ITC 2/52
Information Technology and Control
Vol. 52 / No. 2 / 2023
pp. 336-357
DOI 10.5755/j01.itc.52.2.32804

Forecasting Secondhand Tanker Price Through Wavelet Neural Networks Based on Adaptive Genetic Algorithm

Received 2022/11/18 | Accepted after revision 2022/12/23

HOW TO CITE: Ma, X. (2023). Forecasting Secondhand Tanker Price Through Wavelet Neural Networks Based on Adaptive Genetic Algorithm. *Information Technology and Control*, 52(2), 336-357. https://doi.org/10.5755/j01.itc.52.2.32804

Forecasting Secondhand Tanker Price Through Wavelet Neural Networks Based on Adaptive Genetic Algorithm

Ma Xingyu
School of Economics and Management, Shanghai Maritime University, China;
phone: +86 18810325839; fax: +86 18810325839; e-mail: 202130710069@stu.shmtu.edu.cn

Corresponding author: 202130710069@stu.shmtu.edu.cn

Seaborne crude oil remains the main source of energy in the modern world in terms of volume, accounting for nearly half of all internationally traded crude oil. The shipping market is already characterized by high volatility, coupled with the impact of COVID-19 lockdown and geopolitics events. Price forecasting has become a necessary and challenging task for shipowners and other stakeholders. In the shipping market forecasting literature, the usual focus is on the newbuilding ship price or freight rate. A limited number of literature is for secondhand tanker price. On the other hand, there is few literature that use wavelet neural networks based on adaptive genetic algorithm (AGA-WNN) to predict shipping market. This paper mainly studies the application of the hybrid model to secondhand price prediction of 5 kinds of tanker sizes. The performance of AGA-WNN on time series of 10 and 15 years is compared with the basic performance provided by the six benchmark models, using three error metrics and two statistical tests. We can point out that AGA-WNN provides encouraging and promising results, outperforming the baseline models in both accuracy and robustness. It can be said that AGA-WNN gives the best overall predictive performance.

KEYWORDS: Secondhand Tanker Price, Forecasting, Wavelet Neural Networks, Adaptive Genetic Algorithm.

1. Introduction

The characteristics of high volatility in shipping market can be attributed to intrinsic characteristics and external environment uncertainty [41], for example, economy cycle, trade war, national policies, black swan events and so on. Recently, the Russia-Ukraine conflict has made a significant impact on the oil and tanker markets, with geopolitical events affecting a range of factors including energy prices, tanker markets and seaborne oil trading patterns. Tanker freight rates on ex-Russia routes spiked since late February 2022. Aframax earnings on the Baltic-UKC route surged to over $230,000/day in early March, compared to an average of $10,000/day in Jan-Feb [11]. The price of secondhand ships can fluctuate by millions of dollars a month.

Under the backdrop of rising oil prices and COVID-19 lockdown, trade situation is volatile and policy may change further, making forecast difficult. Newbuilding price, secondhand price, charter rate and scrap value are four important markets that influence and determine investment decisions, costs and profits in the shipping industry [18]. An accurate judgment of the inflection point in time series trends allows shipowners to buy low and sell high, achieving better profits. Hence, shipowners need to make more prudent decisions.

Many researchers have done a lot of work in shipping market forecasting. [36] applies the support vector regression (SVR) model, a novel and innovative forecasting framework, to the empirical study of newbuilding ship (dry bulk and tanker) price forecasts for the first time. [33] considers two alternative neural networks specifications, NN-MLP and NN-RBF, to predict the period charter rates of VLCC. The results show that the neural networks can provide better per- formance than the traditional method (ARIMA) in dealing with tanker period or spot charter rates, thus confirming the previous empirical evidence for the spot market from [28] and [30]. The usual focus in literature is on the newbuilding ship price or freight rate while a limited number of literature is for secondhand ship price. Sale-and-purchase transactions of secondhand ships is a main source of profit for shipowners whose profitability depends on the decision-making occasion [9]. Consequently, secondhand ship market plays an important economic role in shipping industry. In this paper, the existing efforts are predominantly aimed at the secondhand tanker market and the results will have practical significance for relevant stakeholders, such as investors who need to determine the appropriate time of investment and withdrawal.

The purpose of our forecasting work is to learn linear or nonlinear functions of historical prices and to give reliable predictions of unseen data. All the models we propose are based on the fact that the temporal patterns of secondhand tanker price contain useful information for predicting their future movements. It is known to all that the shipping market is highly dynamic. Therefore, it is difficult to accurately model such highly volatile features with traditional models. In order to overcome this limitation, a hybrid model consisting of two algorithms (wavelet neural networks and adaptive genetic algorithm) is proposed. This machine learning model has shown significant success in a variety of applications, but not yet in the shipping market.

Several contributions of this paper are summarized as follows: 1) To the author's best knowledge, wavelet neural networks based on adaptive genetic algorithm (AGA-WNN) is applied to secondhand tanker market prediction for the first time. 2) In terms of accuracy, the result of AGA-WNN is much better than the traditional linear models and slightly better than the existing machine learning models. In terms of robustness, the result of AGA-WNN is superior to all the baseline models proposed. 3) All model errors are established according to out-of-sample accuracy. 4) In traditional shipping markets, shipowners often make decisions based on intuition and past experience. However, the volatility of shipping market and the uncertainty of economic development bring high degree of unreliability to the empirical decisions. Data-driven decision-making can objectively and accurately predict the shipping assets value and maritime business risk, which will enable shipping companies to determine better investment options and enter or exit the market at the right time.

The rest of this paper is organized as follows. Section 2 gives a literature review on shipping market, especially the secondhand tanker market, as well as forecasting methods. Section 3 briefly presents the baseline models used in this article, and introduces the hybrid prediction model combining WNN and AGA in detail. Data description and pre-processing steps are described in Section 4. Then, based on three error metrics and two statistical tests, we simulated each model in different time series test sets. Comparison results are shown in Section 5. Finally, the results are summarized. Conclusions, limitations and future research directions are briefly proposed.

2. Literature Review

Forecasting is a common research task. Research in [26] examines the weak form of market efficiency of Indian stock market from April 1, 1997 to March 31, 2010 with the efficient market hypothesis (EMH). The findings reveal that the stock prices fully reflect all information of the past. Therefore, revenues are not independent of time. In fact, nonlinear auto-correlation exists in most time series data. For shipping economics literature, freight rate, charter rate, newbuilding price, secondhand price, scrap value, shipping index, ship energy efficiency, carbon emission, etc., are frequently considered as research objects.

Research in [25] first studied the determinants of tanker freight rates through a supply-demand model. Later, [2, 6, 23] put forward a theoretical model in which the freight market and ship market are in an interdependent framework and secondhand ships are regarded as capital assets. Studies in [3, 4] apply this model to the world dry cargo market and tanker market, respectively, estimating an aggregated econometric model in which freight rate, newbuilding and secondhand price, and fleet size are jointly and dynamically determined. However, Tsolakis et al. criticizes the view of Beenstock. Newbuilding prices are driven more by supply factors, whereas secondhand prices are driven more by demand. In other words, newbuilding prices do not respond as quickly to changes in market conditions as secondhand ones. A simple present value equation, such as the Norwegian models, will suffice. Econometric models often have statistical defects that make their estimates biased. In order to overcome this problem, [38] provides a theoretical Error Correction model for secondhand ship prices.

Study in [24] compares the price risks between different ship sizes in the tanker industry, using autoregressive conditional heteroscedasticity (ARCH) models and presenting the fluctuation of shipping market as a time-varying process. Research in [1] tests the performance of vector equilibrium correction models (VECM) in predicting spot and forward prices of major shipping routes. A strong convergence between the forward rates and the spot rates is shown, i.e., the forward rates do help predict the spot rates. Study in [20] focuses on spot price prediction from two aspects: (1) multivariate models (VAR and VECM) and (2) univariate models (ARIMA, GARCH and E-GARCH), so as to obtain the best prediction model for each ship type (tankers and bulk carriers). In addition, the prediction results are modified by combinatorial method theory. In [8, 37], a multivariate vector autoregressive model (VARX) containing exogenous variables is established to improve the prediction accuracy of BDI. There are also other novel forecasting methods, for example, judgmental forecasting [14], copula-based multivariate models [42], fuzzy time series modelling approach [17] and popular machine learning algorithms [19, 39]. Although it is convenient to use the linear or nonlinear methods mentioned above, there are certain restrictions on the number of dependent and independent variables. Artificial neural network allows parameter estimation based on a large number of independent and dependent variables and has good generalization performance.

Recently, artificial neural networks is one of the most widely used machine learning algorithms, whose most attractive feature is strong nonlinear processing, self-learning and self-adaptation advantages. The generalization ability of ANN provides conditions for its potential in prediction. ANN cannot only be applied in shipping economy market [13, 15, 21, 31, 32, 40], but also provide help for ship technological design. Study in [5], respectively, uses ANN to estimate engine power and fuel consumption, and then estimate carbon dioxide emissions, for the recent tankers, bulk carriers and container ships built from 2015 to present.

Wavelet neural networks (WNN) is an improved version of BP neural networks, which combines the superiorities of wavelet analysis and neural networks. Therefore, WNN has strong approximation and fault-tolerance ability. According to the wavelet multiscale decomposition of time series, [27] reveals different time frequency variation patterns of dry

bulk shipping indices, and uses the wavelet as the activation function of neural networks. The initial time series is decomposed, predicted and combined to obtain the final predicted value. The application effect of neural networks in empirical research is always inspiring.

3. Methodologies

3.1. Research Framework

Firstly, the baseline models are introduced and the wavelet neural network based on adaptive genetic algorithm is described in detail. Then SARIMA in the baseline model is applied to obtain the autocorrelation order of all time series. The initial data set is transformed into a standard form with input independent variables and output dependent variables. After data pre-processing, all prediction models are simulated, three prediction error measures are calculated and two statistical tests are performed to compare the advantages and disadvantages of each model. Finally, the results are analyzed and the limitations are proposed.

To reiterate, the main purpose of this paper is to prove the effectiveness and superiority of the target algorithm for prediction. Time series prediction can be roughly divided into two types. If only the previous values of time series are used to predict its future values, it is called univariate time series prediction. If we use variables other than time series (i.e. exogenous variables) for prediction, it is called multivariable time series prediction. Considering our purpose, data acquisition and final results, we choose the former and determine that the information of time series is sufficient for prediction. The latter can be further discussed in future studies.

3.2. Brief Introduction of Baseline Models

Firstly, we suppose that the target time series is ϕ_t where t = 1, 2, … , N . Our prediction task is to calculate the values of $\phi_{N+1}, \phi_{N+2}, \cdots$ Then, a brief overview of each baseline model is demonstrated.

Seasonal naive forecast: Naive forecast uses the most recent value as its prediction; likewise, seasonal naive forecast anticipate the forecast to be the observation on corresponding position of the previous time period. The model is defined as the following:

$$\phi_t = \phi_{t-12}, \tag{1}$$

where $t > 12$. While the most simple method, seasonal naive forecast is quite robust for highly fluctuant time series because the temporal patterns in history may make no contribution to predict the future. Therefore, seasonal naive forecast is the most common standard for prediction tasks.

SARIMA: Main effective factors of time series generally include long-term trend (T), seasonal fluctuation (S), cyclical fluctuation (C) and irregular fluctuation (I), in other words, $\phi_t = f\left(T_t, S_t, C_t, I_t\right)$. The differential operation can help remove the impact of periodical trends, seasonal variation and multiplicative interaction factor so as to stabilize the time series. As a result, the SARIMA model, integrating seasonal autocorrelation, auto regression model (AR), differencing and moving average model (MA), is another popular forecasting method and has succeeded in many forcasting tasks. In this paper, the model $SARIMA\left(p,d,q\right)\times\left(P,D,Q\right)_{\pi}$ is defined as the following:

$$\nabla^d \nabla_{\pi}^D \phi_t = \frac{\Theta(B)\Theta_{\pi}(B)}{\Psi(B)\Psi_{\pi}(B)}\varepsilon_t, \tag{2}$$

where ∇ is differential operator, B is delay operator, $\nabla_{\pi}^D \phi_t$ represents the time series after D orders differencing with π steps, ε represents residual sequence, $\Psi\left(B\right) = 1 - \psi_1 B - \psi_2 B^2 - \cdots - \psi_p B^p$, $\Psi_{\pi}\left(B\right) = 1 - \psi_1 B^{\pi} - \psi_2 B^{2\pi} - \cdots - \psi_p B^{P\pi}$, $\Theta\left(B\right) = 1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_a B^q$, $\Theta_{\pi}\left(B\right) = 1 - \theta_1 B^{\pi} - \theta_2 B^{2\pi} - \cdots - \theta_Q B^{Q\pi}$, parameters ψ and θ need to be estimated, and order p, q, P and Q need to be optimized. The value of seasonal cycle ϖ generally equal one year.

Holt–Winters: Basing on exponential smoothing with two parameters presented by Holt, Winter had made improvement and constructed Holt–Winters exponential smoothing with three parameters in order to smooth linear trend and seasonal variation. The multiplicative seasonality model is defined as the following:

$$\begin{cases} \hat{\phi}_{t+1} = \alpha \dfrac{\phi_t}{s_{t-\pi}} + (1-\alpha)\left(\phi_{t-1} + r_{t-1}\right) \\ \hat{r}_t = \beta\left(\hat{\phi}_{t+1} - \hat{\phi}_{t-1}\right) + (1-\beta) r_{t-1} \\ \hat{s}_t = \gamma\left(\dfrac{\phi_t}{\hat{\phi}_{t+1}}\right) + (1-\gamma) s_{t-\pi} \end{cases}, \tag{3}$$

where r_t and s_t represent trend factor and season factor, respectively, whose initial value can be determined through several ways. α, β and γ all range from 0 to 1. The smoothing parameters are constantly iterated to adapt non-stationary time series and to predict short-term future.

Support Vector Regression (SVR): SVR belongs to machine learning algorithms. Given sample data $D=\{(X_1, y_1), (X_2, y_2), \ldots, (X_n, y_n)\}, X_i \in \mathbb{R}^p$, SVR should learn a function $\hat{y}=\omega^T X+b$ which makes the maximum margin between sample points and the hyperplane less than upper bound ε. The primal problem could be modeled as the following:

$$\min_{\omega,b,\xi^U,\xi^L} loss = \min_{\omega,b,\xi^U,\xi^L} \frac{1}{2}\omega^T\omega + C\sum_{i=1}^{n}\left(\xi_i^U + \xi_i^L\right), \tag{4}$$

with the constraints defined as,

$$\begin{aligned} & y_i - \left(\omega^T X_i + b\right) \le \varepsilon + \xi_i^U \\ & \left(\omega^T X_i + b\right) - y_i \le \varepsilon + \xi_i^L \\ & C>0, \xi_i^U \ge 0, \xi_i^L \ge 0, \quad i=1,2,\ldots,n \end{aligned},$$

where ξ_i^U, ξ_i^L denote slack variables and C denotes regularization constant. With regard to optimization parameters ω and b, by introducing Lagrange multipliers λ^U, λ^L, μ^U, μ^L, the primal problem with constraints is transformed into the primal problem without constraints. On the other hand, this convex optimization problem meet KKT conditions, so the primal problem without constraints can be further transformed into the dual problem. The dual problem could be modeled as the following:

$$\begin{aligned} & \max_{\lambda^U,\lambda^L,\mu^U,\mu^L} \min_{\omega,b} \mathcal{L}(\omega,b,\xi^U,\xi^L,\lambda^U,\lambda^L,\mu^U,\mu^L) \\ & = \max_{\lambda^U,\lambda^L,\mu^U,\mu^L} \min_{\omega,b} \frac{1}{2}\omega^T\omega + C\sum_{i=1}^{n}\left(\xi_i^U+\xi_i^L\right) \\ & \quad + \sum_{i=1}^{n}\lambda_i^U\left(y_i - \left(\omega^T X_i + b\right) - \varepsilon - \xi_i^U\right) - \sum_{i=1}^{n}\mu_i^U\xi_i^U \\ & \quad + \sum_{i=1}^{n}\lambda_i^L\left(\left(\omega^T X_i + b\right) - y_i - \varepsilon - \xi_i^L\right) - \sum_{i=1}^{n}\mu_i^L\xi_i^L \\ & \quad \lambda^U,\lambda^L,\mu^U,\mu^L \ge 0, \quad i=1,2,\cdots,n \end{aligned} \tag{5}$$

$$\xrightarrow[\text{partial derivative}]{\text{regarding } \omega,b}$$

$$\begin{aligned} & \min_{\lambda^*,\lambda^*} \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\left(\lambda_i^U-\lambda_i^L\right)\left(\lambda_j^U-\lambda_j^L\right)X_i^T X_j \\ & \quad + \sum_{i=1}^{n}\left(\varepsilon-y_i\right)\lambda_i^U + \sum_{i=1}^{n}\left(\varepsilon+y_i\right)\lambda_i^L \\ & \sum_{i=1}^{n}\left(\lambda_i^U-\lambda_i^L\right)=0, \ 0\le\lambda_i^U,\lambda_i^L\le C, \quad i=1,2,\ldots,n \end{aligned}$$

If introducing the kernel trick, SVR will transform the input space X into higher dimensional space Z through nonlinear function $\psi(X)$. The kernel SVR model could be defined as the following:

$$\begin{aligned} & \min_{\lambda^U,\lambda^L} \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\left(\lambda_i^U-\lambda_i^L\right)\left(\lambda_j^U-\lambda_j^L\right)K(X_i,X_j) \\ & \quad + \sum_{i=1}^{n}\left(\varepsilon-y_i\right)\lambda_i^U + \sum_{i=1}^{n}\left(\varepsilon+y_i\right)\lambda_i^L \\ & \quad \sum_{i=1}^{n}\left(\lambda_i^U-\lambda_i^L\right)=0, \\ & \quad 0\le\lambda_i^U,\lambda_i^L\le C, \ i=1,2,\ldots,n, \end{aligned} \tag{6}$$

where the kernel function $K(X_i,X_j)=\psi^T(X_i)\ \psi(X_j)$.

Random forest: Random forest is a new machine learning algorithm with high flexibility, which has been widely used in various classification and regression fields. Random forest adopts the Bagging idea to aggregate a series of decision trees into a forest. To be more specific, random forest constructs multiple independent decision trees, each tree being considered as an evaluator. Then, after obtaining samples, decision tree divides the input space into disjoint areas by the way where the information gain or Gini coefficient decreases fastest. At last, the number of forecast is as same as the number of decision trees and average forecast is regarded as the final output. Random forest algorithm can not only effectively reduce the over-fitting situation, but also realize the parallel operation when there is a large amount of training data.

BP neural networks: BP neural networks have not only relatively simple network structure and high computational accuracy, but also efficient non-linear mapping ability. Through the activation function, the hidden layer nodes transmit the value of input layer after nonlinear transformation to the output layer. Then we can obtain the output through the calculation of another activation function. The loss function is defined as the difference between the output value and the actual value. Once the information propagates from the input layer to the output layer, the loss can be computed. BP neural network learns the weights w and biases b of all layers in a supervised way, which means the back propagation algorithm is used to optimize the weights w and biases b of all layers. The whole process is constantly iterated to make the loss achieve gradient descent.

3.3. Wavelet Neural Networks Based on Adaptive Genetic Algorithm

Wavelet neural networks (WNN) is an improved version of BP neural networks, which replaces hidden layer's Sigmoid activation function with basic wavelet function. WNN combines wavelet transform's excellent character—time precision in high frequency domain and frequency precision in low frequency domain with BP neural network's self-learning and self-adaptation advantages. Therefore, WNN has strong approximation and fault-tolerance ability.

With WNN being applied wider, its defects appear. For example, the number of hidden layer nodes is difficult to determine and initial values of parameters have a great impact on network performance. Genetic algorithm (GA) is a globally random search algorithm referring to natural evolutionary mechanism. It is suitable for dealing with complex nonlinear optimization problems that are difficult to be solved by traditional search algorithms. Therefore, the hybrid prediction model combining WNN and GA is presented to search a globally near-optimal combination of network pa- rameters, which can improves the accuracy of WNN. [22, 29, 34, 35] have attained preferable results in training neural networks with GA, but there are few studies on WNN with adaptive GA in shipping market.

3.3.1. Wavelet Transform

The essence of continuous wavelet transform is an integral transform between different parameter spaces. It is a conjugate function $\hat{\psi}(t)$ after different translations b and scales a that does inner product with the sequence $f(t)$ waiting to be analyzed [16].

$$CWT_{\psi}(a,b) = |a|^{-\frac{1}{2}} \int_{-\infty}^{+\infty} f(t)\hat{\psi}\left(\frac{t-b}{a}\right)dt, \quad a \neq 0;\ a,b \in \mathbb{R} \tag{7}$$

where $\psi(t)$ is a fundamental wavelet function and Equation (7) should meet the following restrictions:

$$\int_R |\psi(t)|^2\, dt < +\infty; \quad \int_R |f(t)|^2\, dt < +\infty$$

$$C_{\psi} = \int_R \frac{|\hat{\psi}(\omega)|^2}{|\omega|} d\omega < +\infty; \quad \hat{\psi}(0)\int_{-\infty}^{+\infty}\psi(t)dt < \infty$$

The formula of inverse wavelet transform is calculated as follows:

$$f(t) = \frac{1}{C_{\psi}}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\frac{1}{a^2}CWT_{\psi}(a,b)\psi\left(\frac{t-b}{a}\right)dadb. \tag{8}$$

In practice, discrete wavelet transform can usually get better results. Equation (8) is discretized as follows:

$$\hat{f}(t) = \sum_{i=1}^{K}\omega_i\psi\left(\frac{t-b_i}{a_i}\right), \tag{9}$$

where K represents the number of basis wavelet functions. The original time series can be fitted by linear superposition of weighted basis wavelet functions.

3.3.2. Wavelet Neural Networks

Wavelet transform and neural networks can be combined in two ways.

Loose type: The original data is processed by wavelet transform to extract the feature vectors which are targeted at the input of neural network, and then we carry out the whole process of neural network.

Compact type: We replaces hidden layer's Sigmoid activation function with basic wavelet function. The training mode and learning pattern of WNN still adopt the idea of BP neural network.

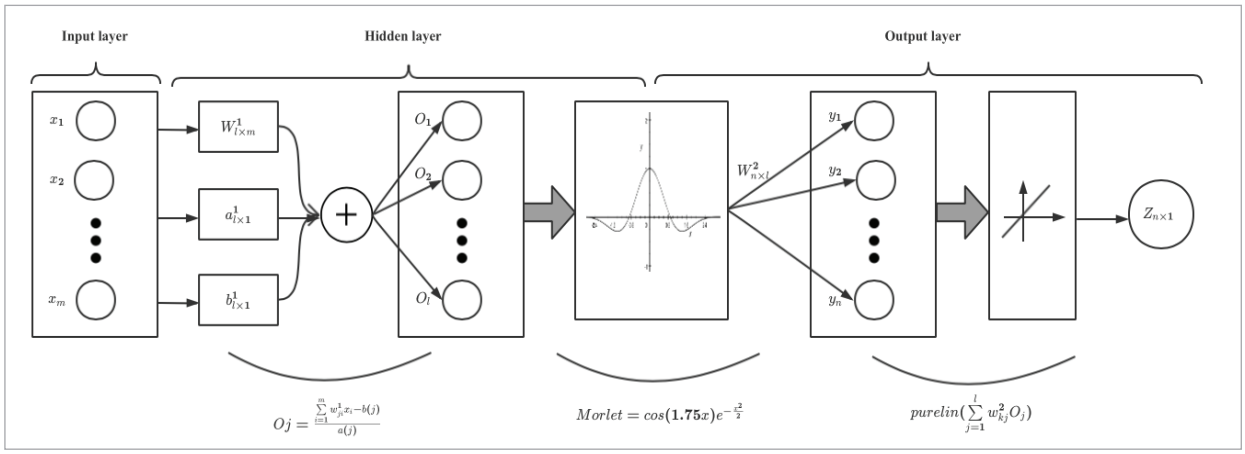
In this paper, the main research object is compact WNN. Hidden layer's activation function uses Morlet wavelet function $y = cos(1.75x)e^{-x^2/2}$ and output layer's activation function adopts purelin $y = x$. The flow chart is shown in Figure 1.

We assume that sample data includes input $X = [x_1, x_2, \ldots, x_m]'$ and expected output $Y = [y_1, y_2, \ldots, y_n]'$. Hidden layer neurons are $O = [O_1, O_2, ..., O_l]'$. The weight matrix between input layer and hidden layer is denoted as W^1 while the weight matrix between hidden layer and output layer is denoted as W^2.

$$W^1 = \left(\omega_{ji}^1\right) = \begin{bmatrix} \omega_{11}^1 & \omega_{12}^1 & \ldots & \omega_{1m}^1 \\ \omega_{21}^1 & \omega_{22}^1 & \ldots & \omega_{2m}^1 \\ \vdots & \vdots & & \vdots \\ \omega_{l1}^1 & \omega_{l2}^1 & \ldots & \omega_{lm}^1 \end{bmatrix}_{l\times m}$$

$$W^2 = \left(\omega_{kj}^2\right) = \begin{bmatrix} \omega_{11}^2 & \omega_{12}^2 & \ldots & \omega_{1l}^2 \\ \omega_{21}^2 & \omega_{22}^2 & \ldots & \omega_{2l}^2 \\ \vdots & \vdots & & \vdots \\ \omega_{n1}^2 & \omega_{n2}^2 & \ldots & \omega_{nl}^2 \end{bmatrix}_{n\times l} \tag{10}$$

Figure 1
The flow chart of compact WNN



Hidden layer's translations b and scales a are denoted, respectively, as $b=[b_1,b_2,...,b_l]'$ and $a=[a_1,a_2,...,a_l]'$. Then, hidden layer's input and output are expressed, respectively, as the following:

$$\begin{aligned} net_j &= (\sum_{i=1}^{m}\omega_{ji}^1 x_i - b_j)/a_j \\ O_j &= Morlet(net_j) = \cos(1.75net_j)e^{-net_j^2/2},\ j=1,2,\ldots,l \end{aligned} \tag{11}$$

Similarly, output layer's input and output are expressed, respectively, as the following:

$$\begin{aligned} net_k &= \sum_{j=1}^{l}\omega_{kj}^2 O_j \\ z_k &= purelin(net_k) = net_k,\ \ k=1,2,\ldots,n \end{aligned} \tag{12}$$

The error between the network output Z and the expected output Y is defined as Equation (13):

$$E=\frac{1}{2}\sum_{k=1}^{n}(y_k-z_k)^2. \tag{13}$$

The partial derivatives of the error with respect to the parameters ω_{kj}^2, ω_{ji}^1, b_j, a_j are calculated as follows:

$$\begin{cases} \dfrac{\partial E}{\partial \omega_{kj}^2} = \dfrac{\partial E}{\partial z_k}\dfrac{\partial z_k}{\partial \omega_{kj}^2} = -(y_k-z_k)O_j \\ \dfrac{\partial E}{\partial \omega_{ji}^1} = \displaystyle\sum_{k=1}^{n}\dfrac{\partial E}{\partial z_k}\dfrac{\partial z_k}{\partial O_j}\dfrac{\partial O_j}{\partial \omega_{ji}^1} \\ \qquad = -\displaystyle\sum_{k=1}^{n}(y_k-z_k)\omega_{kj}^2 Morlet'(net_j)\dfrac{x_i}{a_j} \end{cases} \tag{14}$$

$$\begin{cases} \dfrac{\partial E}{\partial b_j} = \displaystyle\sum_{k=1}^{n}\dfrac{\partial E}{\partial z_k}\dfrac{\partial z_k}{\partial O_j}\dfrac{\partial O_j}{\partial b_j} \\ \qquad = \displaystyle\sum_{k=1}^{n}(y_k-z_k)\omega_{kj}^2 Morlet'(net_j)\dfrac{1}{a_j} \\ \dfrac{\partial E}{\partial a_j} = \displaystyle\sum_{k=1}^{n}\dfrac{\partial E}{\partial z_k}\dfrac{\partial z_k}{\partial O_j}\dfrac{\partial O_j}{\partial a_j} \\ \qquad = \displaystyle\sum_{k=1}^{n}(y_k-z_k)\omega_{kj}^2 Morlet'(net_j)\left(\sum_{i=1}^{m}\omega_{ji}^1 x_i - b_j\right)\dfrac{1}{a_j^2} \end{cases}$$

The iterative formulae of weights, translations and scales are calculated as follows:

$$\begin{cases} \omega_{kj}^2(t+1) = \omega_{kj}^2(t) + \Delta\omega_{kj}^2 = \omega_{kj}^2(t) - \eta^1\dfrac{\partial E}{\partial \omega_{kj}^2} \\ w_{ji}^1(t+1) = \omega_{ji}^1(t) + \Delta\omega_{ji}^1 = \omega_{ji}^1(t) - \eta^1\dfrac{\partial E}{\partial \omega_{ji}^1} \\ b_j(t+1) = b_j(t) + \Delta b_j = b_j(t) - \eta^2\dfrac{\partial E}{\partial b_j} \\ a_j(t+1) = a_j(t) + \Delta a_j = a_j(t) - \eta^2\dfrac{\partial E}{\partial a_j} \end{cases} \tag{15}$$

where η^1 and η^2 represent the learning rate of weight parameters and wavelet parameters, respectively.

3.3.3. Adaptive Genetic Algorithm

Genetic algorithm cannot directly deal with the parameters of the problem space. First, the feasible solutions of the problem space need to be expressed

as chromosomes of the genetic space. Then the initial population is randomly generated. Each individual in the population represents a solution to the problem. The quality of the individual is measured by the fitness function. Through a series of genetic operations, such as selection, recombination, mutation and evolutionary reversal, the father generation produces offspring in which the excellent individuals with high fitness are more likely to be selected to form new population. In the process of iteration, the excellent individual's information is preserved and constantly exchanged to ensure the diversity of the population. Finally, the remaining individuals converge around the optimal solution and the optimal individual is selected as the solution of the problem. The process of WNN based on AGA is shown in Figure 2.

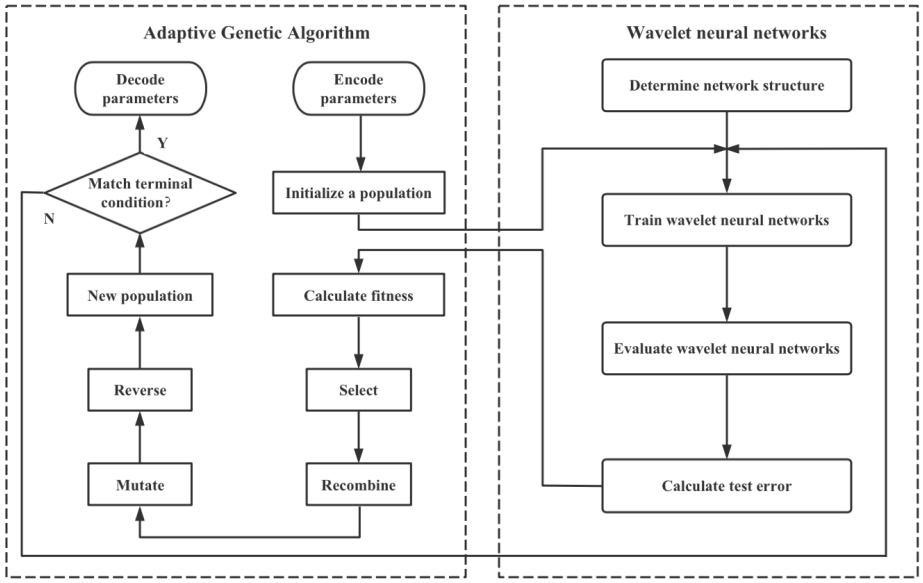
Encode parameters: Before conducting searches, the parameters that need to be encoded include (1) weights ω^1, ω^2 (2) translations b (3) scales a. Given that the number of input layer neurons and output layer neurons are m and n, respectively, the number of hidden layer neurons can be set to $(2m+1)$. The total of the encoded parameters equals $(2m^2+2mn+5m+n+2)$. Each of the parameters is coded with ten binary numbers.

Genetic operations: There are four genetic operations including selection, recombination, mutation and reversal. The purpose of selection is to filter elite individuals from the current population. The individuals with higher fitness scores $(1/MSE_{WNN})$ in the father generation have a higher probability to reproduce offspring. This process reflects "the fittest can survive" in the biological world. In this paper, roulette-wheel selection is applied to establish the offspring. The probability of an individual i being selected is defined as the following:

$$p_i = \frac{F_i}{\sum_{j=1}^{N} F_j}, \tag{16}$$

where F_i represents fitness score of the individual i and N is the capacity of parent population. Both single-point crossover and adaptive mutation are applied to diversify the population so as to find a globally near-optimal solution. Single-point crossover refers that randomly selected two parent members exchange or combine chromosomes at a random position to pass on excellent characteristics to their offspring. For each individual in the offspring, if a random number generated between 0 and 1 is less

Figure 2
The process of WNN based on AGA



than the mutation probability P_m, binary conversion is performed at its random position. In this study, P_m adaptively adjusts according to the heterogeneity of the population $(1/Chrom_\sigma)$, where $Chrom_\sigma$ represents the standard deviation of the parent's fitness scores. A small $Chrom_\sigma$ means high homogeneity in the population, so a high mutation probability is applied to diversify the chromosomes. In order to improve the local searching ability of AGA, continuous evolutionary reversal operations i.e. swapping numbers in two random positions, are introduced after selection, recombination and adaptive mutation operation. Only individuals with improved fitness are accepted. Finally, the inferior members of the parent population are replaced by the elite members of the offspring generation to form a new population. The whole AGA process is iterated repeatedly until the terminal condition is satisfied.

Generally speaking, AGA-WNN has four advantages compared with the baseline models,

1 Strong nonlinear approximation, self-learning and self-adaptation ability.
2 Permission to parameter estimation based on a large number of variables.
3 Excellent generalization ability.
4 Fast convergence and inaccessibility to fall into the local optimal trap.

4. Data

4.1. Data Description

All secondhand tanker price data is collected from the Clarkson Shipping Intelligence Network [10]. This paper aims to research the application effect of the hybrid prediction model — wavelet neural networks based on adaptive genetic algorithm in the tanker market. Attributing to the wide distribution of ship sizes of tankers, only the most representative tanker types are selected for empirical study. The predominant tankers selected carry major global crude oil trade, including Handysize, Panamax, Aframax, Suezmax and VLCC. Therefore, the method and its prediction results can effectively interpret the market situation of second-hand tankers. In this section, we give a concise description of each tanker type, followed by descriptive statistics of its datasets.

Handysize tankers refer to small-sized crude oil tankers with a deadweight tonnage ranging between 10,000 and 50,000 tons. It is because of its strong flexibility and unconstrained draft that handysize tankers can play an important role in many fields, for example offshore waters and offshore drilling platforms, forming a complement to large tankers. Its sales have increased gradually in recent years. In this paper, handysize tankers of 37K DWT are selected for empirical study.

Panamax tankers are subject to the navigation conditions of Panama Canal as the upper limits, in other words, they need to meet some restrictions about the ship's width, draft and so on. Generally, several main indicators are limited as follows: the tanker's total length should not be longer than 274.32 meters, the tanker's width should not be wider than 32.30 meters and the tanker's deadweight tonnage should range between 60,000 and 80,000 tons. The main routes include the Far East to Japan, the Far East to India and Singapore to Japan. In this paper, Panamax tankers of 73K DWT are selected for empirical study.

Aframax tankers refer to medium-sized crude oil tankers with a deadweight tonnage ranging between 80,000 and 120,000 tons. However, considering the operating costs, their actual deadweight tonnage generally vary from 70,000 to 110,000 tons, having an average cargo-carrying capacity of approximately 750,000 barrels [12]. Many non-OPEC oil exporting countries have port facilities that are difficult to accommodate VLCC or ULCC, so that there is a high demand for Aframax tankers, owing to their advantageous size. With the highest average freight index (AFRA) and the best economy, Aframax tankers are the ideal choice for short to medium-haul oil trades, which are also known as the "workhorse" of the world's tanker fleet. The main routes include the Middle East to the Far East, the Mediterranean region to the West Coast of the United States, the Caribbean to the West Coast of the United States, and the Caribbean to Europe. In this paper, Aframax tankers of 105K DWT are selected for empirical study.

Suezmax tankers are the largest vessels that can meet the traffic restrictions of Suez Canal, with a deadweight tonnage ranging between 120,000 and 200,000 tons. In 2009, the Suez Canal was expanded from 18 meters to 20.1 meters to allow a Suezmax tanker with up to 200,000 deadweight tonnage. The main routes

include the Persian Gulf to the Far East, the Caribbean to North America, West Africa to North America, West Africa to the Mediterranean region. In this paper, Suezmax tankers of 150K DWT are selected for empirical study.

Very Large Crude Carriers (VLCC) and Ultra Large Crude Carriers (ULCC) are the main type of long-distance crude oil transportation, compared with other types of tankers with higher performance and economies of scale. VLCC have a deadweight tonnage ranging between 200,000 and 320,000 tons and ULCC have a deadweight tonnage more than 320,000 tons. The main routes are deployed around the Persian Gulf, the Far East, North America, Northern Europe, Indian subcontinent and West Africa. In this paper, VLCC of 300K DWT are selected for empirical study.

The monthly data analyzed are based on 10-year-old and 15-year-old secondhand tanker prices for the five tanker types mentioned above. The longest sequence starts in July 1989 and the shortest sequence is from October 2012. All series end on April 2022. The numbers of data points vary from 115 to 394. Table 1 summarizes the detailed statistics. For the average price level, large tankers are higher than small ones and 10YO vessels are higher than 15YO ones. The standard deviation also follows a similar pattern. Most distributions are right-skewed and leptokurtic.

Jarque-Bera statistics can test whether a set of samples is normally distributed. Under the null hypothesis of normal distribution, the JB statistics calculated according to Equation (17) follows χ^2 distribution with 2 degrees of freedom.

$$JB = \frac{n}{6}\left[skewness^2 + \frac{(Kurtosis - 3)^2}{4} \right] \sim \chi^2(2) \qquad (17)$$

The result demonstrates that the price sequences for all tanker types and ages do not follow a normal distribution at significance level $\alpha = 0.05$.

4.2. Data Pre-processing

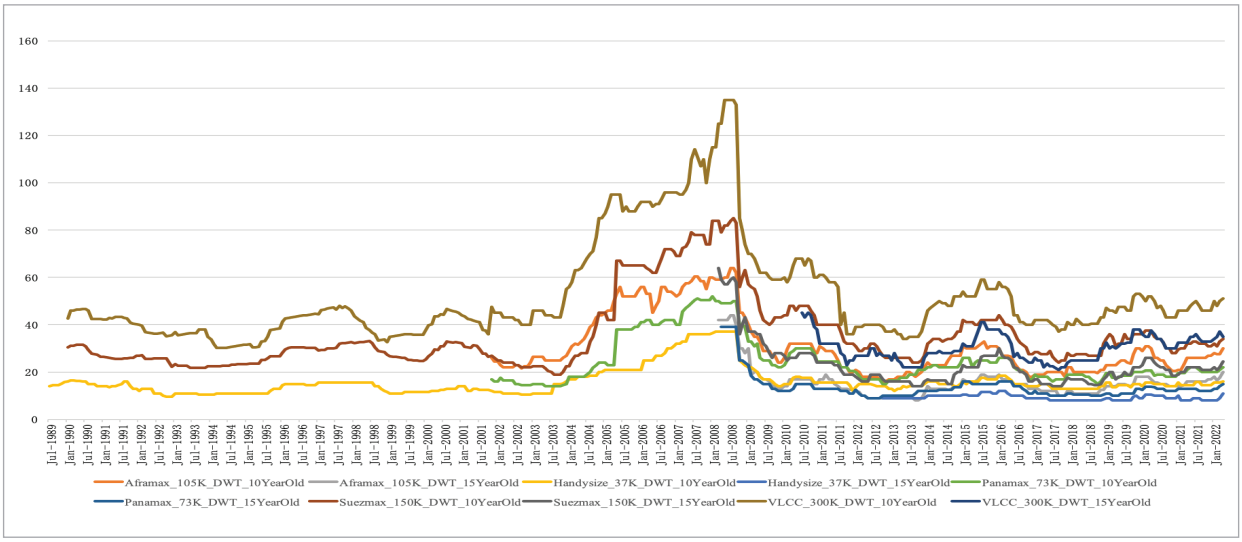
4.2.1. Order Determination

According to Figure 3, the overall pattern of secondhand tanker price series is quite obvious. Especially from 2005 to 2009, the data reflects an unprecedented fluctuation. The factor decomposition of the price series of 10YO secondhand Aframax is taken as an example, as is demonstrated in Figure 4, which is divided into four parts from top to bottom. The first part is the original observations; the second part is the estimated long-term trend which is increasing initially and then decreasing, cyclically, and also shows anomalies in the period 2005-2009; the third part is

Table 1
Descriptive statistics of the datasets

	Start	Stop	Length	Max	Min	Median	Mean	Std	Kurtosis	Skewness	JB	P-value
10YO's Handysize	Jul-1989	Apr-2022	394	37.00	9.50	14.75	15.91	5.80	5.36	2.33	811.0284	0.001
15YO's Handysize	Oct-2012	Apr-2022	115	12.00	8.00	9.00	9.23	1.09	-0.31	0.61	7.5939	0.028
10YO's Panamax	Nov-2001	Apr-2022	246	52.00	14.00	21.00	24.66	10.06	0.77	1.37	81.7490	0.001
15YO's Panamax	Apr-2008	Apr-2022	169	39.00	9.00	12.50	13.62	5.56	13.46	3.54	1542.6200	0.001
10YO's Aframax	Nov-2001	Apr-2022	246	64.00	16.00	26.50	30.95	12.80	0.12	1.17	55.8548	0.001
15YO's Aframax	Mar-2008	Apr-2022	170	44.00	8.00	14.50	15.64	6.65	8.75	2.78	721.7264	0.001
10YO's Suezmax	Jan-1990	Apr-2022	388	85.00	19.00	30.44	35.30	14.79	2.54	1.80	308.9113	0.001
15YO's Suezmax	Mar-2008	Apr-2022	170	64.00	14.00	21.00	22.85	9.43	7.10	2.49	506.0935	0.001
10YO's VLCC	Jan-1990	Apr-2022	388	135	30.23	44.30	51.90	21.20	3.45	1.94	426.3524	0.001
15YO's VLCC	Jul-2010	Apr-2022	142	45.00	21.00	30.00	30.09	5.47	-0.27	0.47	5.6020	0.0495

Figure 3
Time series chart of secondhand tanker prices for various ship types



the estimated seasonal fluctuation; the last part is the estimated random variation which fluctuates around zero. On the whole, Figure 4 is consistent with the information in the original sequence diagram. Given that the temporal patterns of all sequences are highly similar, they are non-stationary time series with seasonal variation and periodic trend, suitable for order determination by SARIMA model.

Figure 4
Decomposition of 10YO's Aframax

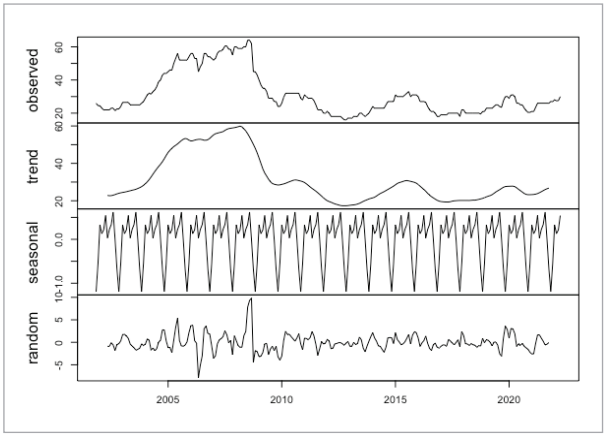


Table 2 shows the autocorrelation orders and underlying network structures of 10 time series, using the *forecast* package in R library. Then, machine learning methods can be applied to analyze modified datasets with standard forms of input independent variables and output dependent variables

4.2.2. Normalization

In the field of machine learning, normalization is a common basic operation of data pre-processing. A proper normalization can help neural network model produce more accurate results. Its advantages can be embodied in two aspects. On the one hand, in some network structures we set above, independent variables have multiple dimensions and different features may range in different scales. In this case, features with higher absolute level will play a more important role, which even affect the final output. In order to solve this problem, when the relative importance of each feature is unclear, normalization makes each input feature in the same magnitude or similar distribution, which is suitable for comparative evaluation. In the network training process, we can ensure that all features are treated equally (meaning we set the same learning rate, initial weight, and activation function). On the other hand, in this paper, the training speed of BP neural networks and wavelet neural networks is determined by the speed of gradient descent algorithm. After data normalization, the contour surface of the loss function's probability distribution is ap-

Table 2
Order determination and underlying structures of all time series

	Order $(p, d, q) \times (P, D, Q)_{\pi}$	SARIMA's formula	WNN's structure (Nodes: Input layer → Hidden layer → Output layer)
10YO's Handysize	(1,1,1)	$\nabla\phi_t = \frac{1+0.8026B}{1-0.8777B}\varepsilon_t$	$(i)\ \phi_t = f(\phi_{t-1}),\ 1\to 3\to 1$ $(ii)\ \phi_t = f(\phi_{t-1},\phi_{t-2}),\ 2\to 5\to 1$
15YO's Handysize	(0,1,0)	$\nabla\phi_t = \varepsilon_t$	$\phi_t = f(\phi_{t-1}),\ 1\to 3\to 1$
10YO's Panamax	(0,1,0)	$\nabla\phi_t = \varepsilon_t$	$\phi_t = f(\phi_{t-1}),\ 1\to 3\to 1$
15YO's Panamax	$(5,2,0)\times(2,0,0)_{12}$	$\nabla^2\phi_t = \frac{(1-0.12635B^{12}-0.00524B^{24})^{-1}}{(1+0.92641B+\ldots+0.13987B^{5})}\varepsilon_t$	$(i)\ \phi_t = f(\phi_{t-1},\phi_{t-2}),\ 2\to 5\to 1$ $(ii)\ \phi_t = f(\phi_{t-12},\phi_{t-24}),\ 2\to 5\to 1$ $(iii)\ \phi_t = f(\phi_{t-1},\llcorner,\phi_{t-7},\phi_{t-12},\phi_{t-24}),\ 9\to 19\to 1$
10YO's Aframax	(1,1,0)	$\nabla\phi_t = \frac{1}{1-0.1572B}\varepsilon_t$	$\phi_t = f(\phi_{t-1},\phi_{t-2}),\ 2\to 5\to 1$
15YO's Aframax	$(5,2,1)\times(2,0,0)_{12}$	$\nabla^2\phi_t = \frac{(1-0.3070B^{12}-0.0149B^{24})^{-1}}{(1-0.0647B-\ldots-0.0619B^{5})}(1+0.9876B)\varepsilon_t$	$(i)\ \phi_t = f(\phi_{t-1},\phi_{t-2}),\ 2\to 5\to 1$ $(ii)\ \phi_t = f(\phi_{t-12},\phi_{t-24}),\ 2\to 5\to 1$ $(iii)\ \phi_t = f(\phi_{t-1},\phi_{t-12},\phi_{t-24}),\ 3\to 7\to 1$ $(iv)\ \phi_t = f(\phi_{t-1},\llcorner,\phi_{t-7},\phi_{t-12},\phi_{t-24}),\ 9\to 19\to 1$
10YO's Suezmax	(0,1,0)	$\nabla\phi_t = \varepsilon_t$	$\phi_t = f(\phi_{t-1}),\ 1\to 3\to 1$
15YO's Suezmax	$(1,2,2)\times(1,0,1)_{12}$	$\nabla^2\phi_t = \frac{(1+0.2635B+0.6619B^{2})(1+0.3551B^{12})}{(1+0.6297B)(1-0.2448B^{12})}\varepsilon_t$	$(i)\ \phi_t = f(\phi_{t-1},\phi_{t-2}),\ 2\to 5\to 1$ $(ii)\ \phi_t = f(\phi_{t-1},\phi_{t-2},\phi_{t-3}),\ 3\to 7\to 1$ $(iii)\ \phi_t = f(\phi_{t-1},\phi_{t-2},\phi_{t-3},\phi_{t-12}),\ 4\to 9\to 1$
10YO's VLCC	(2,1,0)	$\nabla\phi_t = \frac{1}{1-0.0915B-0.1596B^{2}}\varepsilon_t$	$\phi_t = f(\phi_{t-1},\phi_{t-2},\phi_{t-3}),\ 3\to 7\to 1$
15YO's VLCC	(1,0,2)	$\nabla\phi_t = 31.5004 + \frac{1-0.1327B-0.2936B^{2}}{1-0.9215B}\varepsilon_t$	$(i)\ \phi_t = f(\phi_{t-1}),\ 1\to 3\to 1$ $(ii)\ \phi_t = f(\phi_{t-1},\phi_{t-2}),\ 2\to 5\to 1$

proximately circular. The gradient descent direction points toward the center of the circle, resulting in faster convergence and fewer iterations.

Hereby, the max-min normalization is put into practice as Equation (18):

$$x_{normalized} = \frac{x - x_{\min}}{x_{\max} - x_{\min}}, \tag{18}$$

where $x_{\max}$ and $x_{\min}$ represent the maximum and minimum of the original sequence x. $x_{normalized}$ denotes the normalized sequence, which ranges in [0, 1].

5. Empirical Results

5.1. Hyper-parameter Optimization

The sample data of the 10 time series (actually 19) is randomly divided into training set and test set. The optimal hyper-parameters make the validation set separated from the training set get the best performance, that is, the minimum error. Then the test set is retrained based on the optimal hyper-parameters. Hyper-parameters contained in traditional machine learning models are optimized through cross-vali-

dation. BP neural networks adopt Levenberg-Marquardt algorithm while wavelet neural networks use adaptive genetic algorithm to optimize hyper-parameters, iteratively searching for the approximately optimal parameter combination until the termination condition is satisfied. The fitting effects of all models are measured by MSE, MAE and MAPE.

SVR and random forest are simulated using the *Sklearn* package from Python library, where cross-validation is implemented through the *model_selection* module. $CV = 5$. BP neural networks are simulated by neural network toolbox in MATLAB. The values of training epochs, training goal and learning rate equal 2000, 10^{-6} and 0.01, respectively. For wavelet neural networks, the activation function is set to basis wavelet function in hidden layer as well as linear function *purelin* in output layer. The hyper- parameters of adaptive genetic algorithm are presupposed as follows: maximum iteration $MaxGen = 50$, population size $NIND = 150$, precision of variables $PRECI = 10$, generation gap $GGAP = 0.95$, crossover probability $P_x = 0.7$.

5.2. Comparison with Baseline Models

The performance of 7 models on 10 time series (19 structures) is compared based on two perspectives.

1 Error measures.

Three kinds of prediction error measures are calculated to judge the superiority and inferiority of these prediction models. Common prediction error indicators include mean square error (MSE), mean absolute error (MAE) and mean absolute percentage error (MAPE), which are defined as follows:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2, \tag{19}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|, \tag{20}$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right|, \tag{21}$$

where n is the number of samples in the test set. y_i and $\hat{y}_i$ represent raw data and the prediction, respectively.

MSE, MAE and MAPE of all test sets are demonstrated in Tables 3-5, respectively. The numbers in bold indicate that the corresponding model performs best on the corresponding time series. For sequences with multiple structures, we select the better one and mark it with a checkmark ("√"). As a result, 10 structures corresponding to 10 sequences still remain.

Table 3
MSE compared on all structures

		Seasonal Naive	SARIMA	Holt-Winters	SVR	Random forest	BP neural networks	AGA-WNN
10YO's Handysize	structure1 (√)	10.0434	0.4979	0.4192	2.4039	**0.3256**	15.7060	0.3349
	structure2 (×)	21.8814	1.6327	1.9767	2.0644	**0.4601**	24.0407	0.8800
15YO's Handysize		1.2083	114.1502	0.1626	0.1662	0.1343	2.3390	**0.1235**
10YO's Panamax		43.5490	3.1156	3.2066	8.9880	2.6004	71.5110	**2.0859**
15YO's Panamax structure2 (×)	structure1 (√)	27.1827	0.9340	1.6208	26.7814	2.5547	8.4017	**0.5563**
	structure2 (×)	3.1591	45.5823	9.9325	1.1812	0.9386	22.7605	**0.6047**
	structure3 (×)	5.3068	40.5541	24.6567	1.2212	**0.5996**	9.6743	0.9577
10YO's Aframax		156.2703	12.1383	12.9038	**2.6719**	3.8613	62.3163	5.0925
15YO's Aframax structure2 (×)	structure1 (×)	22.3750	7.0320	3.7788	1.2925	3.2285	45.3370	**0.8221**
	structure2 (×)	7.4659	48.4879	18.9533	3.0849	2.8113	53.7523	**1.5444**
	structure3 (√)	9.2045	1.0825	1.2391	1.2761	1.0943	22.2193	**0.8341**
	structure4 (×)	10.4545	**0.4195**	2.0019	1.6585	0.9329	12.1958	5.3153

		Seasonal Naive	SARIMA	Holt-Winters	SVR	Random forest	BP neural networks	AGA-WNN
10YO's Suezmax		59.4843	5.1193	5.5852	2.7341	3.4265	100.7711	**2.1241**
15YO's Suezmax structure2 (×)	structure1 (×)	52.5385	8.2474	9.7118	26.0564	7.0189	67.8454	**4.4440**
	structure2 (×)	69.5481	23.1119	**5.2875**	69.5337	17.3315	429.2047	7.7149
	structure3 (√)	37.5104	34.0081	37.9286	2.6998	2.4625	42.8727	**2.3490**
10YO's VLCC		177.0710	**3.3135**	3.8989	60.5313	10.3000	262.6474	4.4867
15YO's VLCC structure1 (×)	structure1 (×)	48.1818	3.0838	2.0531	4.8042	3.8642	57.3450	**1.5329**
	structure2 (√)	33.9524	1.4360	3.7300	2.6101	3.6991	37.3826	**1.0255**
Standard deviation on the selected structures (σ)		58.1720	33.6080	10.8640	18.0808	2.7263	72.7867	1.6273

Table 4
MAE compared on all structures

		Seasonal Naive	SARIMA	Holt -Winters	SVR	Random forest	BP neural networks	AGA-WNN
10YO's Handysize	structure1 (√)	2.2839	0.4644	0.4968	0.6317	**0.3412**	17.9148	0.4142
	structure2 (×)	2.5754	0.8709	0.9817	0.6746	**0.4631**	24.8626	0.8189
15YO's Handysize		0.8056	7.4406	**0.1711**	0.2281	0.1917	3.9216	0.2635
10YO's Panamax		4.7770	1.2495	**0.0568**	1.5885	1.1508	33.4315	1.1397
15YO's Panamax structure2 (×)	structure1 (√)	2.5962	0.7049	*0.9031*	1.8416	0.8762	10.4205	**0.5931**
	structure2 (×)	1.3636	4.0366	*2.4107*	0.8644	0.7823	18.6808	**0.6528**
	structure3 (×)	1.8864	4.2262	*3.7039*	0.8756	**0.5208**	9.6922	0.8220
10YO's Aframax		9.0270	2.7064	*2.7393*	**1.2273**	1.4413	35.2682	1.7478
15YO's Aframax structure2 (×)	structure1 (×)	3.5962	1.7993	*1.6427*	0.8149	0.8927	22.4614	**0.7466**
	structure2 (×)	2.2045	5.0490	*3.6010*	1.3282	1.3036	28.3095	**1.0059**
	structure3 (√)	2.2273	0.6848	*0.9075*	0.8009	0.7316	15.6608	**0.6068**
	structure4 (×)	2.5455	**0.4719**	*0.8402*	0.8638	0.8203	12.7344	1.8327
10YO's Suezmax		5.3828	1.6643	*1.8155*	1.1877	1.3644	55.2468	**1.1067**
15YO's Suezmax structure2 (×)	structure1 (×)	5.0000	2.2703	*2.4213*	1.9161	**1.5509**	29.4822	1.7770
	structure2 (×)	5.4038	2.8861	*1.7637*	3.2871	**1.7097**	43.0683	1.9766
	structure3 (√)	4.5208	4.4267	*4.6894*	1.3091	**1.1103**	23.4216	1.1465
10YO's VLCC		7.7695	**1.0596**	1.1145	4.1012	1.7070	79.7420	1.6255
15YO's VLCC structure1 (×)	structure1 (×)	5.2273	1.2851	**0.8746**	1.3355	1.4260	25.3621	0.9491
	structure2 (√)	4.9524	0.9211	1.5232	1.2584	1.5002	20.0925	**0.7817**
Standard deviation on the selected structures ( )		2.4403	2.1048	1.3223	0.9975	0.4764	21.6918	0.4715

Table 5
MAPE compared on all structures

		Seasonal Naive	SARIMA	Holt-Winters	SVR	Random forest	BP neural networks	AGA-WNN
10YO's Handysize	structure1 (√)	0.1408	0.0303	0.0330	0.0300	**0.0207**	1.1853	0.0284
	structure2 (×)	0.1552	0.0567	0.0635	0.0362	**0.0304**	1.6571	0.0562
15YO's Handysize		0.0818	0.8046	**0.0177**	0.0232	0.0199	0.4183	0.0293
10YO's Panamax		0.1807	0.0569	0.0568	0.0565	**0.0497**	1.3921	0.0518
15YO's Panamax structure2 (×)	structure1 (√)	0.2012	0.0557	0.0718	0.0705	0.0498	0.8306	**0.0446**
	structure2 (×)	0.1199	0.3408	0.2121	0.0708	0.0653	1.5513	**0.0570**
	structure3 (×)	0.1662	0.3494	0.3167	0.0713	**0.0441**	0.7761	0.0681
10YO's Aframax		0.3252	0.0968	0.0966	**0.0458**	0.0530	1.2949	0.0636
15YO's Aframax structure2 (×)	structure1 (×)	0.2767	0.1302	0.1283	0.0662	0.0643	1.6745	**0.0545**
	structure2 (×)	0.1773	0.4050	0.3213	0.1088	0.1142	2.3808	**0.0860**
	structure3 (√)	0.1795	0.0499	0.0670	0.0632	0.0530	1.1377	**0.0423**
	structure4 (×)	0.2033	**0.0357**	0.0604	0.0711	0.0650	0.9453	0.1483
10YO's Suezmax		0.1547	0.0514	0.0566	0.0392	0.0417	1.7389	**0.0366**
15YO's Suezmax structure2 (×)	structure1 (×)	0.2293	0.1010	0.1126	0.0593	**0.0578**	1.1666	0.0856
	structure2 (×)	0.2619	0.1093	0.0848	0.0855	**0.0609**	1.5664	0.0778
	structure3 (√)	0.2448	0.2395	0.2555	0.0705	**0.0593**	1.2916	0.0624
10YO's VLCC		0.1546	**0.0197**	0.0211	0.0682	0.0307	1.5036	0.0340
15YO's VLCC structure1 (×)	structure1 (×)	0.1951	0.0404	**0.0298**	0.0412	0.0441	0.8072	0.0301
	structure2 (√)	0.1657	0.0299	0.0522	0.0414	0.0468	0.6191	**0.0242**
Standard deviation on the selected structures (σ)		0.0619	0.2285	0.0650	0.0165	0.0143	0.3862	0.0132

The prediction model — WNN based on adaptive genetic algorithm (AGA-WNN) is superior to other prediction models on most datasets, and even improve significantly on certain time series. Taking the 15-year-old Suezmax secondhand prices (structure 3) as an example, AGA-WNN's accuracy is up to about 93 percent higher than that of SARIMA while about 13 percent higher than that of SVR regarding MSE. For time series where AGA-WNN does not rank first, the performance is still competitive and remains at the forefront of rankings. For instance, SVR reaches the minimum MAPE for the 10-year-old Aframax secondhand price, but AGA-WNN ranks top 3 with a narrow gap, only 1.78%. These two examples are visualized in Figures 5 and 7. Generally, AGA-WNN never worst perform, compared to the baseline models. It is reasonable to adopt AGA-WNN as the primary prediction model for the secondhand tanker market.

Figure 5

15YO's Suezmax (structure 3): Comparison results on test set

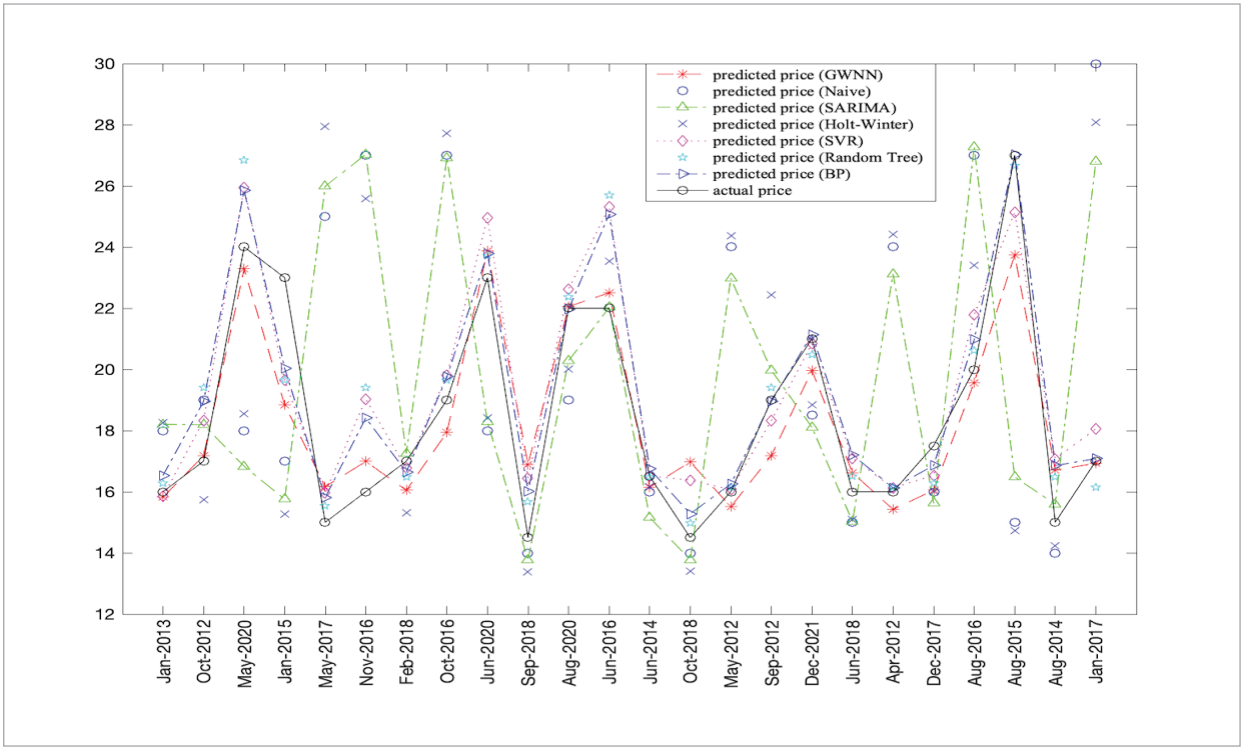


Figure 6

15YO's Suezmax (structure 3): Evolutionary process

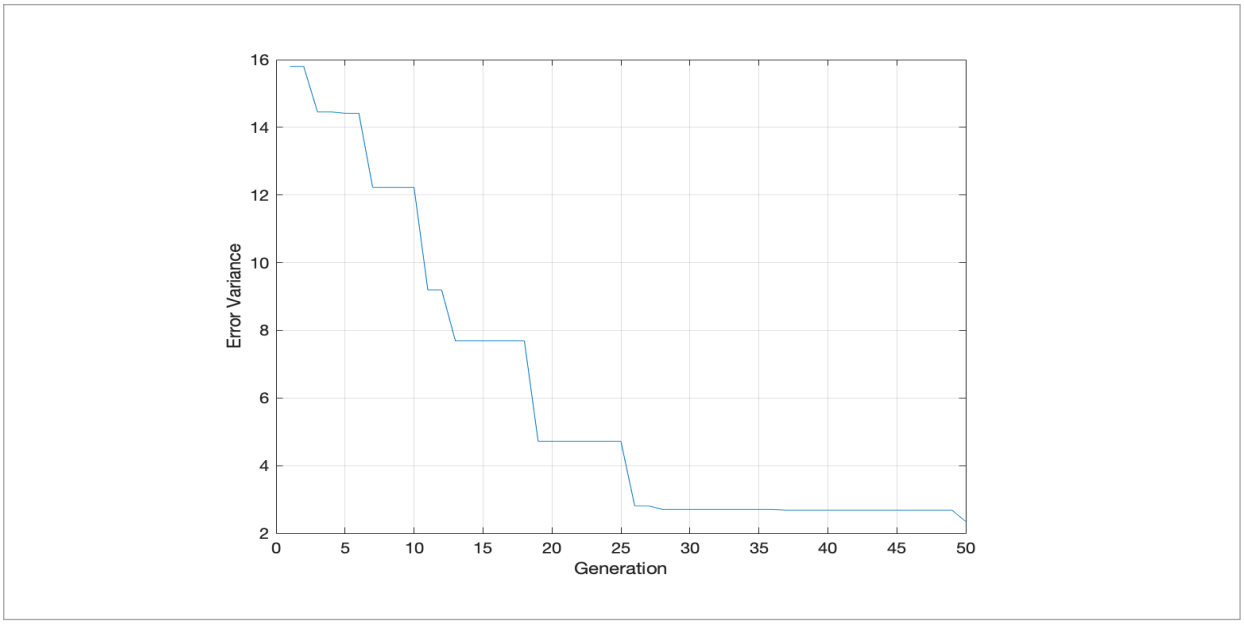


Figure 7

10YO's Aframax: Comparison results on test set

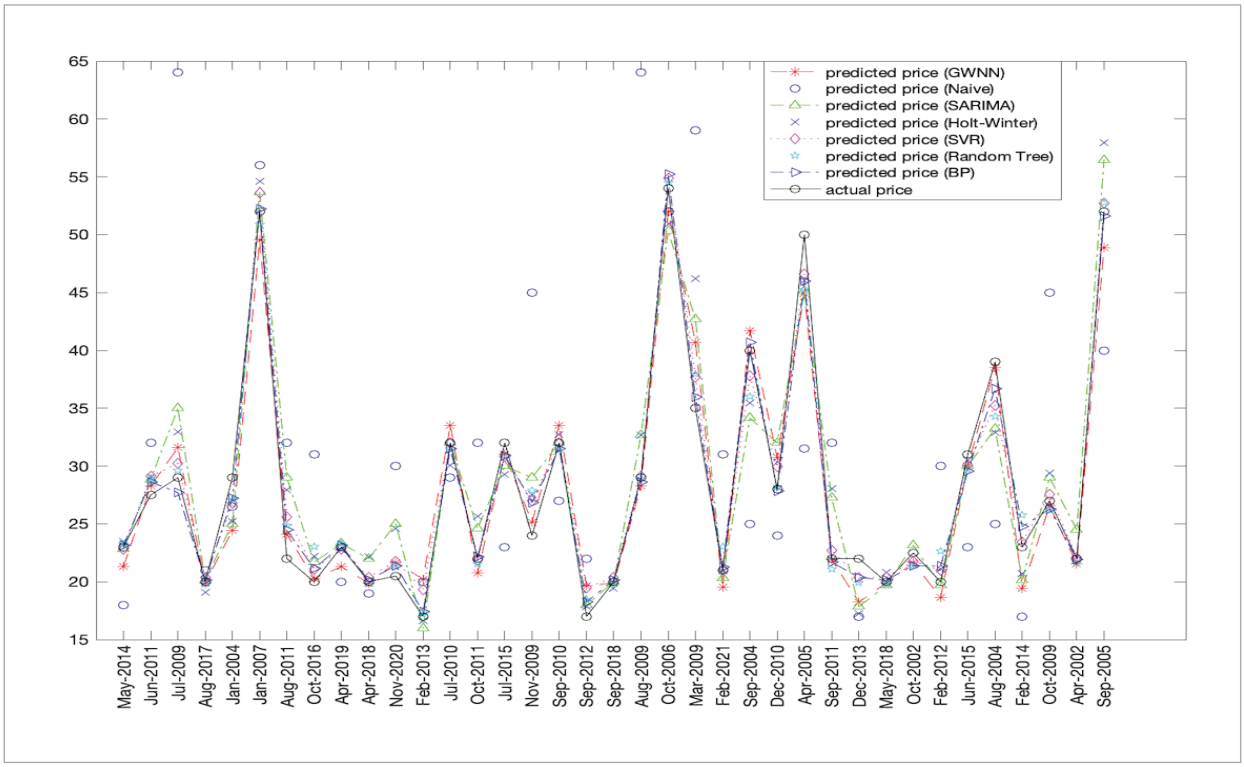
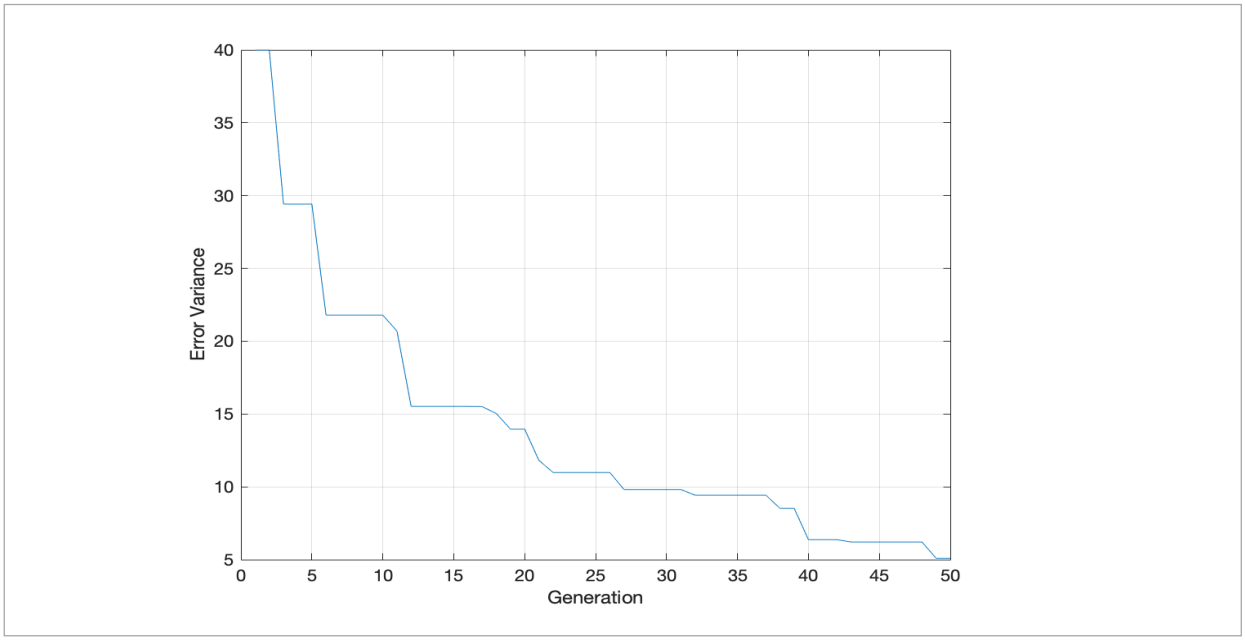


Figure 8

10YO's Aframax: Evolutionary process



2 Statistical tests.

The generalization ability of the models can be easily judged by comparing the error indicators directly. Meanwhile, in order to make the comparison more convincing, we can conduct statistical tests on the remaining time series structures. In this paper, two kinds of statistical tests, *Friedman* test and *Nemenyi* post-hoc test, are utilized to further confirm the significant differences among these models.

The test statistics of the *Friedman* test is shown in Equation (23):

$$\tau_{\chi^2} = \frac{k-1}{k}\frac{12N}{k^2-1}\sum_{i=1}^{k}\left(\gamma_i - \frac{k+1}{2}\right)^2 = \frac{12N}{k(k+1)}\left(\sum_{i=1}^{k}\gamma_i^2 - \frac{k(k+1)^2}{4}\right), \quad (22)$$

$$\tau_F = \frac{(N-1)\tau_{\chi^2}}{N(k-1)-\tau_{\chi^2}} \sim F(k-1,(k-1)(N-1)), \quad (23)$$

where N and k denote the number of datasets and algorithms, respectively. γ_i is the average rank of the algorithm i.

$$\gamma_i \sim N\left((k+1)/2,(k^2-1)/12\right).$$

When both N and k are large, τ_{χ^2} obeys the χ^2 distribution with (k – 1) degrees of freedom. However, τ_F is more frequently used.

As calculated on our data, MSE's $\tau_F = 19.6364$, MAE's $\tau_F = 16.9794$, MAPE's $\tau_F = 22.5789$, which are all greater than $F_{0.05}(6,\ 54) = 2.2720$. Thus we have adequate reasons to reject the null hypothesis that all algorithms have same performance, at the significance level $\alpha = 0.05$. The above diagrams can also prove that machine learning methods generally prevail over traditional models. However, we cannot directly determine whether the slight differences among AGA-WNN, SVR and random forest are significant based on three error metrics or *Friedman* test. Therefore, the *Nemenyi* test is introduced to further distinguish random two algorithms, providing a critical value (CD) of the difference between the average ranks.

$$CD = q_\alpha\sqrt{\frac{k(k+1)}{6N}}. \quad (24)$$

If the difference between the average ranks of two algorithms exceeds

$$CD = q_{0.05}\sqrt{\frac{7\times(7+1)}{6\times10}} = 2.949\times0.966 = 2.849,$$

the null hypothesis that two algorithms have same performance will be rejected at the corresponding significance level $\alpha = 0.05$. The average ranks based on the three error metrics are shown in Table 6, visualized in Figures 9-11. From top to bottom, the models are ranked from worst to best in terms of performance.

The results of the *Nemenyi* test show that although AGA-WNN and random forest are the top two models, they are not significantly different from SVR, SARIMA and Holt-Winter. The NFL Theorem (No Free Lunch Theorem) can explain that the expected performance of all algorithms is same, in other words, one algorithm will not perform best in all problems. Therefore, for several algorithms with similar accuracy on certain datasets, error variance becomes one of the most important factors. The error variance of AGA-WNN is smaller than that of other models. As a result, AGA-WNN is recommended to predict secondhand tanker price because of its robustness.

Table 6

Average ranks on remaining structures

Error metrics	Seasonal Naive	SARIMA	Holt -Winters	SVR	Random forest	BP neural networks	AGA-WNN
MSE	6.0	3.3	4.0	3.9	2.7	6.6	1.5
MAE	5.8	3.3	3.8	3.6	2.6	6.9	2.0
MAPE	5.8	3.8	4.2	3.2	1.9	6.9	2.2

Figure 9
Nemenyi test diagram based on MSE

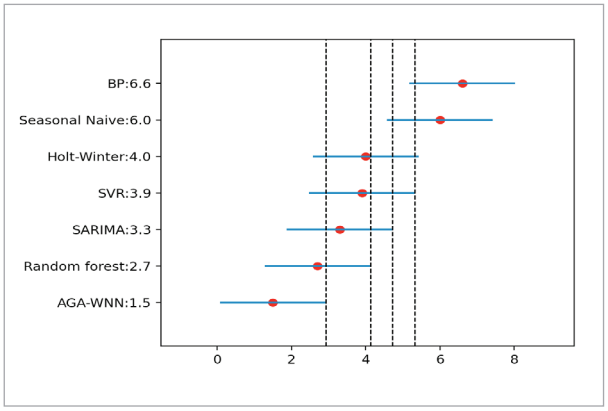


Figure 10
Nemenyi test diagram based on MAE

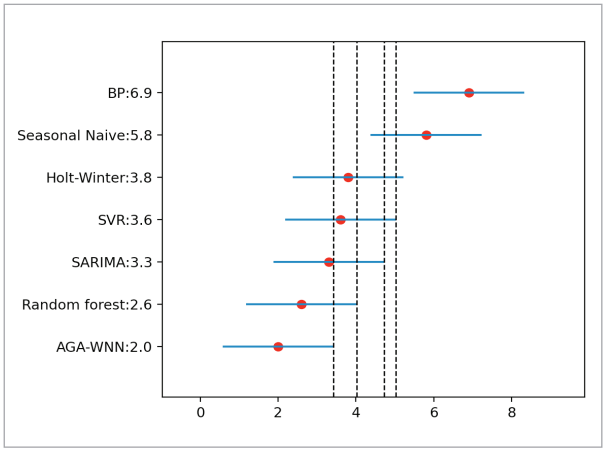
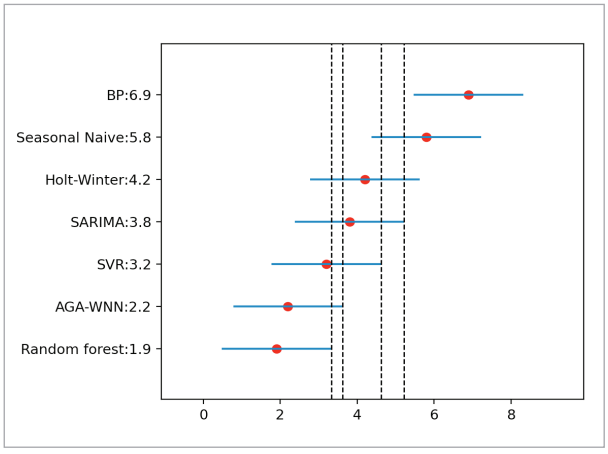


Figure 11
Nemenyi test diagram based on MAPE



6. Discussion and Conclusion

The secondhand ship market offers shipowners and investors the chance to buy and sell ships directly, making it easier for them to enter or exit the freight market. Therefore, the key lies in the time of investment. Asset transactions in which buy low and sell high can generate considerable profits. High freight rate tend to be accompanied by high ship value. While this is a piece of bad news for new investors in terms of raising costs, it provides an opportunity for shipowners to make money, whether operating or selling their ships. It is essential to make accurate predictions driven by data.

In this paper, a hybrid prediction model composed of wavelet neural networks and adaptive genetic algorithm is proposed to forcast the secondhand tanker price for different ship types and ship ages. A series of time series datasets include 10 year-old and 15 year-old Handysize, Panamax, Aframax, Suezmax and VLCC/ULCC. The proposed hybrid model is simulated and compared with common machine learning algorithms and traditional prediction models. Three kinds of prediction error metrics are calculated, and two kinds of post-hoc statistical tests based on the error are performed, in order to judge the overall performance of AGA-WNN. The results demonstrate that AGA-WNN is the best model compared with the baseline models, with comparative superiority in accuracy and robustness. Therefore, AGA-WNN can be considered as an applicable data-driven tool to help relevant stakeholders in the shipping market monitor market trends in time, make reasonable management or decisions, and avoid unnecessary loss caused by subjective judgment.

Although the hybrid prediction model has achieved remarkable success in multiple time series, it still has some limitations. Firstly, this paper mainly studies the prediction of nonlinear models, which have higher time complexity and slower running speed, compared with linear models. It is necessary and challenging to develop new parameter efficiently iterative algorithms in neural networks. Secondly, in addition to the five types of tankers investigated in the empirical study, we will also study the performance of AGA-WNN on other types of ships in the future, for example, dry bulk and container market. Furthermore, in terms of data frequency, this paper researches the

low-frequency monthly data. In regards to variable dimension, only the information of time series itself is extracted for prediction. In the future, we can aim at the high-frequency data, incorporating more exogenous variables that affect decision-making, for example, newbuilding ship price, time charter rates, scrap value and so on, to develop a predictive model with strong robustness. Last but not least, this article only adopts AGA to solve WNN's shortcomings regarding slow convergence and accessibility to fall into the local optimal trap. We can also combine WNN and other intelligent search algorithms, such as particle swarm optimization algorithm [7].

Acknowledgement

Data sharing agreement

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

Conflict of Interest

The authors have no conflicts of interest to declare.

References

1. Batchelor, R., Alizadeh, A., Visvikis, I. Forecasting Spot and Forward Prices in the International Freight Market. International Journal of Forecasting, 2007, 23(1), 101-114. https://doi.org/10.1016/j.ijforecast.2006.07.004
2. Beenstock, M. A Theory of Ship Prices. Maritime Policy and Management, 1985, 12(3), 215-225. https://doi.org/10.1080/03088838500000028
3. Beenstock, M., Vergottis, A. An Econometric Model of the World Market for Dry Cargo Freight and Shipping. Applied Economics, 1989, 21(3), 339-356. https://doi.org/10.1080/758522551
4. Beenstock, M., Vergottis, A. An Econometric Model of the World Tanker Market. Journal of Transport Economics and Policy, 1989, 263-280.
5. Cepowski, T., Chorab, P. The Use of Artificial Neural Networks to Determine the Engine Power and Fuel Consumption of Modern Bulk Carriers, Tankers and Container Ships. Energies, 2021, 14(16), 4827. https://doi.org/10.3390/en14164827
6. Charemza, W., Gronicki, M. An Econometric Model of World Shipping and Shipbuilding. Maritime Policy & Management, 1981, 8(1), 21-30. https://doi.org/10.1080/03088838100000019
7. Chen, Q., Song, Y., Zhao, J. Short-term Traffic Flow Prediction Based on Improved Wavelet Neural Network. Neural Computing and Applications, 2021, 33(14), 8181-8190. https://doi.org/10.1007/s00521-020-04932-5
8. Chen, S., Meersman, H., Voorde, E. V. D. Forecasting Spot Rates at Main Routes in the Dry Bulk Market. Maritime Economics & Logistics, 2012, 14(4), 498-537. https://doi.org/10.1057/mel.2012.18
9. Chou, H.-C., Chen, D.-H. The Use of Technical Analysis in Sale-and-purchase Transactions of Secondhand Ships. Maritime Economics & Logistics, 2019, 21(2), 223-240. https://doi.org/10.1057/s41278-017-0096-2
10. Clarkson. (n.d.). Shipping Intelligence Network. http://www.clarksons.net/sin2022/. (accessed 28 April 2022)
11. Clarkson.Oil and Tanker Trades Outlook, 2022.
12. Duru, O., Clott, C., Mileski, J. P. Us Tanker Transport: Current Structure and Economic Analysis. Research in Transportation Business & Management, 2017, 25, 39-50. https://doi.org/10.1016/j.rtbm.2017.04.001
13. Duru, O., Gulay, E., Girgin, S. C. Predictability of Second-hand Bulk Carriers with a Novel Hybrid Algorithm. The Asian Journal of Shipping and Logistics, 2021, 37(4), 291-300. https://doi.org/10.1016/j.ajsl.2021.07.002
14. Duru, O., Yoshida, S. Judgmental Forecasting in the Dry Bulk Shipping Business: Statistical vs. Judgmental Approach. The Asian Journal of Shipping and Logistics, 2009, 25(2), 189-217. https://doi.org/10.1016/S2092-5212(09)80002-3
15. Eslami, P., Jung, K., Lee, D., Tjolleng, A. Predicting Tanker Freight Rates Using Parsimonious Variables and a Hybrid Artificial Neural Network With an Adaptive Genetic Algorithm. Maritime Economics & Logistics, 2017, 19(3), 538-550. https://doi.org/10.1057/mel.2016.1
16. Esquível, M. L., Krasii, N. P. A Wavelet-based Neural Network Scheme for Supervised and Unsupervised Learning. Neural Computing and Applications, 2021, 33(20), 13433-13448. https://doi.org/10.1007/s00521-021-05968-x

17. Gao, R., Duru, O. Parsimonious Fuzzy Time Series Modelling. Expert Systems with Applications, 2020, 156, 113447. https://doi.org/10.1016/j.eswa.2020.113447

18. Gao, R., Liu, J., Du, L., Yuen, K. F. Shipping Market Forecasting by Forecast Combination Mechanism. Maritime Policy & Management, 2021, 1-16. https://doi.org/10.1080/03088839.2021.1945698

19. Gao, R., Liu, J., Zhou, Q., Duru, O., Yuen, K. F. Newbuilding Ship Price Forecasting by Parsimonious Intelligent Model Search Engine. Expert Systems with Applications, 2022, 117119. https://doi.org/10.1016/j.eswa.2022.117119

20. Geomelos, N., Xideas, E. Forecasting Spot Prices in Bulk Shipping Using Multivariate and Univariate Models. Cogent Economics & Finance, 2014, 2(1), 923701. https://doi.org/10.1080/23322039.2014.932701

21. Gkerekos, C., Lazakis, I., Theotokatos, G. Machine Learning Models for Predicting Ship Main Engine Fuel Oil Consumption: A Comparative Study. Ocean Engineering, 2019, 188, 106282. https://doi.org/10.1016/j.oceaneng.2019.106282

22. Han, J.-X., Ma, M.-Y., Wang, K. Product Modeling Design Based on Genetic Algorithm and BP Neural Network. Neural Computing and Applications, 2021, 33(9), 4111-4117. https://doi.org/10.1007/s00521-020-05604-0

23. Hawdon, D. Tanker Freight Rates in the Short and Long Run. Applied Economics, 1978, 10(3), 203-218. https://doi.org/10.1080/758527274

24. Kavussanos, M. G. Price Risk Modelling of Different Size Vessels in the Tanker Industry Using Auto-regressive Conditional Heterskedastic (Arch) Models. Logistics and Transportation Review, 1996, 32(2), 161.

25. Koopmans, T. C. Tanker Freight Rates and Tank Ship Building, 1939. Bohn.

26. Kushwah, S. V., Negi, P., Sharma, A. The Random Character of Stock Market Prices: A Study of Indian Stock Exchange. Integral Review: A Journal of Management, 2013, 6(1).

27. Leonov, Y., Nikolov, V. A Wavelet and Neural Network Model for the Prediction of Dry Bulk Shipping Indices. Maritime Economics & Logistics, 2012, 14(3), 319-333. https://doi.org/10.1057/mel.2012.10

28. Li, J., Parsons, M. G. Forecasting Tanker Freight Rate Using Neural Networks. Maritime Policy & Management, 1997, 24(1), 9-30. https://doi.org/10.1080/03088839700000053

29. Liu, A., Zhang, Y., Zhao, H., Wang, S., Sun, D. Neural Network Control System of Cooperative Robot Based on Genetic Algorithms. Neural Computing and Applications, 2021, 33(14), 8217-8226. https://doi.org/10.1007/s00521-020-04952-1

30. Lyridis, D., Zacharioudakis, P., Mitrou, P., Mylonas, A. Forecasting Tanker Market Using Artificial Neural Networks. Maritime Economics & Logistics, 2004, 6(2), 93-108. https://doi.org/10.1057/palgrave.mel.9100097

31. Munim, Z. H., Schramm, H.-J. Forecasting Container Freight Rates for Major Trade Routes: A Comparison of Artificial Neural Networks and Conventional Models. Maritime Economics & Logistics, 2021, 23(2), 310-327. https://doi.org/10.1057/s41278-020-00156-5

32. Şahin, B., Gürgen, S., Ünver, B., Altin, I. Forecasting the Baltic Dry Index by Using an Artificial Neural Network Approach. Turkish Journal of Electrical Engineering & Computer Sciences, 2018, 26(3), 1673-1684. https://doi.org/10.3906/elk-1706-155

33. Santos, A. A. P., Junkes, L. N., Jr, F. C. M. P. Forecasting Period Charter Rates of VLCC Tankers Through Neural Networks: A Comparison of Alternative Approaches. Maritime Economics & Logistics, 2014, 16(1), 72-91. https://doi.org/10.1057/mel.2013.20

34. Serrano, W. Genetic and Deep Learning Clusters Based on Neural Networks for Management Decision Structures. Neural Computing and Applications, 2020, 32(9), 4187-4211. https://doi.org/10.1007/s00521-019-04231-8

35. Solans, V., Rochman, D., Brazell, C., Vasiliev, A., Ferroukhi, H., Pautz, A. Optimisation of Used Nu-clear Fuel Canister Loading Using a Neural Network and Genetic Algorithm. Neural Computing and Ap-plications, 2021, 33(23), 16627-16639. https://doi.org/10.1007/s00521-021-06258-2

36. Syriopoulos, T., Tsatsaronis, M., Karamanos, I. Sup-port Vector Machine Algorithms: An Application to Ship Price Forecasting. Computational Economics, 2021, 57(1), 55-87. https://doi.org/10.1007/s10614-020-10032-2

37. Tsioumas, V., Papadimitriou, S., Smirlis, Y., Zahran, S. Z. A Novel Approach to Forecasting the Bulk Freight Market. The Asian Journal of Shipping and Logistics, 2017, 33(1), 33-41.https://doi.org/10.1016/j.ajsl.2017.03.005

38. Tsolakis, S., Cridland, C., Haralambides, H. Econo-metric Modelling of Second-hand Ship Prices. Mari-time Economics & Logistics, 2003, 5(4), 347-377. https://doi.org/10.1057/palgrave.mel.9100086

39. Vorkapić, A., Radonja, R., Martinčić-Ipšić, S. Predicting Seagoing Ship Energy Efficiency from the Operational Data. Sensors, 2021, 21(8), 2832. https://doi.org/10.3390/s21082832

40. Yang, Z., Mehmed, E. E. Artificial Neural Networks in Freight Rate Forecasting. Maritime Economics & Logistics, 2019, 21(3), 390-414. https://doi.org/10.1057/s41278-019-00121-x

41. Yu, H., Fang, Z., Lu, F., Murray, A. T., Zhang, H., Peng, P., Mei, Q., Chen, J. Impact of Oil Price Fluctu-ations on Tanker Maritime Network Structure and Traffic Flow Changes. Applied Energy, 2019, 237, 390-403. https://doi.org/10.1016/j.apenergy.2019.01.011

42. Zhang, Y. Investigating Dependencies Among Oil Price and Tanker Market Variables by Copula-based Multivariate Models. Energy, 2018, 161, 435-446. https://doi.org/10.1016/j.energy.2018.07.165