Research Article

Forecasting Dry Bulk Freight Index with Improved SVM

Qianqian Han,[1] Bo Yan,[2] Guobao Ning,[3] and B. Yu[4]

[1] *School of Accountancy, Shandong University of Finance and Economics, Jinan, Shandong 250014, China*
[2] *Transportation Management College, Dalian Maritime University, Dalian 116026, China*
[3] *School of Automotive Studies, Tongji University, Shanghai 201804, China*
[4] *School of Traffic and Transportation, Beijing Jiaotong University, Beijing 100044, China*

Correspondence should be addressed to Guobao Ning; guobao_tj@163.com and B. Yu; yubjjt@126.com

An improved SVM model is presented to forecast dry bulk freight index (BDI) in this paper, which is a powerful tool for operators and investors to manage the market trend and avoid price risking shipping industry. The BDI is influenced by many factors, especially the random incidents in dry bulk market, inducing the difficulty in forecasting of BDI. Therefore, to eliminate the impact of random incidents in dry bulk market, wavelet transform is adopted to denoise the BDI data series. Hence, the combined model of wavelet transform and support vector machine is developed to forecast BDI in this paper. Lastly, the BDI data in 2005 to 2012 are presented to test the proposed model. The 84 prior consecutive monthly BDI data are the inputs of the model, and the last 12 monthly BDI data are the outputs of model. The parameters of the model are optimized by genetic algorithm and the final model is conformed through SVM training. This paper compares the forecasting result of proposed method and three other forecasting methods. The result shows that the proposed method has higher accuracy and could be used to forecast the short-term trend of the BDI.

1. Introduction

The BDI published by the Baltic Exchange is used as an important evaluation factor for the dry bulk market in shipping industry. It is usually consulted by shipping operators and investors to forecast the trend of dry bulk market. However, as the price of dry bulk market changes almost every day and the affecting factors of the price are complicated, the prediction of the trend of dry bulk market becomes of difficulty. Since 2001, the BDI has experienced a huge fluctuation. The value of BDI was less than 1000 points at that time and increased to more than 11000 points in May, 2008. Five months later, it decreased to less than 800 points. Therefore, research on the law of shipping market freight fluctuation and the forecasting of the trend of BDI is of special significance for operators and investors to manage the market trend and avoid price risk in shipping industry.

Since the Baltic Freight Index (BFI) was established in 1985, many researchers and shipping scholars have made in-depth research on the volatility and trend prediction of BFI and subsequent BDI. However, none of the forecasting methods is widely used in BDI prediction. The remarkable work in this field has been done by Kavussanos, who worked on the dry bulk market prices issues as early as the 1990s. Kavussanos and Visvikis [1] used VECM-GARCH model to investigate the lead-lag relationship in both return and volatilities between spot and future markets. Cullinake [2] is also a pioneer in developing BFI index forecasting method with ARIMA model. After that, some prediction techniques, such as statistical regression and neural network, are widely used in BDI prediction. However, the method of statistical regression is appropriate for stability, normality, and independence time series and not appropriate for complex time series. Neural network has a good nonlinear approximation capability, but the model structure is difficult to determine. It is prone to excessive training or insufficient training for the neural network, which induces some shortages such as trapping in local minimum, being sensitive to initial value,

and overreliance on design skills. Li and Parsons [3] used the neural network to forecast the tanker freight rate and then compared with ARMA model. They proved that the neural network model has higher prediction accuracy for a long time series. Cullinane et al. [4] forecasted the spot freight rate index with a simple single variable in ARMA model. A new model-fractional integrated autocorrelation moving average model (ARFIMA) is presented by Granger [5] in data forecasting. Based on the ARFIMA approach, Henry [6] showed that almost a half of the new-coming seventeen stock markets in the world in the 1990s have long memory, which is similar to their searches by Crato [7], Jonahan [8], and so on.

In recent years, econometric models are popular in freight rate index forecasting, for instance, ARIMA model, VAR model, and VECM model. Empirical analysis shows that the econometric models have higher accuracy than the traditional prediction methods for nonstationary time series. Veenstra and Franses [9] developed VAR model in forecasting BDI based on cointegrated process of the high time series and the method of unit root test. Kavussanos and Alizadeh-M [10] presented a single variable seasonal ARIMA-SARIMA model, multivariable seasonal cointegrating, and VAR to analyze the seasonal characteristics of dry bulk shipping market.

Tvedt [11] used the unit root tests to reject the random walk hypothesis of freight rate, confirming a state of stationary in freight rates forecasting for the classical shipping market models. Two years later, a rejection of applicability of the expectations theory in freight market was presented by Adland and Cullinane [12]. They also proved that the risk premium is time varying, depending on the freight market conditions and time charter duration. A fuzzy-DELPHI adjustment process for improvement of accuracy was proposed by Duru et al. [13]. They also illustrate its performance in the validation of adjustments of statistical forecasts in the BDI through an empirical study. Zhang et al. [14] employed R/S and GPH tests to model long memory of volatility of the indices based on the investigation of fluctuation features of dry bulk shipping market with the BDI.

A new machine learning method, namely, support vector machine (SVM), is widely employed in many fields, for instance, handwriting recognition, three-dimension objects recognition, faces recognition, text images recognition, voice recognition, and regression analysis [15–23]. The SVM based on statistical learning theory has good fitting ability for complex nonlinear function. At the same time, it can avoid trapping problem of overfitting learning. A support vector machine-based (SVM) model is developed to predict the baseline travel and dwell times of buses based on recent data by Yu et al. [24, 25]. The authors [24, 25] also presented a hybrid model based on support vector machine (SVM) and Kalman filtering technique to predict bus arrival times. Yu et al. [26] also adopted support vector machine (SVM), artificial neural network (ANN), k nearest neighbor's algorithm (k-NN), and linear regression (LR) for the bus arrival time prediction. van Gestel et al. [27] research the financial time series prediction problems based on the least squares SVM model. Cao and Tay [28] research the parameters of SVM selection and test their approach with empirical analysis of financial time series. The SVM model is widely used in the stock market. Kim [29] compared the forecasting result of stock price index between SVM model with the BP neural network model and the CBR model. He proved that the SVM model has better prediction of financial time series. Huang et al. [21] applied SVM to forecast the movement direction of NIKKEI 225 index.

It is suitable for nonlinear time series prediction by the SVM since the full consideration with the randomness of the data sequence. From the application of SVM especially in financial index prediction, as well as freight data, it is no hard to be employed in another index prediction in terms of applicability. However, SVM is rarely studied and used in the field of BDI forecast. Therefore, this paper will analyze the applicability of SVM model in BDI forecasting and make an empirical test.

BDI data sequences can be regarded as signals changed over time. The signals usually have characteristics, such as periodic and seasonal, of dry bulk shipping market freight index fluctuation. The noise of the signal reflects the influence factors of freight index or random events. To grasp the rule of BDI data variation, the random disturbance factors should be eliminated. Therefore, denoising processing of the signal is necessary for an accuracy data forecasting. Some methods, such as a self-adaptive filtration method, the Kalman filter method, and average moving method, are often used to denoise signal of data. Yu et al. [30] adapt an adaptive filtration method in bus arrival time prediction model. However, the method of wavelet transform is proposed to denoise the raw signal in this paper.

Wavelet transform is the most widely used multiscale analysis method till now. The root of wavelet transform is scaling and translation in the signal analysis. It is a milestone in the history of the development of Fourier analysis. In 1982, a French oil exploration technician called J. Morfet tried to deal with irregular signals with wavelet method more effectively. Research of wavelet analysis was becoming popular after that. It is widely used in image processing, voice processing, and signal processing. In recent years, the wavelet transform method has been employed in areas of finance and economy, and the empirical result showed a good effect. Esteban et al. [31] predicted time series with wavelet method. Firstly, they decomposed the time series into two different cycles in sequence with wavelet. Then, ARMA model is present in regression of each data cycle. Lastly, the two forecasting values are added to get the final prediction results. They proved that their proposed approach had better predicted result than ARMA model.

Moreover, some researchers also tried to establish the combined model with the wavelet transform and SVM [32–39]. Wu [40] proposed novel robust wavelet support vector machine, which is based on wavelet theory and the modified support vector machine. They also designed swarm optimization algorithm to select the optimal parameters of their proposed model. Two years later, Wu and Law [41] made an in-depth research based on the wavelet support vector machine proposed by Wu [40]. Wavelet transform is also used by Guo to map the sample data into several time-frequency domains. He then developed the SVM model to forecast the

gross value of textile products in Japan precisely. A wavelet transform and SVM combined model is developed by Hsieh and Chen to predict the dissolved oxygen density in water-quality process. Their results showed a higher accuracy of the combined model than BP neural network model. Liu and Fan [42] stated that the performance of SVM can be improved with the introducing of discrete wavelet transform. Wang et al. [43] used the wavelet transform to decompose the dam deformation time series into different frequency components and then forecast the series with a SVM model. A wavelet kernel function for SVM is presented by Wei and Lin [17]; they also denoised the signal with multiscale interpolation and sparse attributes. The performance proved that their proposed model was accurate and convergent.

Although there are many studies of the combination of wavelet transform and SVM, few have been made in dry bulk freight index. Therefore, this paper constructs a wavelet transform and SVM combined forecast model. It removes the random factors in BDI series with wavelet and then establishes a SVM model. The numerical analysis shows that our method has better predicting results than the commonly used prediction methods. Of course as a prediction method, it should be tested with large numbers of tests while evaluating the accuracy of its prediction, which is, however, the shortage of this paper for the time limited.

This paper is organized as follows. Section 2 reviews the data on the shipping freight market and analyzes the future of the BDI data. Section 3 presents the decomposition and reconstruction of wavelet. The forecasting models and procedures are proposed in Section 4. A case study is shown in Section 5 and the performance of several prediction results is compared. Besides, the conclusion of this paper and the recommendations for future studies are provided in this part.

2. Characters and Influence Factors of BDI

The volatility of freight is directly reflected by the fluctuation of freight indexes, for example, BDI. In terms of market structure, freight price depends on the supply (ship owner) and the demand (cargo). Concerning market economies, freight price of general cargo is mainly influenced by three external factors: an act of war or natural disaster, the global economy, and the market speculation.

In retrospect, force majeure, for example, war factors, is the major power driving the fluctuation of the world shipping market, especially in the turbulent times. In 1956, the outbreak of Suez Crisis drastically increased the shipping market risks throughout the world. Shipping lines and area changed a lot, and supply in dry bulk shipping market rapidly went down, which led to the high volatility of freight price. In 1973, the third Middle East war broke out; Arab countries firstly used the "oil weapon," resulting in the sharp increasing of fuel price and freight price consequently. The First World War, the Second World War, the Middle East wars, hurricane, tsunamis, and other natural disasters brought high risks to maritime transport market. Firstly, wars and natural disasters such as force majeure occurrence or even expectation of those events can affect the confidence of both ship owners and shippers; secondly, once sailing area is limited, such as the close of Suez Canal during the second Middle East war, the average travel distance will increase and the supply capacity will drop significantly; besides, the rise of oil prices due to wars will also increase shipping costs.

Shipping derivation has shown that the world economic situation and the development of international trade play a decisive role in the shipping market existence and changes. Therefore, the economic cycle and trade demand are the durable and fundamental influences in the shipping market. The most remarkable presentation of the impact of economic environment on maritime shipping was the terrible hit of the global economic crisis to the shipping industry. Economic crisis led to slower global economic growth and commodity prices falling sharply. In the first half of 2009, the fixed capital formation and manufacturing output of the world's major economies have double-digit decline. Steel mills and other enterprises, in order to cope with shrinking demand, take measures of limiting production or semiproduction, which led to the demand on iron ore and coal dropping significantly. Dry Freight Index experienced unprecedented volatility in the six months from the highest point in history falling to the lowest point. The demand of iron ore which was the largest dry bulk seaborne trade at that time decreased significantly. China, as the largest importer, unloaded 30 million tons of iron ore imports in Nov., 2008, down by 20.7%, which was the first negative growth. The impact of economic crisis on supply capacity is mainly reflected in the shipbuilding market. Global economic downturn led to sharp decline in shipbuilding demand, and some ship owners began to cancel the order because of the shortage of money. Because there are one to two years of construction time from ordering to delivery, the impact of the economic crisis on the shipbuilding industry has one to two years of lag extending to freight market. Therefore, new ships to be delivered two years after ordering will substantially decrease, resulting in shrinking supply capacity.

Freight derivatives were created in order to avoid the risk of emergencies in shipping market. Major functions of freight derivatives reside in hedging and price discovery. Freight derivatives include the Baltic Freight Index futures (BIFFEX), forward freight agreements (FFA), and the shipping options (freight option). Volatile freight rates since 2004 have given speculators profit opportunities. Investment banks as Goldman Sachs, Morgan Stanley, and other financial institutions and hedge funds have entered into speculative market; some shipping companies also use their information superiority to engage in the market.

According to dry bulk freight index trend, freight index, it disorderly changes in random variation, so it is difficult to grasp the change regulation. In order to better grasp the inherent regulation of fluctuations, it can be divided into two categories: the first category is the one in which there is a pattern existing, for example, the world economy with cyclical characteristics, coal, iron ore, grain production capacity, and shipping capacity with seasonal fluctuations; the second category is sudden and random factors, for example, natural climate political events, average travel distance, scientific and technological development, country's international

trade policy, sudden changes in trade structure, economic interest transferred, exchange rate fluctuations, ship archive, operational productivity, international shipping norms, and market rumors.

After the above analysis, both the two factors have effect on the dry bulk freight index. To grasp more accurate freight index fluctuation characteristics, what needs to be done is to dig out the historical dry bulk freight index data and then use data processing methods to eliminate the disorder characteristics caused by the second category of factors. Based on that, the most suitable methods are used for BDI prediction.

If the time series of BDI can be regarded as a kind of changeable signal with time elapsing, then there is rich information in the signal. The first category includes the information of cyclical fluctuations of BDI. As the cycle is long-term, the first category factors have lower frequencies and are located in low frequency range. The second category factors are stochastic, irregular, and unexpected. Though those factors occur not very often, the frequency can be still relatively high if aggregating the second category factors into monolithic. So the high frequency range includes the second category factors. The discussion on cyclical fluctuation of BDI is based on thought as follows: (1) signal reconstruction. Extracting BDI signal process attempts to remove stochastic, irregular and unexpected factors, and noise from the BDI signal by separating the low and the high frequency part. (2) BDI is an output of a complex function as there are so many factors impacting on the dry bulk freight market. In order to analyze the BDI signal accurately, this paper applies the SVM onto the prediction of the reconstructed signal based on the results of extracting BDI signal process.

3. Adopting the Wavelet Transform to Denoise the BDI

Useful signal is commonly presented as stationary signals or low frequency signals, while noise signal is usually unstable and has high frequency. Therefore, the characteristics of BDI ensure the application of wavelet analysis to eliminate noise signal. When using wavelet analysis to remove noise signal from shipping indexes, such noise signal is mainly included in high frequency wavelet coefficients, for which the threshold method can be used for decomposing wavelet coefficients. Each layer of decomposed wavelet coefficients should be reconstructed to eliminate the noise. The purpose of removing noise signals from BDI signal $S(t)$ is to obtain actual signal $f(t)$ from $S(t)$, by which the authenticity of data can be ensured.

The one-dimension model of BDI signal with noises can be presented as follows:

$$S(t) = f(t) + \sigma e(t), \quad t = 0, 1, \ldots, n-1, \tag{1}$$

where $f(t)$ is the real signal; $e(t)$ is the noise; σ is the noise intensity; $S(t)$ is the signal with noises.

The process of wavelet noise reduction is the process of decomposition and reconstruction for signal. Original function or signal is split into several relevant pieces without
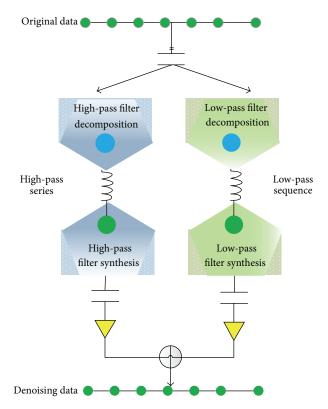

Figure 1: Decomposition and reconstruction of wavelet transform denoising process.

losing much information. Those pieces are such wavelet which changes in scale and decays in time. The wavelet reconstruction is the process where those pieces are combined to restore the real features.

The decomposition and reconstruction of wavelet are shown as in Figure 1.

BDI is one-dimension time series and the wavelet denoising process against such kind of signal is usually expressed as the procedure presented as follows.

Step 1. Preprocessing the data, which include noises, for using in next steps.

Step 2. Wavelet denoising process to the one-dimension signal. Selecting a suitable wavelet mother function and setting an appropriate decomposing layer N. Decomposing $S(t)$ into N layers.

Step 3. Quantizing the threshold of wavelet decomposition coefficients. Selecting a suitable threshold for the high frequency coefficient of each layer.

Step 4. Inverse transform of one-dimension wavelet. Based on the coefficient of Nth layer and the quantized high frequency coefficients from 1st to Nth layer, reconstructing the one-dimension signal. The reconstructed signal is the denoised signal.

Theoretical base of wavelet denoising is presented as follows.

$\psi(t)$ is a function where Fourier transform exists. If its Fourier transform $\widehat{\psi}(t)$ meets the condition $\int_{-\infty}^{\infty}(\widehat{\psi}(t)^2/w)dw < \infty$, the function can be a wavelet function. Supposing $j \in Z$ and $\psi_{2^j}(t)$ is the dyadic stretching transformation of $\psi(t)$ against factor 2^j, then $\psi_{2^j}(t)$ can be expressed as

$$\psi_{2^j}(t) = \frac{1}{2^j}\psi\left(\frac{t}{2^j}\right). \tag{2}$$

Wavelet transform of function $f(t)$ with scale 2^j at position t can be defined as the convolution of $f(t)$ and $\psi_{2^j}(t)$, presented as

$$W_{2^j}f(t) = f \times \psi_{2^j}(t). \tag{3}$$

For wavelet function $\psi(t)$, supposing there exist constants A and B and $0 < A \le B < \infty$, then we can get

$$\forall \omega \in R, \quad A \le \sum_{j=-\infty}^{\infty}\widehat{\psi}\left(2^j\omega\right) \le B(2), \tag{4}$$

where $\widehat{\psi}(t)$ is the Fourier transform of $\psi(t)$. Then $\psi(t)$ can be called dyadic wavelet function and the corresponding wavelet transform can be called dyadic wavelet transform.

For any function $\chi(t)$ with Fourier transform, if its Fourier transform meets Subject (5)

$$\sum_{j=-\infty}^{\infty}\widehat{\psi}\left(2^j\omega\right)\widehat{\chi}\left(2^j\omega\right) = 1, \tag{5}$$

then it can be called reconstruction wavelet. It can be easily found that there are countless functions $\chi(t)$ meeting Subject (5).

The dyadic wavelet transform is complete and stable. The "complete" means that the function can be restored by its dyadic wavelet transform. In terms of energy, "stable" means that the total ability of dyadic wavelet transform has limitation which is close to the energy of the function. Function $f(t) \in L^2(R)$ can be restored by its dyadic wavelet transform and the corresponding reconstruction wavelet on the basis of

$$f(t) = \sum_{-\infty}^{\infty}W_{2^j}f \times \chi_{2^j}(t), \tag{6}$$

$$A\|f\|^2 \le \sum_{j=-\infty}^{\infty}\|W_{2^j}f(t)\|^2 \le B\|f\|^2. \tag{7}$$

In practical application, the measurable resolution of signal is limited, so it is impossible to conduct wavelet transform on all scales 2^j ($-\infty < i < \infty$). Therefore, 2^j should be set as a limited value. The wavelet transform is confined between a limited maximum scale $j = J$ and a limited minimum scale $j = 1$. 2^I is the highest resolution and 2^J is the lowest resolution. With respect to resolution, it is relevant to frequency. That is to say, the higher the frequency

is, the higher the resolution is, and vice versa. To express the signal resolution decomposition of wavelet transform, a real function $\varphi(t)$ is introduced hereafter, whose Fourier transform should meet Subject (8). Consider

$$\widehat{\varphi}(t)^2 = \sum_{j=1}^{+\infty}\widehat{\psi}\left(2^j\omega\right)\widehat{\chi}\left(2^j\omega\right). \tag{8}$$

According to (3) and (6), it can be easily obtained that

$$\widehat{\varphi}(0)^2 = \lim_{k \to 0}\widehat{\psi}\left(2^k\omega\right)^2 = \lim_{k \to 0}\widehat{\psi}\left(2^j\omega\right)\widehat{\chi}\left(2^j\omega\right) = 1$$

$$\widehat{\varphi}(\infty)^2 = \lim_{k \to +\infty}\widehat{\psi}\left(2^k\omega\right)^2 = \lim_{k \to +\infty}\widehat{\psi}\left(2^j\omega\right)\widehat{\chi}\left(2^j\omega\right) = 1. \tag{9}$$

Equation (9) indicates that the energy of $\widehat{\varphi}(\omega)$ gathers in the low frequency range, so $\varphi(t)$ is a smooth function with low-pass characteristics. A smooth operator S_{2^j} is defined as follows:

$$S_{2^j}f(t) = f * \varphi_2(t),$$

$$\varphi_2(t) = \frac{1}{2^j}\varphi\left(\frac{t}{2^j}\right), \tag{10}$$

where $S_{2^j}f(t)$ denotes the low-pass filtering component of signal $f(t)$ when the resolution is 2^j. The high frequency component of $f(t)$ is not presented in $S_{2^j}f(t)$ but in the dyadic wavelet transform $\{W_{2^j}f(t)\}_{1 \le j \le J}$ between scales 2^I and 2^J, so $W_{2^j}f(t)$ stands for the detailed component and $S_{2^j}f(t)$ means the low-pass smooth component of the signal. The signal details (the high frequency ingredient) contained in $S_{2^j}f(t)$ decrease with 2^j increasing, and the lost information can still be restored by the wavelet transform $W_{2^j}f(t)$.

The time series is defined as $S_2^{d_0}f$, and the low-pass smooth component at scale 2^j is defined as $S_2^{d_j}f$. According to (7), $S_2^{d_j}f$ can be split into the low and the high half frequency denoted by $S_2^{d_j}+1f$ and $W_2^{d_j}+1f$, respectively. The d is the concrete signal. The decomposition algorithm of $S_2^{d_j}f$ is shown as follows:

$j = 0$,
while ($j < J$),
$W_2^{d_j}+1f = (1/\lambda_j)S_2^{d_j}f * G_j$,
$S_2^{d_j}+1f = S_2^{d_j} * H_j$,
$j = j + 1$,
the end.

The reconstruction algorithm of $S_2^{d_0}f$ is shown as follows:

$j = J$,
while ($j > 0$),
$S_2^{d_j}+1f = \lambda_jW_2^{d_j}f * K_{j-1} + S_2^{d_j}f * H_{j-1}$,
$j = j - 1$,
the end,

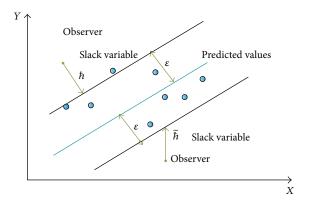where G_j, H_j, and K_j are a group of corresponding filters.

Figure 2: The ε-insensitivity tube of SVM.

4. Forecasting BDI with SVM

4.1. Support Vector Machine.
The support vector machine is a kind of machine learning system, with the purpose of maximizing the margin distance between different categories of problems [44–46]. The model of SVM is as follows:

$$f(x) = \omega \times \varphi(x) + b, \tag{11}$$

where ω is the weight vector; b is error; $\varphi(x)$ is a kernel function to deal with the nonlinear problem, with mapping the nonlinear input to a high dimensional space by a nonlinear function, to make the input linear.

The least square method in conventional regression model takes the square error as the loss function in accordance with minimizing empirical risks. Vapnik et al. [44] took the ε-insensitivity as the loss function in SVM model, and the ε-insensitivity loss is shown as

$$L_\varepsilon(f(x) - y) = \begin{cases} |f(x) - y| - \varepsilon & |f(x) - y| \geq \varepsilon \\ 0 & \text{others,} \end{cases} \tag{12}$$

where parameter ε determines the area of ε-insensitivity (Figure 2). When the predicted value $f(x)$ is within the tube area, the loss is zero; otherwise, the loss is the difference between the prediction error and the tube area radius ε. $\hbar$ and $\tilde{\hbar}$ are slack variables, indicating the prediction errors in different directions:

$$L_\varepsilon(f(x) - y) = \begin{cases} |f(x) - y| - \varepsilon = \hbar & |f(x) - y| \geq 0 \\ |f(x) - y| - \varepsilon = \tilde{\hbar} & |f(x) - y| < 0 \\ 0 & \text{others,} \end{cases} \tag{13}$$

where $\hbar$ is the training error which is higher than the area boundary; $\tilde{\hbar}$ is the training error which is lower than the area boundary.

In the input space, SVM uses the minimize-adjustment-risk function to calculate the weight vector and the error. The function is shown as

$$R(C) = C\frac{1}{N}\sum_{i=1}^{n} L_\varepsilon(f(x_i), y_i) + \frac{1}{2}\|w\|^2, \tag{14}$$

where $L_\varepsilon(f(x_i), y_i)$ is the ε-insensitivity loss function; $C(1/N)\sum_{i=1}^{n} L_\varepsilon(f(x_i), y_i)$ is the empirical error; $(1/2)\|w\|^2$ is the adjustment item. Then the SVM model can be figured out with minimizing

$$\text{Min}: \frac{1}{2}w^T w + C\sum_i \left(\hbar + \tilde{\hbar}\right)$$

$$\text{subject to} \begin{cases} y_i - w^T x_i - b \leq \varepsilon + \hbar \\ w^T x_i + b - y_i \leq \varepsilon + \tilde{\hbar} \\ \hbar\tilde{\hbar} \geq 0, \end{cases} \tag{15}$$

where $i = 1, 2, \ldots, n$ is the number of samples for training; $\hbar + \tilde{\hbar}$ is empirical risks; $(1/2)w^T w$ is structure risks which can avoid excessive learning; C is correction factor, indicating the balance between the experimental risk and the structure risk. Larger C means the model pays more attention to the experimental risk, otherwise, more attention to the structure risk. When C, ε, and the kernel function k which meets Mercer's condition are determined appropriately, the model can be solved with Lagrangian multiplier method.

Besides, in the process of artificial intelligent model construction, different data will lead to different combinations of best parameters. Therefore, the trial-and-error method is widely used to search the best parameter combination. With synthetically considering Cherkassky and Ma's suggestions [47] in parameter setting, this paper firstly applies Cherkassky and Ma's method [47] to estimate training data to calibrate several suggested parameter combinations (C and ε) of SVM model. Then the exponent search method is employed to select the best parameter combination, based on minimizing the mean square error. The method can prevent the risk of simple suggested parameter combination and also reduce the trial-and-error times.

4.2. Combined Model.
In this paper, wavelet transform decomposes the original sequence of BDI layer by layer and then gets a low frequency signal layer and N high frequency detailed layers (N is a decomposition level). Fluctuation of international dry bulk shipping market is included in the low frequency part of the BDI. The impact of random factors such as incidents is included in the high frequency part. But the high frequency part is not an irregular mutational factor. Therefore, it needs to denoise each layer sequence of low and high frequencies, respectively. A denoised BDI sequence is retained by reconstructing. The process of sequence denoising not only filters random factors but also makes the predictive model robust.
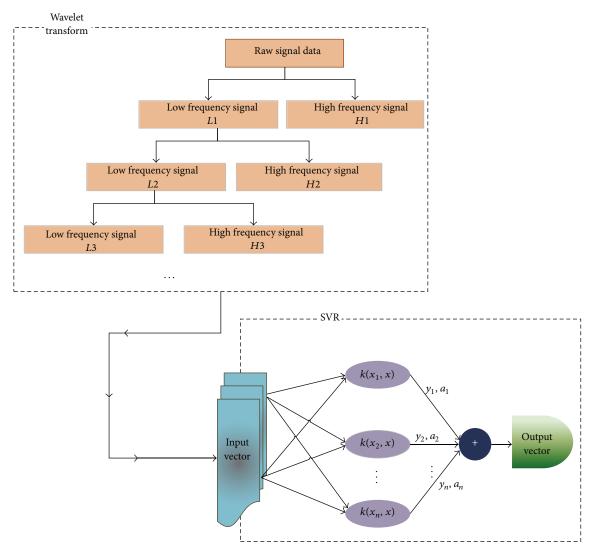
FIGURE 3: Structure of the wavelet transform-SVM combined model.

Wavelet transform has characteristics of time-frequency localization and zoom features, while support vector machine has nice tolerance of self-learning adaptive fault, generalization ability, and robustness. Through operation functions such as scaling and translation, wavelet transform is able to analyze functions or signals with multiscale refinement. Wavelet SVM is combined by the wavelet analysis and SVM can deal with nonlinear function approximation uniquely. This research uses wavelet transform to analyze BDI sequence and then trains the time series by SVM to get trained models and predictions. Figure 3 shows the structure of hybrid forecasting model.

FIGURE 4: Historical data of monthly averaged BDI (2005/1–2012/12).

5. Case Study

Since 2001, the BDI has experienced a huge fluctuation. The value of BDI was less than 1000 points at that time and increased to more than 11000 points in May, 2008. Five months later, it 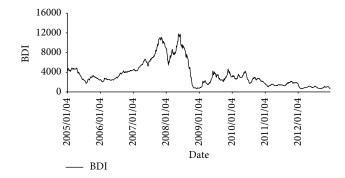decreased to less than 800 points. This paper takes data of the BDI published by the Baltic Exchange from January 2005 to December 2012 as the empirical objective. Besides, the daily BDI data is replaced by month data; that is, the objective data is the average BDI for each month.
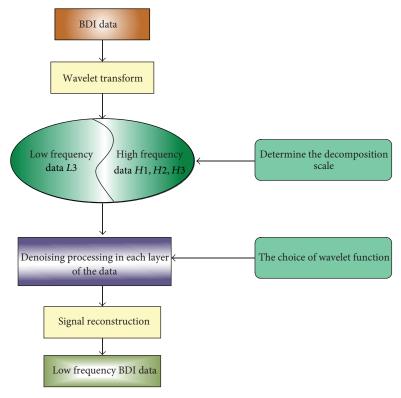
FIGURE 5: The wavelet transforming process of BDI series.

So there are 96 data of BDI. Among them, the 84 prior consecutive monthly BDI data are the inputs of the model, and the last 12 monthly BDI data are the outputs of model. The parameters of the model are selected and the final model is conformed through SVM training. Figure 4 shows the fluctuation phenomenon of monthly data.

5.1. Process Data. To avoid the training error resulting from dimension in sample data or a large dimension data value, the whole data should be normalized and processed before the SVM training. Consider

$$S_i' = 2 \cdot \frac{S_i - S_{\min}}{S_{\max} - S_{\min}} - 1, \tag{16}$$

where S_i' is normalized value. S_i is raw value. $S_{\min}$ is the minimum value in a sequence of samples. $S_{\max}$ is the maximum value in a sequence of samples.

5.2. Wavelet Analysis. The denoising process of original BDI sequence is presented by wavelet transform, which is shown in Figure 5. Figure 5 shows the wavelet transform process of BDI series. Firstly, the raw BDI data split into the high frequency data and the low frequency data decomposed with the wavelet transform. Then, by use of some tech-methods, such as threshold, each sequence will be processed with manic elimination. Go around and around until the final low frequency sequence is chosen.

Two problems, which wavelet function should be selected in denoising process and how to determine the decomposition scale, should be solved. Different wavelet function will get different wavelet transform analysis results, which is important for the effect of denoising. There is no acknowledged method about how to choose the optimal wavelet functions and decomposition scale for signal denoising. So this paper settles the above two problems with experiment.

The purpose of denoising is to remove the mutation factors and random effects in the sequence. So the denoised sequence should not be too smooth or existing obvious step phenomenon. Considering the orthogonality, symmetry, smoothness, and other characteristics of the wavelet function, the best wavelet function and the decomposition scale are determined. The paper used the wavelet toolbox of MATLAB to make the test.

The commonly used wavelet functions are *Haal wavelet*, *dbN wavelet*, *symN wavelet*, *biorN wavelet*, *coifN wavelet*, *dmey wavelet*, and so on. We make transformation analysis for the BDI sequence with the same scale and the same order number with different wavelet function. This paper will take three layers of decomposition. So the ₦N is selected as 3. After the experience, the *dbN wavelet* is selected as the one in denoising BDI sequence.

Then, different coefficients of dbN wavelet function are used to analyze wavelet transform. The coefficients of dbN wavelet function are usually selected from 1 to 6. Through effective comparison, the coefficient of dbN wavelet function is settled as 3.
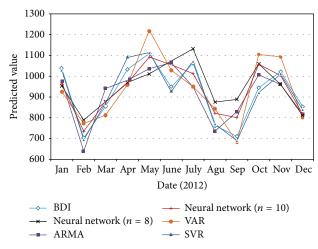
FIGURE 6: Forecasting results of four prediction models.

5.3. The Wavelet-SVM to Forecast BDI Sequence. The 84 prior consecutive monthly BDI data are the inputs of the model, and the last 12 monthly BDI data are the outputs of model. The SVM function with output close to the last 12 monthly BDI data will be selected. The parameters in SVM, which greatly influence the performance of SVM, need to be optimized and set by users. Heuristic algorithms have been successfully used in many complex problems [48–51]. Genetic algorithm (GA) is a common heuristic algorithm which has been widely used in lots of literatures [46, 52]. Therefore, GA is also used to optimize the three parameters C and ε for SVM. Due to lots of literatures about GA for references [46, 52], the process of GA has not been introduced in this paper. Before the implementation of GA, there are four GA parameters, namely, p_c, p_m, p_{size}, and T_{max}, which need to be predetermined. In general, p_c varies from 0.3 to 0.9. p_m varies from 0.01 to 0.1; p_{size} is the population size which is set according to the size of the samples. T_{max} is the maximum number of generation. At last, after the optimization of GA, the two parameters of SVM were optimized as (5.5 and 0.02) with the best optimization value.

Then, the trained model is presented for one-step prediction on the last 12 monthly data. To test the forecasting effect of mixed-model, three traditional econometric methods, ARIMA model, VAR model, and neural network model, are proposed for one-step prediction on the same sample data. Since the above three models use the raw BDI sequence as the input sample for index forecast, it has a strong comparability. Compare the results (Table 2) of one-step prediction with the actual value of BDI. For easy understanding and comparing, the actual and predicted values are antinormalized so that the data back to the real market freight index level. Figure 6 shows the compared results of the four predicted models.

As can be seen from Figure 6, the predicted results obtained from three models have the same trend with the actual value of BDI. However, among them, the deviation between the prediction results of neural network and the real value is the maximum. This is because that the international dry bulk market in 2007 and 2008 has always been in volatile

mood, causing the artificial neural network falling into the overlearning problem in the case of small samples. Therefore, it amplifies the up and down magnitude of BDI values for the BDI forecast after 2008. ARMA and VAR itself are suitable for short-term time series prediction, and results are better than the neural network model obviously. However, as can be seen in Figure 6, at some turning points, Wavelet-SVM model is more close to the true value than the ARMA model. Table 1 shows the forecasting value of each prediction model.

This paper uses root mean square error (RMSE) to test training effect and forecasting precision of the various forecasting methods:

$$\text{RMSE} = \left(\sum_{i=1,\dots,N} \frac{\left(S_{fi} - S_{ri}\right)^2}{N} \right)^{1/2}, \qquad (17)$$

where S_r is the actual value of BDI index and S_f is the prediction value.

By calculating the RMSE of the above four models with the forecasting result, we see that the wavelet-SVM hybrid prediction model has the best prediction accuracy. The large deviation among the four models is related with the fall of BDI under the influence of the economic crisis in 2008. BDI value fell more than 90% from more than 17000 points in May, 2008, to less than 700 points in end of 2008. Therefore, seeing from the predicted trend and the prediction accuracy of each forecasting model, wavelet SVM is the most suitable method in short-term prediction of BDI.

6. Conclusions

Research on the law of shipping market freight fluctuation and the forecasting of the trend of BDI is of special significance for operators and investors to manage the market trend and avoid price risk in shipping industry. Therefore, this paper constructs a wavelet transform and SVM combined forecast model. It removes the random factors in BDI series with wavelet and then establishes a SVM model. The BDI data in 2005 to 2012 are presented to test the proposed model. The 84 prior consecutive monthly BDI data are the inputs of the model, and the last 12 monthly BDI data are the outputs of model. The parameters of the model are selected and the final model is conformed through SVM training. This paper compares the forecasting result of proposed method with three other forecasting methods (VAR model, ARMA model, and neural network). The result shows that the proposed method has higher accuracy and could be used to forecast the short-term trend of the BDI. In further research, we will be devoted to improving the prediction accuracy and to forecasting the BDI with long-term period.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

TABLE 1: Forecasting results of five prediction models.

	BDI	ARMA		VAR		Neural network ($n = 8$)		Neural network ($n = 10$)		SVM	
		Predict value	Relative error	Predict value	Relative error	Predict value	Relative error	Predict value	Relative error	Predict value	Relative error
Jan. 12	1039.381	977.1347	0.059888	925.1847	0.109869	953.5296	0.082599	963.4226	0.07308	1029.505	0.009502
Feb. 12	702.619	639.3836	0.09	772.8565	0.099965	789.4652	0.123603	735.1234	0.046262	694.6194	0.011385
Mar. 12	855.381	941.0745	0.100182	812.9358	0.049621	877.5513	0.025919	876.2803	0.024433	870.5223	0.017701
Apr. 12	1032.905	980.6357	0.050604	959.705	0.070868	969.8216	0.061074	968.0003	0.062837	1090.248	0.055517
May 12	1109.762	1036.372	0.066132	1221.098	0.100325	1012.03	0.088066	1092.27	0.015762	1114.378	0.00416
June 12	947	1065.348	0.124972	1029.246	0.086849	1072.941	0.132989	1055.275	0.114335	925.786	0.022401
July 12	1064.048	948.7609	0.108347	950.0095	0.107174	1132.184	0.064036	1012.775	0.048186	1075.023	0.010315
Aug. 12	763.5714	735.4512	0.036827	843.5304	0.104717	875.5779	0.146688	822.0825	0.076628	770.095	0.008543
Sep. 12	710.381	829.6896	0.16795	682.384	0.039411	888.8169	0.251183	800.9311	0.127467	690.1364	0.028498
Oct. 12	944.619	1007.365	0.066424	1104.807	0.169579	1059.535	0.121653	1057.672	0.119681	921.4623	0.024514
Nov. 12	1021.714	960.0282	0.060375	1093.676	0.070432	959.9639	0.060438	1003.558	0.017771	1010.307	0.01165
Dec. 12	855.6875	817.9351	0.04419	799.9326	0.065158	811.3679	0.051794	818.812	0.043095	840.0157	0.018315

Table 2: RMSE of the three prediction models.

Model	RMSE
ARMA	78.96465
VAR	90.56454
Neural network ($n = 8$)	96.95657
Neural network ($n = 10$)	65.97173
SVM	21.67031

Acknowledgments

References

[1] M. G. Kavussanos and I. D. Visvikis, "Market interactions in returns and volatilities between spot and forward shipping freight markets," *Journal of Banking and Finance*, vol. 28, no. 8, pp. 2015–2049, 2004.

[2] K. Cullinake, "A short adaptive forecasting modal for BIFFEX speculation: a Box-Jenkins approach," *Maritime Policy & Management*, vol. 2, pp. 91–114, 1992.

[3] J. Li and M. G. Parsons, "Forecasting tanker freight rate using neural networks," *Maritime Policy & Management*, vol. 24, no. 1, pp. 9–30, 1997.

[4] K. P. B. Cullinane, K. J. Mason, and M. Cape, "A comparison of models for forecasting the Baltie Freight Index: Box-Jenkins revisited," *International Journal of Maritime Economies*, vol. 1, no. 2, pp. 15–39, 1999.

[5] C. W. J. Granger, "Long memory relationships and the aggregation of dynamic models," *Journal of Econometrics*, vol. 14, no. 2, pp. 227–238, 1980.

[6] O. T. Henry, "Long memory in stock returns: Some international evidence," *Applied Financial Economics*, vol. 12, no. 10, pp. 725–729, 2002.

[7] N. Crato, "Some international evidence regarding the stochastic behavior of stock returns," *Applied Financial Economics*, vol. 4, no. 1, pp. 33–39, 1994.

[8] H. W. Jonahan, "Long memory in emerging stock market returns," Federal Reserve System Working Paper 650, 1999.

[9] A. W. Veenstra and P. H. Franses, "A co-integration approach to forecasting freight rates in the dry Bulk shipping sector," *Transportation Research Part A*, vol. 31, no. 6, pp. 447–458, 1997.

[10] M. G. Kavussanos and A. H. Alizadeh-M, "Seasonality patterns in dry bulk shipping spot and time charter freight rates," *Transportation Research Part E*, vol. 37, no. 6, pp. 443–467, 2001.

[11] J. Tvedt, "A new perspective on price dynamics of the dry bulk market," *Maritime Policy and Management*, vol. 30, no. 3, pp. 221–230, 2003.

[12] R. Adland and K. Cullinane, "A time-varying risk premium in the term structure of bulk shipping freight rates," *Journal of Transport Economics and Policy*, vol. 39, no. 2, pp. 191–208, 2005.

[13] O. Duru, E. Bulut, and S. Yoshida, "A fuzzy extended DELPHI method for adjustment of statistical time series prediction: An empirical study on dry bulk freight market case," *Expert Systems with Applications*, vol. 39, no. 1, pp. 840–848, 2012.

[14] H. Zhang, F. Wei, and Z. Zhang, "Modeling volatility of baltic dry bulk freight index," in *Proceedings of the IEEE International Conference on Automation and Logistics (ICAL '08)*, vol. 9, pp. 1089–1094, September 2008.

[15] B. L. Koley and D. Dey, "Automatic detection of sleep apnea and hypopnea events from single channel measurement of respiration signal employing ensemble binary SVM classifiers," *Measurement*, vol. 46, no. 7, pp. 2082–2092, 2013.

[16] M. G. Poddar, V. Kumar, and Y. P. Sharma, "Linear-nonlinear heart rate variability analysis and SVM based classification of normal and hypertensive subjects," *Journal of Electrocardiology*, vol. 46, no. 4, p. e25, 2013.

[17] Y. C. Wei and C. H. Lin, "A robust video text detection approach using SVM," *Expert Systems with Applications*, vol. 39, no. 12, pp. 10832–10840, 2012.

[18] X. M. Chen, H. B. Gong, and J. N. Wang, "BRT vehicle travel time prediction based on SVM and Kalman filter," *Journal of Transportation Systems Engineering and Information Technology*, vol. 12, no. 4, pp. 29–34, 2012.

[19] O. Duru, "A fuzzy integrated logical forecasting model for dry bulk shipping index forecasting: an improved fuzzy time series approach," *Expert Systems with Applications*, vol. 37, no. 7, pp. 5372–5380, 2010.

[20] B. Yu and Z. Z. Yang, "An ant colony optimization model: the period vehicle routing problem with time windows," *Transportation Research Part E*, vol. 47, no. 2, pp. 166–181, 2011.

[21] W. Huang, Y. Nakamori, and S.-Y. Wang, "Forecasting stock market movement direction with support vector machine," *Computers and Operations Research*, vol. 32, no. 10, pp. 2513–2522, 2005.

[22] K. K. Seo, "An application of one-class support vector machines in content-based image retrieval," *Expert Systems with Applications*, vol. 33, no. 2, pp. 491–498, 2007.

[23] B. Wohlberg, D. M. Tartakovsky, and A. Guadagnini, "Subsurface characterization with support vector machines," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 1, pp. 47–57, 2006.

[24] B. Yu, Z. Z. Yang, K. Chen, and B. Yu, "Hybrid model for prediction of bus arrival times at next station," *Journal of Advanced Transportation*, vol. 44, no. 3, pp. 193–204, 2010.

[25] B. Yu, J. B. Yao, and Z. Z. Yang, "An improved headway-based holding strategy for bus transit," *Transportation Planning and Technology*, vol. 33, no. 3, pp. 329–341, 2010.

[26] B. Yu, W. H. K. Lam, and M. L. Tam, "Bus arrival time prediction at bus stop with multiple routes," *Transportation Research Part C*, vol. 19, no. 6, pp. 1157–1170, 2011.

[27] T. van Gestel, J. A. K. Suykens, D. E. Baestaens et al., "Financial time series prediction using least squares support vector machines within the evidence framework," *IEEE Transactions on Neural Networks*, vol. 12, no. 4, pp. 809–821, 2001.

[28] L. J. Cao and F. E. Tay, "Support vector machine with adaptive parameters in financial time series forecasting," *IEEE Transactions on Neural Networks*, vol. 14, no. 6, pp. 1506–1525, 2003.

[29] K. J. Kim, "Financial time series forecasting using support vector machines," *Neurocomputing*, vol. 55, no. 1-2, pp. 307–319, 2003.

[30] B. Yu, B. Yu, J. Lu, and Z. Z. Yang, "An adaptive bus arrival time prediction model," *Proceedings of the Eastern Asia Society for Transportation Studies*, vol. 7, 2009.

[31] M. Esteban, C. Ariño, and J. M. Díaz-Cruz, "Chemometrics for the analysis of voltammetric data," *TrAC Trends in Analytical Chemistry*, vol. 25, no. 1, pp. 86–92, 2006.

[32] Z. Z. Yang, L. J. Jin, and M. H. Wang, "Forecasting Baltic Panamax index with Support Vector Machine," *Journal of Transportation Systems Engineering and Information Technology*, vol. 11, no. 3, pp. 50–57, 2011.

[33] P. Du, K. Tan, and X. Xing, "Wavelet SVM in Reproducing Kernel Hilbert Space for hyperspectral remote sensing image classification," *Optics Communications*, vol. 283, no. 24, pp. 4978–4984, 2010.

[34] I. Turkoglu and E. Avci, "Comparison of wavelet-SVM and wavelet-adaptive network based fuzzy inference system for texture classification," *Digital Signal Processing*, vol. 18, no. 1, pp. 15–24, 2008.

[35] G. Y. Chen and W. F. Xie, "Pattern recognition with SVM and dual-tree complex wavelets," *Image and Vision Computing*, vol. 25, no. 6, pp. 960–966, 2007.

[36] H. Keskes, A. Braham, and Z. Lachiri, "Broken rotor bar diagnosis in induction machines through stationary wavelet packet transform and multiclass wavelet SVM," *Electric Power Systems Research*, vol. 97, pp. 151–157, 2013.

[37] Y. Zheng, L. Zhu, and X. Zou, "Short-term load forecasting based on Gaussian wavelet SVM," in *Proceedings of the 1st International Conference on Smart Grid and Clean Energy Technologies (ICSGCE '11)*, pp. 387–393, September 2011.

[38] B. Yu, Z. Yang, and S. Li, "Real-time partway deadheading strategy based on transit service reliability assessment," *Transportation Research Part A*, vol. 46, no. 8, pp. 1265–1279, 2012.

[39] V. Fernandez, "Wavelet- and SVM-based forecasts: an analysis of the U.S. metal and materials manufacturing industry," *Resources Policy*, vol. 32, no. 1-2, pp. 80–89, 2007.

[40] Q. Wu, "The forecasting model based on wavelet ν-support vector machine," *Expert Systems with Applications*, vol. 36, no. 4, pp. 7604–7610, 2009.

[41] Q. Wu and R. Law, "An intelligent forecasting model based on robust wavelet ν-support vector machine," *Expert Systems with Applications*, vol. 38, no. 5, pp. 4851–4859, 2011.

[42] F. Y. Liu and M. Fan, "A hybrid support vector machines and discrete wavelet transform model in futures price forecasting," in *Advances in Neural Networks*, vol. 3973 of *Lecture Notes in Computer Science*, pp. 485–490, 2006.

[43] X. Wang, Q. Fan, C. Xu, and Z. Li, "Dam deformation prediction based on wavelet transform and support vector machine," *Geomatics and Information Science of Wuhan University*, vol. 33, no. 5, pp. 469–507, 2008.

[44] V. Vapnik, M. R. Müller, A. J. Smola, G. Rätsch, B. Schölkopf, and J. Kohlmorgen, "Predicting time series with support vector machines," in *Artificial Neural Networks*, vol. 1327 of *Lecture Notes in Computer Science*, pp. 999–1004, Springer, Berlin, Germany, 1997.

[45] B. Z. Yao, C. Y. Yang, J. B. Yao, and J. Sun, "Tunnel surrounding rock displacement prediction using support vector machine," *International Journal of Computational Intelligence Systems*, vol. 3, no. 6, pp. 843–852, 2010.

[46] J. B. Yao, B. Z. Yao, L. Li, and Y. L. Jiang, "Hybrid model for displacement prediction of tunnel surrounding rock," *Neural Network World*, vol. 22, no. 3, pp. 263–275, 2012.

[47] V. Cherkassky and Y. Ma, "Practical selection of SVM parameters and noise estimation for SVM regression," *Neural Networks*, vol. 17, no. 1, pp. 113–126, 2004.

[48] B. Yao, C. Yang, J. Hu, J. Yao, and J. Sun, "An improved ant colony optimization for flexible job shop scheduling problems," *Advanced Science Letters*, vol. 4, no. 6-7, pp. 2127–2131, 2011.

[49] B. Z. Yao, P. Hu, X. H. Lu, J. J. Gao, and M. H. Zhang, "Transit network design based on travel time reliability," *Transportation Research Part C*, 2014.

[50] B. Z. Yao, P. Hu, M. H. Zhang, and S. Wang, "Artificial bee colony algorithm with scanning strategy for periodic vehicle routing problem," *SIMULATION: Transactions of the Society for Modeling and Simulation International*, vol. 89, no. 6, pp. 762–770, 2013.

[51] B. Z. Yao, P. Hu, M. H. Zhang, and X. M. Tian, "Improved ant colony optimization for seafood product delivery routing problem," *Promet: Traffic & Transportation*, vol. 26, no. 1, pp. 1–10, 2014.

[52] A. C. Lorena and A. C. P. L. F. de Carvalho, "Evolutionary tuning of SVM parameter values in multiclass problems," *Neurocom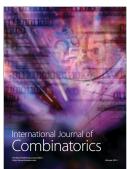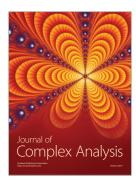puting*, vol. 71, no. 16-18, pp. 3326–3334, 2008.