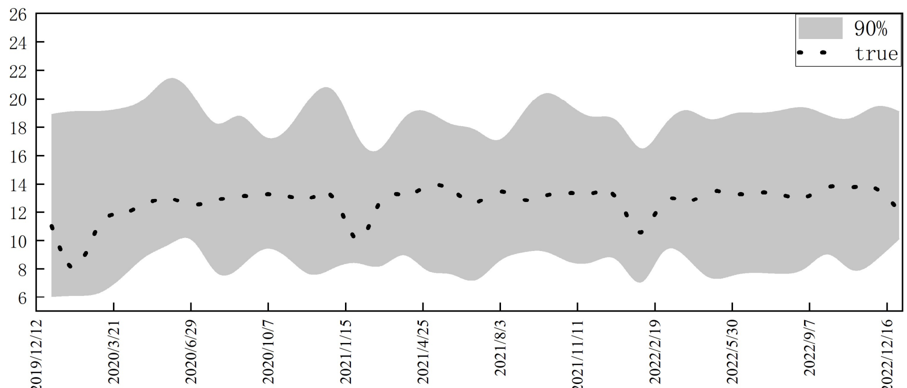


Figure 4 China's port cargo throughput forecasts on test data with the ADM model (s_i = 22/22)

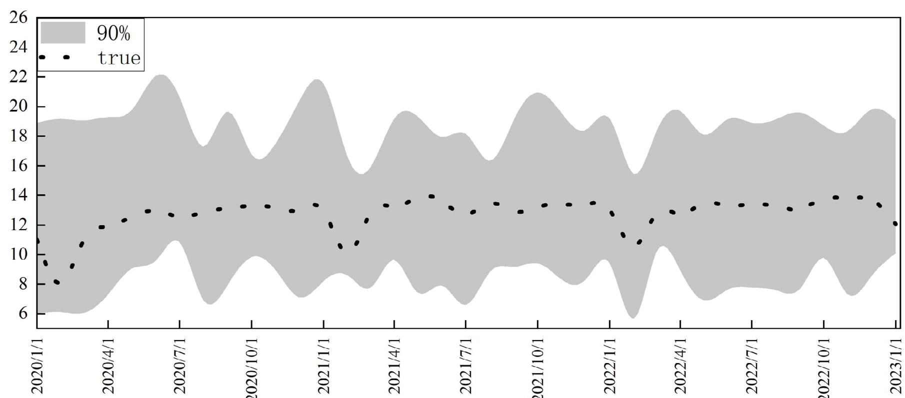
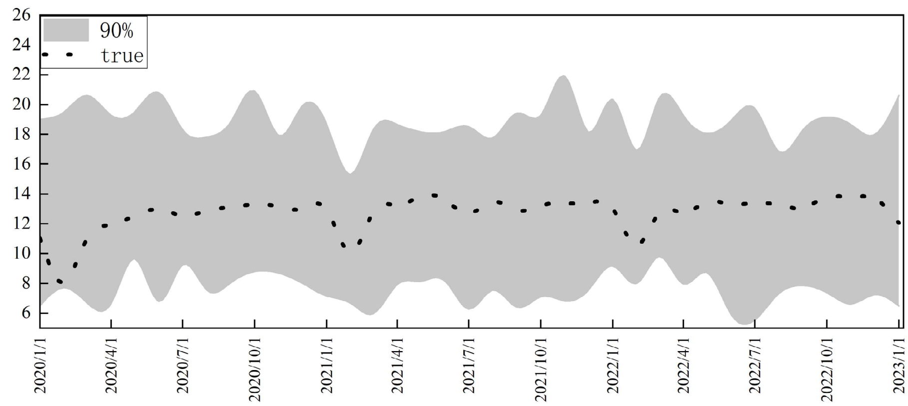


Figure 5 China's port cargo throughput forecasts on test data with the ADM model (s_i = 44/22)



4.3.3 The predicted results and analysis among models

This study applies the Diebold-Mariano (DM) test to evaluate if there is a statistically significant difference in the out-of-sample forecast accuracy between each pair of models to establish whether the comparison of forecast accuracy between the models has statistical relevance.

Table 5 The DM test results of the forecasting model

Model 2	s_i	*Model 1*	
		AD model	*MD model*
ADM model	1/22	2.7217***	3.3526***
	22/22	2.5942***	3.2535***
	44/22	2.7118***	3.3503***

Notes: (1) The numbers in the table represent the DM statistic;
(2) '***' indicates significance at the 1% level, namely, $p < 0.01$.

The Diebold-Mariano test's statistically significant findings show that Model 2 outperforms Model 1 in terms of performance. At the 1% significance level, Table 5 demonstrates that the ADM model performs much better than the AD model and MD model. The ADM model outperforms the AD model in terms of predictive capacity, demonstrating the value of aligning mixed-frequency data and concentrating on the useful information conveyed by high frequency variables. A filter-based attention mechanism for feature selection may capture significantly more valuable influencing aspects and boost the temporal correlation of port throughput data, as shown by the ADM model's greater forecast accuracy when compared to the MD model.

5 Conclusions

As a key indicator in tracking and evaluating maritime transport performance, port throughput plays an essential role in optimisation decision-making, resource allocation and control planning. To capture the uncertainty and nonlinearity of port throughput against exogenous and endogenous disruptions, a robust forecasting method (ADM) employing the MIDAS, attention mechanism, and deep learning methods was introduced in this study. Then, the proposed method was compared with the baseline model (DeepAR) and the candidate models (AD model, MD model, and ADM model). Moreover, six performance indicators, comprising MAE, RMSE, $\Delta_{Model}^{\mathrm{MAE}}$, $\Delta_{Model}^{\mathrm{RMAE}}$, PICP and PINAW, were employed to evaluate the models. The DM tests were used to further confirm the key empirical findings. The evaluation results indicated that the proposed method outperforms all other models. Our model combined the MIDAS and attention mechanism, improving forecasting accuracy and efficiency by effective use of endogenous and exogenous mixed-frequency data of port throughput, tracking and fitting the fluctuation trend of port throughput, and quantifying port uncertainty.

Moreover, we evaluate the ADM model in real-world applications to forecast China's monthly port cargo throughput. As expected, the practice-based validation results demonstrate that the ADM model outperforms benchmark models. the proposed model exhibits strong skills for detecting latent nonlinear connections between variables in mixed-frequency sampling data. It can automatically detect the temporal weighting of sequences, and promptly collect external disturbance elements for multi-step ahead probabilistic forecasts. In summary, the proposed ADM model provides more timely, accurate, and comprehensive predictions for decision-making. However, the applicability of ADM model to other port remains untested, given China's unique maritime policies and infrastructure. Future studies should consider validating the model's effectiveness across a broader spectrum of international ports and integrating real-time data to enhance responsiveness to global economic events.

The proposed ADM model in this study is highly adaptable; it may be expanded to related prediction studies in the industries of energy, transportation, and electricity based on the same theory and prediction approach. In addition, other frequency alignment techniques, like R-MIDAS, can be used in model construction, considering the existence of accurate real-time prediction of low frequency response variables based on high frequency explanatory variables, which is also worth further investigation in future studies.

Acknowledgements

The research is supported by the National Social Science Foundation of China under Grant No. 20BGL112.

References

Dong, M., Wu, H., Hu, H., Azzam, R., Zhang, L., Zheng, Z. and Gong, X. (2020) 'Deformation prediction of unstable slopes based on real-time monitoring and deepar model', *Sensors*, Vol. 21, No. 1, p.14.

Du, P., Wang, J., Yang, W. and Niu, T. (2019) 'Container throughput forecasting using a novel hybrid learning method with error correction strategy', *Knowledge-Based Systems* [online] https://doi.org/10.1016/j.knosys.2019.07.024.

Franses, P.H. (2016) 'A note on the mean absolute scaled error', *International Journal of Forecasting*, Vol. 32, No. 1, pp.20–22.

Fung, M.K. (2002) 'Forecasting Hong Kong's container throughput: an error-correction model', *Journal of Forecasting*, Vol. 21, No. 1, pp.69–80.

Gagliardini, P., Ghysels, E. and Rubin, M. (2017) 'Indirect inference estimation of mixed-frequency stochastic volatility state space models using MIDAS regressions and ARCH models', *Journal of Financial Econometrics*, Vol. 15, No. 4, pp.509–560.

García, T.R., Cancelas, N.G. and Soler-Flores, F. (2014) 'The artificial neural networks to obtain port planning parameters', *Procedia-Social and Behavioral Sciences*, Vol. 162, pp.168–177.

Ghysels, E., Santa-Clara, P. and Valkanov, R. (2006) 'Predicting volatility: getting the most out of return data sampled at different frequencies', *Journal of Econometrics*, Vol. 131, No. 2, pp.59–95.

Ghysels, E., Sinko, A. and Valkanov, R. (2007) 'MIDAS regressions: further results and new directions', *Econometric Reviews*, Vol. 26, No. 1, pp.53–90.

Gosasang, V., Chandraprakaikul, W. and Kiattisin, S. (2011) 'A comparison of traditional and neural networks forecasting techniques for container throughput at Bangkok port', *The Asian Journal of Shipping and Logistics*, Vol. 27, No. 3, pp.463–482.

Justice, V., Bhaskar, P., Pateman, H., Cain, P. and Cahoon, S. (2016) 'US container port resilience in a complex and dynamic world', *Maritime Policy & Management*, Vol. 43, No. 2, pp.179–191.

Li, X., Hu, Y., Su, X. and Shao, B. (2023) 'A data-driven newsvendor problem with shifting demand: a deep autoregressive model with attention mechanism', *Journal of Engineering Science & Technology Review*, Vol. 16, No. 3, pp.74–83.

Min, K.C. and Ha, H.K. (2014) 'Forecasting the Korea's port container volumes with SARIMA model', *Journal of Korean Society of Transportation*, Vol. 32, No. 6, pp.600–614.

Moscoso-López, J.A., Urda, D., Ruiz-Aguilar, J.J., Gonzalez-Enrique, J. and Turias, I.J. (2021) 'A machine learning-based forecasting system of perishable cargo flow in maritime transport', *Neurocomputing*, Vol. 452, pp.487–497.

Patil, G.R. and Sahu, P.K. (2016) 'Estimation of freight demand at Mumbai Port using regression and time series models', *KSCE Journal of Civil Engineering*, Vol. 20, No. 5, pp.2022–2032.

Rashed, Y., Meersman, H., Van de Voorde, E. and Vanelslander, T. (2017) 'Short-term forecast of container throughout: an ARIMA-intervention model for the port of Antwerp', *Maritime Economics & Logistics*, Vol. 19, No. 4, pp.749–764.

Ruiz-Aguilar, J.J., Urda, D., Moscoso-López, J.A., Gonzalez-Enrique, J. and Turias, I.J. (2020a) 'A freight inspection volume forecasting approach using an aggregation/disaggregation procedure, machine learning and ensemble models', *Neurocomputing*, Vol. 391, pp.282–291.

Ruiz-Aguilar, J.J., Urda, D., Moscoso-López, J.A., González-Enrique, J. and Turias, I.J. (2020b) 'Container demand forecasting at border posts of ports: a hybrid sarima-som-svr approach', in *Optimization and Learning: Third International Conference*, *OLA 2020*, Cádiz, Spain, 17–19 February, Springer International Publishing, pp.69–81.

Salinas, D., Flunkert, V., Gasthaus, J. and Januschowski, T. (2017) 'DeepAR: Probabilistic forecasting with autoregressive recurrent networks', *International Journal of Forecasting*, [online] https://doi.org/10.1016/j.ijforecast.2019.07.001.

Schulze, P.M. and Prinz, A. (2009) 'Forecasting container transhipment in Germany', *Applied Economics*, Vol. 41, No. 22, pp.2809–2815.

Shankar, S., Ilavarasan, P.V., Punia, S. and Singh, S.P. (2020) 'Forecasting container throughput with long short-term memory networks', *Industrial Management & Data Systems*, Vol. 120, No. 3, pp.425–441.

Shih, S.Y., Sun, F.K. and Lee, H. (2019) 'Temporal pattern mechanism for multivariate time series forecasting', *Machine Learning*, Vol. 108, No. 9, pp.1421–1441.

Silver, E.A., Pyke, D.F. and Thomas, D.J. (2016) *Inventory and Production Management in Supply Chains*, CRC Press, Boca Raton [online] https://www.taylorfrancis.com/books/mono/10.1201/9781315374406/inventory-production-management-supply-chains-edward-silver-david-pyke-douglas-thomas.

Tang, L., Yu, L. and He, K. (2014) 'A novel data-characteristic-driven modelling methodology for nuclear energy consumption forecasting', *Applied Energy*, Vol. 128, pp.1–14.

Tian, X., Xu, L., Liu, L. and Wang, S. (2010) 'Analysis and forecasting of port logistics using TEI@ I methodology', in Wang, J. and Wang, S. (Eds.): *Business Intelligence in Economic Forecasting: Technologies and Techniques*, pp.248–264, IGI Global, Hershey, Pennsylvania, USA, https://doi.org/10.4018/978-1-61520-629-2.

United Nations Conference on Trade and Development (UNCTAD) (2022) *Review of Maritime Transport 2021*, United Nations Publications, Geneva [online] https://unctad.org/system/files/official-document/rmt2022_en.pdf (accessed 25 April 2024).

Xiao, J., Xiao, Y., Fu, J. and Lai, K.K. (2014) 'A transfer forecasting model for container throughput guided by discrete PSO', *Journal of Systems Science and Complexity*, Vol. 27, No. 1, pp.181–192.

Xiao, Y., Xie, M., Hu, Y. and Yi, M. (2023) 'Effective multi-step ahead container throughput forecasting under the complex context', *Journal of Forecasting* [online] https://doi.org/10.3390/w14223672.

Xie, Q., Lang, Q. and Liu, X. (2023) 'Stock price forecasting based on feature fusion deepar model', *2023 35th Chinese Control and Decision Conference (CCDC)*, IEEE, pp.4242–4248.

Xu, L., Zou, Z. and Zhou, S. (2022a) 'The influence of COVID-19 epidemic on BDI volatility: an evidence from GARCH-MIDAS model', *Ocean & Coastal Management* [online] https://doi.org/10.1016/j.ocecoaman.2022.106330.

Xu, S., Zou, S., Huang, J., Yang, W. and Zeng, F. (2022b) 'Comparison of different approaches of machine learning methods with conventional approaches on container throughput forecasting', *Applied Sciences* [online] https://doi.org/10.3390/app12199730.

Xu, Q., Liu, S., Jiang, C. and Zhuo, X. (2021) 'QRNN-MIDAS: a novel quantile regression neural network for mixed sampling frequency data', *Neurocomputing*, Vol. 457, pp.84–105.

Xu, Q., Wang, L., Jiang, C. and Liu, Y. (2020) 'A novel (U) MIDAS-SVR model with multi-source market sentiment for forecasting stock returns', *Neural Computing and Applications*, Vol. 32, No. 10, pp.5875–5888.

Yu, L., Wang, Z. and Tang, L. (2015) 'A decomposition – ensemble model with data-characteristic-driven reconstruction for crude oil price forecasting', *Applied Energy*, Vol. 156, pp.251–267.

Zhang, K., Qu, Z., Dong, Y., Lu, H., Leng, W., Wang, J. and Zhang, W. (2019) 'Research on a combined model based on linear and nonlinear features – a case study of wind speed forecasting', *Renewable Energy* [online] https://doi.org/10.1016/j.renene.2018.05.093.