CIT. Journal of Computing and Information Technology, Vol. 29, No. 1, March 2021, 27–38
doi: 10.20532/cit.2021.1005190

Prediction of Spot Price of Iron Ore Based on PSR-WA-LSSVM Combined Model

Xiangjun Cai[1] and Shihua Luo[2]

[1]Zhejiang college of Security Technology, Wenzhou, China
[2]School of Statistics, Jiangxi University of Finance and Economics, Nanchang, China

Aiming at the problems that the existing single time series models are not accurate and robust enough when it comes to forecasting the iron ore prices and the parameters of the traditional LSSVM model are difficult to determine, we propose a combined model based on Phase Space Reconstruction (PSR), wavelet transform and LSSVM (PSR-WA-LSSVM) to tackle these issues. ARIMA model, LSTM model, PSR-LSSVM model, and PSR-WA-LSSVM models were used for contrast simulation to forecast the spot price data of 61.5%PB powder from January 30, 2019, to February 1, 2021, in Ningbo Zhoushan port. The experimental results show that the PSR-WA-LSSVM combination model achieves better prediction results. At the same time, the model has a good performance in the multi-step prediction of the iron ore price.

ACM CCS (2012) Classification: Computing methodologies → Modeling and simulation → Model development and analysis → Model verification and validation

Keywords: iron ore spot price, price projection, combination model

1. Introduction

With the increase of economic globalization and the speeding up of domestic urbanization, demand for bulk commodities constantly increases while the industry is developing rapidly in our country. Constrained by the domestic environment and resources, our reliance on imported iron ore is increasing. As early as 2003, we surpassed Japan as the largest iron ore importer worldwide, importing iron ore up to 148 million tons that year. While in 2016, our country imported more than 1 billion tons of iron ore, up to 1.024 billion tons, creating a new highest record. As the largest iron ore consumer worldwide though, our country always has no say in the pricing of iron ore, only forced to accept the ever-rising price of iron ore.

From 1980 to 2009, the price of iron ore was determined through negotiation between major iron ore suppliers and their major customers in the world, namely the long-term agreement pricing (annually pricing) mechanism. The agreement protects the interests of both negotiation participants. Since the global iron ore long-term agreement pricing mechanism broke down in 2010, the pricing mode of iron ore was changing constantly, the price fluctuated increasingly dramatically. In 2011, the monthly pricing of iron ore gradually became a trend. In 2013, iron ore futures were officially listed and traded on the Dalian Commodity Exchange, and iron ore was given more financial attributes. The fluctuation of iron ore prices has seriously affected the interests of our country's iron ore manufacturers, consumer companies, and the country herself. Therefore, forecasting research on iron ore prices has attracted more and more attention. Spot trade is one of the important trading modes in iron ore import of our country. Comparing the spot price at the port with discounted features, the basic price can correctly reflect the spot price at the mine and the domestic port, so the spot price at the port is an important indicator reflecting iron ore price. In this paper, we selected Ningbo Zhoushan Port, the port with the largest cargo throughput in the world, with the spot price of 61.5% PB powder iron ore, as the subject investigated for the proposed prediction research. The main goal of our research is to grasp the fluctuating trend of iron ore.

2. Related Work

In recent years, scholars at home and abroad have made many attempts to predict the price of iron ore. In 2009, Zhu Bingwei [1] conducted wavelet analysis based on iron ore shipping price prediction research, attempting to predict price using the wavelet-ARIMA model. In 2014, Wang Qiang [2] *et al*. predicted the monthly price of iron ore in 2014 – 2015 by using exponential smoothing and combining three supply-demand conditions. In 2018, Fu Zhijie [3] analyzes the relationship and the issues between iron ore market from demand and supply side. He empirically explains how and why the import iron ore price in China fluctuates through Baltic Dry Index, Dollar Index, iron ore production and volume of import iron ore, and also proves that Vector Error Corrected Model (VECM) outperforms ARIMA model in forecast ability assessment. All these researches achieved valuable results. But from the angle of systems theory, the iron ore price forming mechanism is a nonlinear system that has many influencing factors, namely, a great difference in time lag and intensity of different influencing factors, and multi-variable coupling. Compared with traditional regression models, time series-based prediction often reveals the operating law of iron ore itself.

Above mentioned ARIMA model, exponential smoothing, VECM, and others, all achieved good fitting and prediction results, but with a deficiency in accuracy and robustness of the single model. During the years, the time series mixed prediction model drew much attention, and a lot of research emerged in many fields such as power [4], finance [5], navigation [6], and others. By combining ARIMA, GARCH, *etc.* with intelligent algorithms, such as SVM [7–8], neural networks [9–14], it is possible to predict linear parts and nonlinear parts of series respectively. By combining the overall prediction of paired series, such a combination attempt achieves significant prediction accuracy and improvement in robustness across the different fields.

Considering that the existing single time series models have certain space to improve accuracy and robustness in the prediction of iron ore price, it is difficult to determine parameters for the traditional least square support vector model (LSSVM). In this paper, a combined model based on Phase space reconstruction (PSR) and wavelet transform and LSSVM (PSR-WA-LSSVM) is proposed and is used to predict the spot price of powder iron ore. Comparison simulation shows that compared with a single ARIMA, single LSTM, and PSR-LSSVM model, PSR-WA-LSSVM combined model obtains a better prediction accuracy, with a hit rate of prediction increasing dramatically.

3. From Construction of Classic ARIMA Model to PSR-WA-LSSVM Combined Model

ARIMA model and SVM model have their advantages, respectively: the former is a classic selection for equipped to deal with linear problems, and the latter is very effective in dealing with nonlinear problems. LSSVM transforms inequation constraint of SVM method to equation constraint, dramatically improving the effectiveness of Lagrange multiplier α, optimizing the structure and solving process of the whole problem. In this paper, both theories and simulations are conducted in such a logical framework.

3.1. ARIMA Model

Autoregressive moving average (ARMA) is regularly used in the prediction of random phenomena in fields such as economics, physics, and others., and is very suitable for the prediction of stationary time series. ARIMA model is the optimization of ARMA, mainly transforming non-stationary series into stationary series through the order difference, then using parameters p, q to construct the model.

Suppose that Z_t is a group of non-stationary time series, then the ARIMA model of Z_t can be expressed as:

$$\begin{aligned}\nabla^d Z_t = \varphi_1 \nabla^d Z_{t-1} + \dots + \varphi_p \nabla^d Z_{t-p} - \\ - \theta_1 \alpha_{t-1} - \dots - \theta_q \alpha_{t-q} + \alpha_t\end{aligned} \tag{1}$$

where ∇^d is a different factor. After the stationary series is obtained, autocorrelation function

ACF, partial autocorrelation function PAC are calculated first. Secondly, orders of the model can be preliminarily determined through truncation, and parameter identification of time series can be obtained using the least square estimate. The least-square method finds the best function match of the data by minimizing the sum of squares of the error. Thus, the least square method can quickly obtain anonymous data and minimize the sum of squares of errors between the obtained data and the actual data. It can also estimate the parameters very well. The formula is as follows:

$$\sum_{t=1}^{N} \alpha_t^2 = \sum_{t=1}^{N} \left(\theta_q^{-1}(Z)_{\varphi_p}(Z) \nabla^d y_t \right)^2 \tag{2}$$

When the above equation takes the minimum value, estimate parameters φ_1, φ_2, ..., φ_p, θ_1, θ_2, ..., θ_q are obtained. Finally, when constructing different parameters combined ARIMA model, the optimal model is selected by testing the Akaike information criterion (AIC).

3.2. Least Squares Support Vector Machine Model (LSSVM)

SVM based on improved LSSVM uses the least-square linear system as loss function, uses an equation to replace inequation constraint condition of SVM, transforms solution of quadratic programming problem into a solution of linear equation set, simplifies the calculation, and increases the convergence speed of the algorithm. Its basic principle is defined as follows:

For a given sample set $x_i \in R^d$, it selects a suitable nonlinear function φ to map sample data from the original characteristic space to a high dimension characteristic space. The regression function can be constructed as follows:

$$f(x) = (\omega, \varphi(x)) + b \tag{3}$$

where ω is the weight vector and b is deviation. Based on the structural risk minimization (SRM) principle, the target of least squares support vector machine optimization is as follows:

$$\begin{cases} \min \frac{1}{2}\omega^2 + \frac{1}{2}\gamma \sum_{i=1}^{l} e_i^2, \\ \text{s.t. } \omega^T \varphi(x_i) + b + e_i = y_i, \ i = 1, ..., l \end{cases} \tag{4}$$

where e_i is an error, $e \in R^{l\times 1}$ is error vector, and γ is the regularization parameter. If we introduce the Lagrange multiplier α_i, $\alpha_i \in R^{l\times 1}$, then the Lagrange polynomial of its dual problem is as follows:

$$\min J = \frac{1}{2}\omega^2 + \frac{1}{2}\gamma \sum_{i=1}^{1} e_i^2 - \\ - \sum_{i=1}^{l} \alpha_i \left(\omega^T \varnothing(x_i) + b + e_i - y_i \right) \tag{5}$$

According to Karush-Kuhn-Tucker condition, a formula can be obtained as follows:

$$\begin{cases} \frac{\partial J}{\partial \omega} = 0 \ \rightarrow \ \sum_{i=1}^{l} \alpha_i \varphi(x_i) = 0 \\ \frac{\partial J}{\partial b} = 0 \ \rightarrow \ \sum_{i=1}^{l} \alpha_i = 0 \\ \frac{\partial J}{\partial e_i} = 0 \ \rightarrow \ \alpha_i = \gamma e_i, \ i = 1, 2, ..., l \\ \frac{\partial J}{\partial \alpha_i} = 0 \ \rightarrow \ \omega^T \varphi(x_i) + b + e_i - y_i = 0, \\ \qquad i = 1, 2, ..., l \end{cases} \tag{6}$$

The problem is transformed into the solution of linear equation set, and the formula is as follows:

$$\begin{bmatrix} 0 & I^T \\ I & A \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix} \tag{7}$$

where $I = [1, ..., 1]^T$, $\alpha = [\alpha_1, \alpha_2, ..., \alpha_l]^T$, $y = [y_1, y_2, ..., y_l]$ LSSVM regression prediction model can be expressed as:

$$f(x) = \sum_{i=1}^{N} \alpha_i K(x, x_i) + b \tag{8}$$

where K is a kernel function. Different support vector machines can be constructed by using a different kernel function. There are many kinds of kernel functions meeting the Mercer condition, typically including radial base kernel function, and the function in question is $K(x, x_i) = \exp\{-x - x_i^2 / 2\sigma^2\}$, and a linear kernel function is $K(x, x_i) = x \cdot x_i^T$.

3.3. PSR-WA-LSSVM Combined Model

Complex iron ore is full of unstable factors which influence the iron ore prices at different levels in different dimensions. Thus, iron ore prices have certain regularity on the one hand, also have strong randomness on the other hand, which is a signal formed by the superposition of multiple spectrums. As a classic prediction model, the ARIMA model can deal with linear problems well, but the iron ore price series has nonlinear characteristics, resulting in a certain decline in prediction accuracy. While the LSSVM model has strong generalization ability, a single LSSVM model can only fit the nonlinear part of the system, and the non-stationarity of data will make the prediction result not so ideal. As a result, it is a more suitable prediction method to use wavelet transformation to separate linear components and the high-frequency random component of iron ore price series, making it possible to distinguish obvious separate characteristics, as opposed to constructing the LSSVM model to conduct prediction aiming at the signal of every frequency band. The following subsections will introduce the basic principle and construction process of the PSR-WA-LSSVM iron ore price combined prediction model.

3.3.1. Wavelet Multi-Resolution Analysis

In 1988, based on a large amount of previous work, S. Malla and Y. Meyer presented the concept of multi-resolution analysis, spatially vividly describing a multi-resolution characteristic of a wavelet. Later, Mallat proposed Tower's multi-resolution analysis and quick reconstruction algorithm of signal, namely the Mallat algorithm. The algorithm is like Fourier quick transformation, quick and concise, so it is used widely [15–17].

However, there is a difference between the wavelet and Fourier. For example, we just need to map to wavelet bases which only take zero values in these regions if we are only interested in some high-frequency fragments of the signal, as shown in Figure 1. Furthermore, it is impossible to do this by a Fourier transformation, and, thus, a wavelet is also called a mathematical microscope.

To understand multi-resolution analysis more directly, it is explained with a 3-layer multi-resolution decomposition.

As shown in Figure 2, multi-resolution analysis is a further decomposition of low-frequency component, where S is the original series, A_3 is the low-frequency signal of third-order decomposition dimension, while D_3, D_2, D_1 are high-frequency signals with the decomposition of 3, 2, 1 respectively, S signal can be obtained from A_3, D_3, D_2, D_1 through reconstruction of the Mallat algorithm.

It shall be noted that the wavelet base and the number of decomposition layers shall be selected according to specific signal variation. When the signal fluctuates strongly, sparse sampling points will result in distortion during the restoration of the original signal.

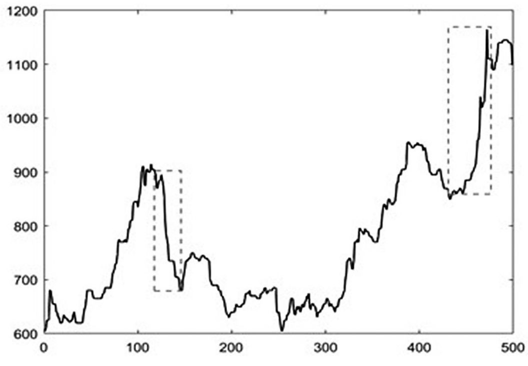

Figure 1. Schematic diagram of the difference between wavelet and Fourier change.

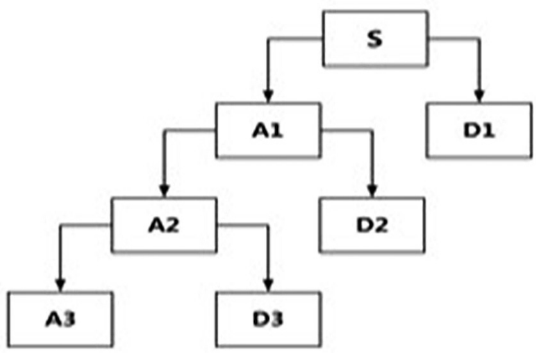

Figure 2. Wavelet transformation multi-resolution analysis structure diagram.

3.3.2. Parameter Optimization

Phase space reconstruction theory (PSR) was proposed by Takens *et al.* It uses the time series method to extract the one-dimensional feature information of the original time series into the high-dimensional space, increase the feature dimension, and maximize the information contained in the sequence. The key is to determine the optimal embedding dimension m and time delay τ [18].

It can be known from Kolmogorov's theorem that any time series can be viewed as an input/output system decided by a nonlinear mechanism. Suppose that $x(t)$ is known, then for prediction $x(t+1)$, the following mapping can be established:

$$f: R_m \to R : \hat{x}_{t+1} = f(x_{t-(m-1)\tau}, \ldots, x_{t-\tau}, x_t) \quad (9)$$

where m is the nest dimension, and τ is the time delay. Meanwhile, according to LSSVM optimization regression theory, it can be known that kernel function parameter σ^2 and regularity parameter γ are important parameters of the model, having a great influence on prediction accuracy. So, it is significant for increasing the accuracy of the prediction model to select suitable nest dimension m, time delay τ, kernel function parameter σ^2, and regularity parameter γ.

During the time series prediction, many methods can be used to select an interval of delay time τ and nest dimension m. There is no strict theoretical basis yet, so adjustment is often made according to the prediction effect, and those adjustments with optimum prediction effect are selected. Methods that separately solve the interval of delay time τ and nest dimension m include autocorrelation function method, average Mutual Information method, Virtual adjacent point method, method of reconstruction development, and others [19]. In addition to methods that separately solve phase space parameters, some methods use optimization algorithms such as Fruit Fly Optimization Algorithm, genetic algorithms to jointly optimize interval of delay time τ, nest dimension m LSSVM kernel function parameter σ^2, and regularity parameter γ. However, these methods often need a long computing time and easily fall into a locally optimal solution.

The paper adopts the method that jointly optimizes interval of delay time τ, nest dimension m, and support vector machine with optimal kernel function parameter of least squares σ^2 and regularity parameter γ. To avoid such problems as requiring a long computation time, and selecting the local optimal solution, the paper also adopts the enumeration method to give the value to internal of delay time τ and nest dimension m. It also uses the tunelssvm function to provide kernel function parameter σ^2 as well as regularity parameter and uses a tenfold crossing test. This paper uses RMSE to evaluate the prediction error of the model and selects τ, m, σ^2, γ according to error. The test shows that the selected method is a relatively effective parameter optimizing method.

3.3.3. Build Model

By performing the steps described in previous subsections, a phase space reconstruction, wavelet transformation, and LSSVM combined prediction model is built. The model architecture is shown in Figure 3.

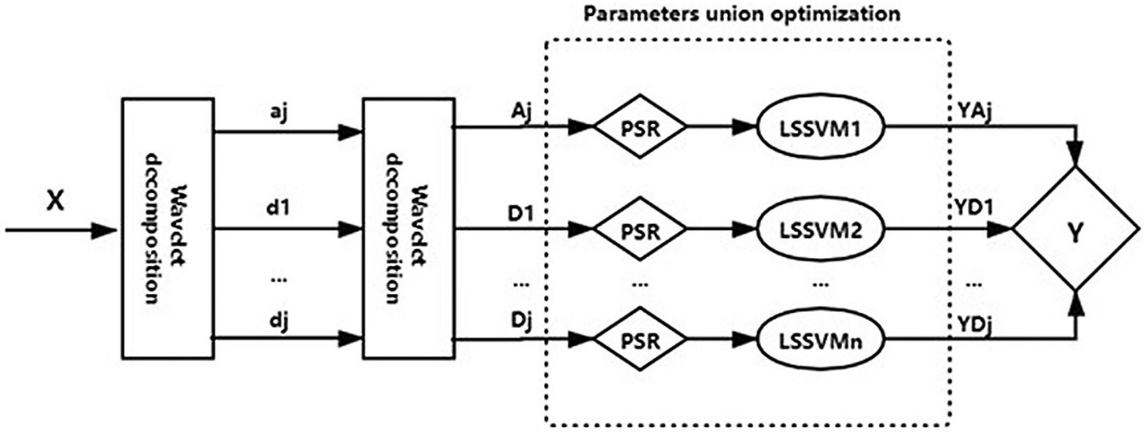

Figure 3. Architecture diagram of the PSR-WA-LSSVM prediction model.

Here $\{X(t), t = 1, 2, ..., n\}$ is iron ore spot price at the port at different times, J is decomposition dimension, $a_J(t)$ is J dimension low-frequency signal series after wavelet decomposition, $d_1(t), d_2(t), d_3(t), ..., d_J(t)$ is the series obtained by the high-frequency signal of every dimension single-branch reconstructing signal series of every dimension respectively. Parameters $A_J(t), D_1(t), D_2(t), D_3(t), ..., D_J(t)$ are in relation as follows: $\hat{X} = A_J(t) + D_1(t) + D_2(t) + ... + D_J(t)$, having in mind that there is a reconstruction error between $\hat{X}$ and original series X. The RMSE indicator is used to express the magnitude of reconstruction error Δ, build an LSSVM model aiming at every series to obtain the predicted value of every dimension $Y_{A_J}, Y_{D_1}, Y_{D_2}, ..., Y_{D_J}$. The final predicted value is their sum:

$$Y = Y_{A_J} + Y_{D_1} + Y_{D_2} + ... + Y_{D_J}. \qquad (10)$$

4. Model Verification

The paper selects 61.5% PB powder spot price index (sample size is 500) at Ningbo Zhoushan Port from January 30, 2019, to February 1, 2021, as a sample of simulation. The original price series is shown in Figure 4. The data comes from the benefit prediction information system of Xingang.

As shown in the figure, the sequence has strong volatility. This is because the price of iron ore is strongly affected by external factors. In fact, due to the impact of COVID-19 and various countries' fiscal policies, iron ore prices have risen rapidly in recent years. Through the stationarity test, it is found that the series is not stationary. After the first-order difference, the sequence can reach a stationary state. At the same time, through the period-power spectrogram and autocorrelation diagram, it can be found that the sequence does not have periodicity. In the following experiments, the data is normalized to [0, 1] to improve the speed and accuracy of prediction.

4.1. ARIMA Model

The autocorrelation and partial autocorrelation function diagram of the original series are shown in Figure 6. According to Section 3.1, the ARIMA model is built for the original series to obtain optimal model ARIMA (1, 1, 1). Prediction is made based on the model sings one step forward static prediction, namely, after data is predicted once, the predicted value is replaced by the real value in estimation interval to conduct the next prediction till the prediction process is over. Unless otherwise stated, the following models can all make a prediction using the method.

4.2. LSSVM Model

Before building a model, the original series is processed by normalization. Secondly, RBF is used as a kernel function to optimize parameters of the support vector machine based on the algorithms shown in Section 3.2. The First 400 data points are used as a training set, 50 data points as the validation set, and the remaining

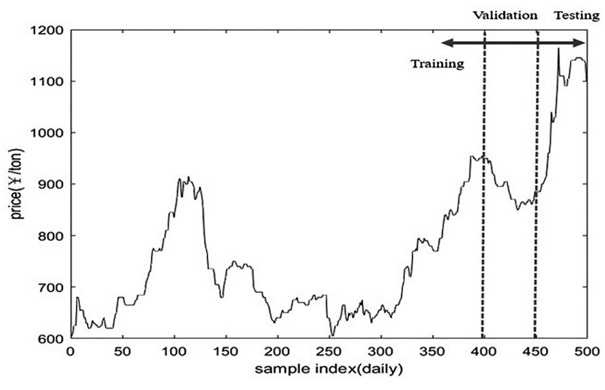

Figure 4. Time sequence diagram of the spot price of 61.5%PB silt in Ningbo Zhoushan port, 2019.1.30-2021.2.1.

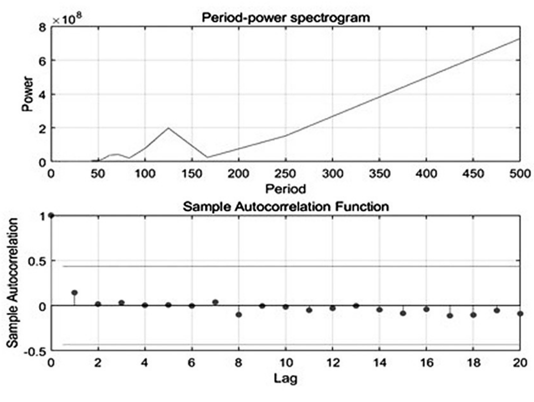

Figure 5. Period-power spectrogram and Autocorrelogram.

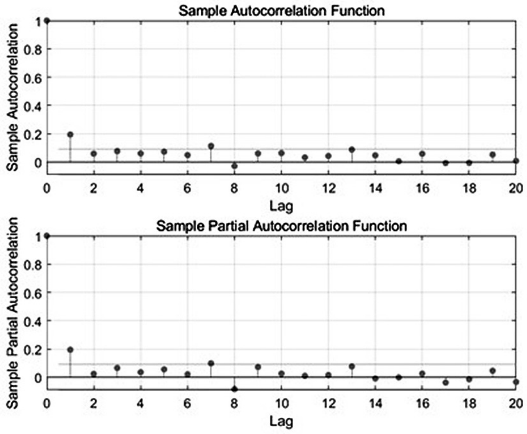

Figure 6. Diagrams of the autocorrelation function, partial autocorrelation function.

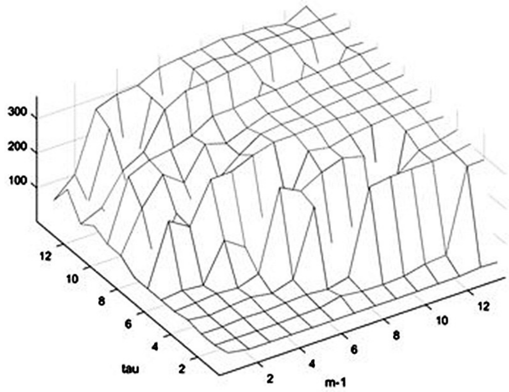

Figure 7. RMSE distribution when τ , m values vary.

50 data points as the test set. We take τ = [1, 2, 3, ..., 15], m = [2, 3, 4, ..., 15] respectively to predict the validation set. To reduce accidental error, every pair of parameter combinations is tested 10 times and then the minimal RMSE value is outputted. The corresponding interval of delay time and nest dimension m are optimal parameters, as shown in Figure 7. Here the selected optimal parameter is $\tau = 7$, $m = 2$.

We set parameters $\tau = 7$, $m = 2$, spatially reconstruct training set data and build support vector machine model, adopt tenfold crossing test, and finally determine σ^2 and φ. In this case σ^2=170105.133489448,γ=36933.6870006593. We use the stated parameters to build a prediction model used to predict the test set, and reversely normalize the predicted data to get the final result.

4.3. PSR-WA-LSSVM Model

During actual production, single-step prediction can trace and predict price variation in real-time on a daily basis, with higher timeliness and accuracy. On the other hand, multi-step prediction aims to predict the price of a period of multiple days, which has more significance for steel enterprises when it comes to grasping the trend to adjust the operating strategies. Thus, the uses of the PSR-WA-LSSVM model are dual and allow to make single-step and multi-step predictions.

4.3.1. Single-Step Prediction

Wavelet transformation is conducted on the iron ore price series first during wavelet decomposition and noise elimination. It is very important to select the appropriate wavelet base and the number of decomposition layers, which decides reconstruction error Δ to a large extent. This paper adopts urgent supported Daubechies wavelet series, Coiflets wavelet series, Symlets wavelet series, Biorthogonal wavelet series to conduct noise elimination processing for historical spot price data of iron powder, then calculates RMSE indicator to evaluate the effect of noise elimination. Smaller RMSE values represent the better effect of noise elimination. In the test, the number of wavelet decomposition layers is 5 so the db4 wavelet with the best effect of noise elimination is obtained. Then we use the wavelet to conduct tests of the different number of decomposition layers and evaluate with RMSE indicator.

Through the above test, it is possible to select db 4 wavelet to conduct 3 layers of wavelet decomposition to test the data, use ddencmp function to produce default threshold value, and use soft threshold value to conduct global noise elimination of the series. Then, single-branch reconstruction of every frequency band is performed, and finally A_3, D_3, D_2, D_1, are obtained as shown in Figure 8.

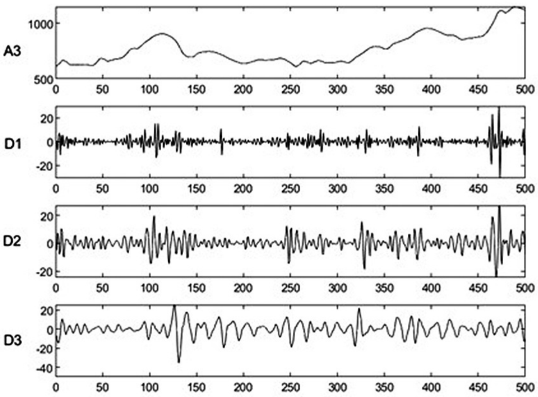

Figure 8. Decomposition diagram of wavelet.

The relation between these values is: $\hat{A} = A_3 + D_3 + D_2 + D_1$. $\hat{A}$ is the series after wavelet decomposition, noise elimination, the error of reconstruction with original series A is $\Delta = 6.5741^{-11}$. After the signal of every frequency is obtained, using the LSSVM prediction method mentioned in Section 2.3 the result of predicting to predict A_3, D_3, D_2, D_1 is shown in Figure 9.

It can be seen in Figure 9 that after wavelet noise elimination, decomposition, the LSSVM model has a good prediction score for each series.

4.3.2. Multi-Step Prediction

According to the PSR-WA-LSSVM model dynamic five-step prediction of iron ore spot price series (calculated on a working day, that is, the price fluctuation forecast of the "next week") is conducted, namely, after data is predicted once, the next step of prediction continues using predicted value within the estimation interval. Having the interval of 5 data points as a dynamic prediction period, there are 10 periods and a total of 50 data points.

First, according to the method in Section 2.3.1, wavelet decomposition, noise elimination of the original series is conducted. Within the first dynamic prediction period, with the first 445 data points as the training set, the latter 5 data points as the validation set, and 5 data points after the validation set as test set, first, we predict validation set data using one step forward static prediction method and take the parameter associated with the minimal prediction error RMSE m, τ, σ^2, γ as an optimal parameter. Finally, an optimal parameter is used to conduct a dynamic 5-step prediction of the period. After prediction, test set data of the period is added to the training set, while the validation set and test set move 5 data points into the future. Then, the next period of prediction is conducted according to the above methods and repeated until the prediction process is completed. The prediction results are shown in Figure 10.

4.4. Statistical Analysis

This section presents a comparison of the predictive capacity of the proposed model to the classic ARIMA, classic LSTM, and PSR-

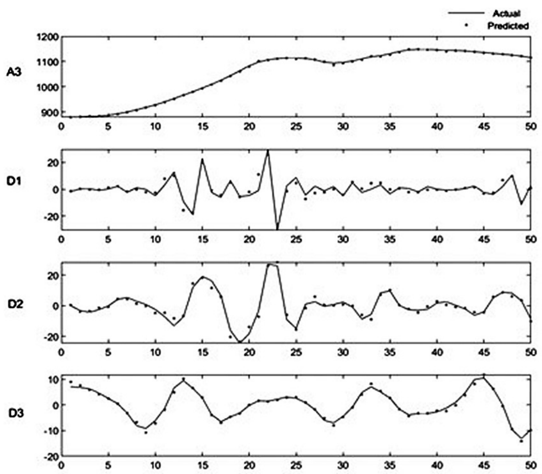

Figure 9. Single-step prediction diagram of wavelet every frequency band signal.

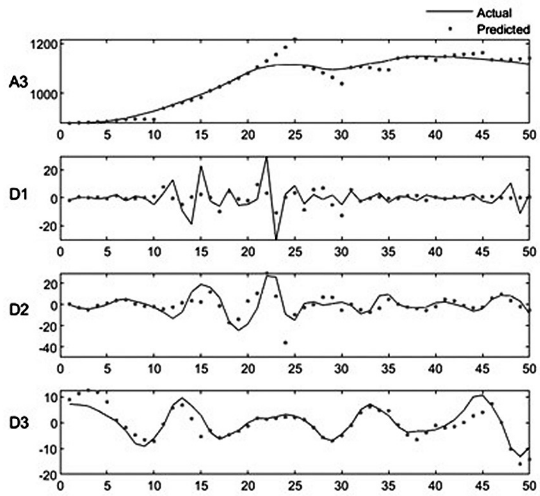

Figure 10. 5-step prediction diagram of wavelet for every frequency band signal.

Table 1. Performance comparison of single-step prediction with different models.

Models	RMSE	MAPE	D_{Rate}	$\frac{1}{3}_{Rate}$
ARIMA	18.8857	1.0024	88%	92%
LSTM	19.7376	1.1916	80%	92%
PSR-LSSVM	18.8029	0.9980	86%	92%
PSR-WA-LSSVM	4.1444	0.3003	100%	100%

LSSVM using the same data sets. The classic LSTM model parameters are set as follows: the neural network has a hidden layer with 120 neurons, the batch size is 1, whereas, the iteration optimization algorithm is set to 100. The other model's parameters have been described in the corresponding chapters. The experiment work uses MATLAB (R2020b) software. The following indicator to evaluate the result used:

1. Root mean square error (RMSE):

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(A_i - F_i\right)^2} \tag{11}$$

2. Mean absolute Percentage of Error (MAPE):

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{A_i - F_i}{A_i}\right| \tag{12}$$

3. N times standard deviation hit rate (N_{Rate}):

$$a_t = \begin{cases} 1, & \left|A_i - F_i\right| < N\,\mathrm{sqrt}\left(\mathrm{var}\left(\{A_n\}\right)\right) \\ 0, & \left|A_i - F_i\right| \geq N\,\mathrm{sqrt}\left(\mathrm{var}\left(\{A_n\}\right)\right) \end{cases} \tag{13}$$

$$N_{rate} = \frac{1}{n}\sum_{t=1}^{n} a_t \times 100\% \tag{14}$$

4. Correctness rate of prediction direction (D_{Rate}):

$$b_t = \begin{cases} 1, & \left(A_i - A_{i-1}\right)\left(F_i - A_{i-1}\right) \geq 0 \\ 0, & \left(A_i - A_{i-1}\right)\left(F_i - A_{i-1}\right) < 0 \end{cases} \tag{15}$$

$$D_{rate} = \frac{1}{n}\sum_{t=1}^{n} b_t \times 100\% \tag{16}$$

where, $\{A_n\}$ is the test set, A_i is the actual value of the i-th prediction point, F_i is the predicted value of the i-th prediction point, and n is the number of prediction points.

For specific indicator evaluation of prediction results of every model see Table 1. It can be found that firstly, RMSE of PSR-WA-LSSVM model prediction is lower by 78.1% compared with the ARIMA model, lower by 78.0% compared with PSR-LSSVM model, and lower by 79.0% compared with LSTM model respectively. Secondly, MAPE of the PSR-WA -LSSVM model is 0.3003, lower by 70.0% compared with the ARIMA model, lower by 70.0% compared with the PSR-LSSVM model, and lower by 74.8% compared with the LSTM model respectively. Thirdly, D_{Rate} of PSR-WA -LSSVM model is 100%, up 12% compared with ARIMA mode, up 14% compared with PSR-LSSVM model, and up 20% compared with the LSTM model respectively. Finally, the prediction accuracy of the PSR-WA-LSSVM model under 1/3 standard deviation is 100%, up 8% both compared with ARIMA, PSR-LSSVM, and LSTM model respectively. Multiple error indicators show that the newly constructed PSR-WA-LSSVM model has an even smaller error, model prediction performance increases significantly.

The prediction results of every model are shown in Figure 11. The results demonstrated the superiority of the PSW-WA-LSSVM model over all other models. Furthermore, the numerical results indicated that the PSR–WA-LSSVM model provided the lowest RMSE, MAPE, of 4.1444, 0.3003.

PSR-WA-LSSVM combined model keeps better prediction accuracy when the original price series is stationary. More importantly, starting from the 6th prediction point, the iron ore price series begins to fluctuate dramatically, for example, at prediction points 13-15, 20-24, 27-29, *etc.*, the original price series direction varies frequently. In contrast with the lagging of a single model, the PSR-WA-LSSVM model accurately predicts variation of original price series direction. In general, the PSR-WA-LSSVM model keeps a smaller deviation from the actual value during the whole prediction period, showing excellent price variation direction prediction ability when the original price series fluctuates dramatically, with a better comprehensive prediction effect.

To summarize, compared with the single ARIMA model, PSR-LSSVM model, and LSTM model, the PSR-WA-LSSVM model achieves a better effect during single-step prediction.

The performance of the PSR-WA-LSSVM model in the five-step prediction is shown in Figure 12. It can be concluded that the model predicts the overall trend of the original price series well. It can be found from Table 2 that the direction prediction accuracy of the PSR-WA-LSSVM model (D_{Rate}) is 86%. In general, the model shows better prediction ability in the multi-step prediction of iron ore spot price and can provide a certain reference for steel enterprises to better grasp the market trend, select investment, and operating strategy.

5. Conclusion

When the LSSVM model is used to predict nonlinear time series, values of nest dimension m, delay time τ, kernel function parameter σ^2, and regularity parameter γ significantly influence the prediction accuracy of the model. Parameters m and τ are typically no more than 15, while the variation ranges of σ^2 and γ are larger. The test indicates that a convenient method to find parameters is by using the enumeration method to give m, τ in turn, using tunelssvm function to give σ^2 and γ, and using the training set to conduct tenfold validation. The PSR-LSSVM model tends to fluctuate greatly when the iron ore price sequence changes drastically, but the wavelet can help solve this problem well. Data of 61.5% PB powder spot price index (sample size is 500) at Ningbo Zhoushan Port from January 30, 2019, to February 1, 2021, is predicted respectively by using four single-step prediction models. The test result shows that compared with the ARIMA model, LSTM model, and PSR-LSSVM model, PSR-WA-LSSVM combined model achieves better prediction results. PSR-WA-LSSVM was used to conduct a five-step prediction of data for the 61.5% PB powder spot price index, having the sample size set to 500, at the Ningbo Zhoushan Port from January 30, 2019, to February 1, 2021. The test found that PSR-WA-LSSVM combined model achieved 86% direction prediction accuracy, with a certain practical reference value.

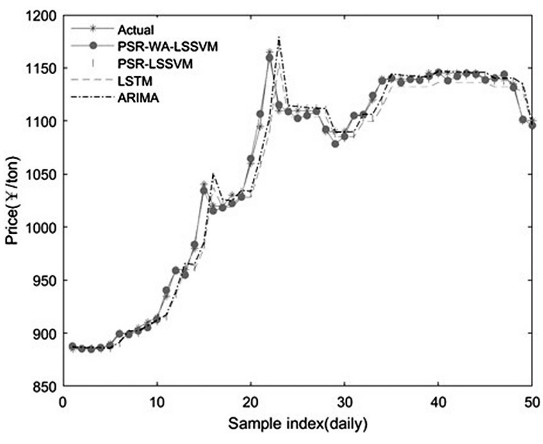

Figure 11. Contrast chart of single-step prediction effect.

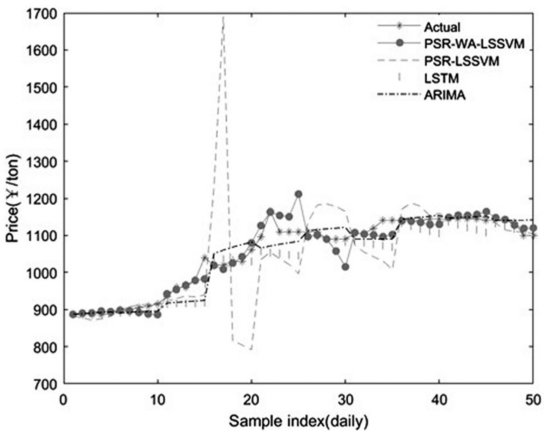

Figure 12. Contrast chart of five-step prediction effect.

Table 2. Performance comparison of five-step prediction with different models.

Models	RMSE	MAPE	D_{Rate}	$\frac{1}{3}_{Rate}$
ARIMA	32.4925	2.1733	70%	74%
LSTM	37.6885	2.4431	74%	72%
PSR-LSSVM	125.8011	6.1899	72%	56%
PSR-WA-LSSVM	25.8862	1.5528	86%	84%

Granted, this article also has some research limitations. One of the most obvious is that the consideration of iron ore price factors is not enough. It is only based on the iron ore price sequence data and does not include external economic and environmental factors. Therefore, future research can introduce domestic and international economic environment and other variables into consideration based on the above method analysis. In addition, how to choose more appropriate multi-scale analysis and deep learning methods to establish the optimal prediction model based on the characteristics of the iron ore price sequence data will also be one of the directions for continued research in the future.

References

[1] B. W. Zhu, "Prediction of Seaborne Price of Iron Ore Based on Wavelet Analysis", [D]//Dalian: Dalian Maritime University, 2009. (in Chinese) http://dx.doi.org/10.7666/d.y1548775

[2] Q. Wang *et al.*, "Prediction of Iron Ore Price in 2015 – exponential smoothing method", *China land and resources economy*, vol. 27, no. 2, pp. 51–54+61, 2014. http://dx.doi.org/10.3969/j.issn.1672-6995.2014.02.013

[3] Z. Fu, "The Mechanism of Imported Iron Ore Price in China", *Modern Economy*, vol. 9, no. 11, pp. 1908–1931, 2018. https://doi.org/10.4236/me.2018.911120

[4] J. Wang *et al.*, "A Trend Fixed on Firstly and Seasonal Adjustment Model Combined with the e-SVR for Short-Term Forecasting of Electricity Demand", *Energy Policy*, no. 37, pp. 4901–4909, 2009. https://doi.org/10.1016/j.enpol.2009.06.046

[5] P. F. PAI and C. S. LIN, "A Hybrid ARIMA and Support Vector Machines Model in Stock Price Forecasting", *Omega*, vol. 33, no. 6, pp. 497–505, 2005. https://doi.org/10.1016/j.omega.2004.07.024

[6] G. Xie and S. Y. Wang, "Hybrid Approaches Based on LSSVR Model for Container Throughput Forecasting: A Comparative Study", *Applied Soft Computing*, vol.13, no.5, pp. 2232–2241, 2013. https://doi.org/10.1016/j.asoc.2013.02.002

[7] B. Z. Zhu and Y. M. Wei, "Carbon Price Forecasting with a Novel Hybrid ARIMA and Least Squares Support Vector Machines Methodology", Omega, vol. 41, no.3, pp. 517–524, 2013. https://doi.org/10.1016/j.omega.2012.06.005

[8] K. Y. Chen and C. H. Wan,"A Hybrid SARIMA and Support Vector Machines in Forecasting the Production Values of the Machinery Industry in Taiwan", *Expert Systems With Applications*, vol.32, no.1, pp. 254–264, 2007. https://doi.org/10.1016/j.eswa.2005.11.027

[9] T. Taskaya and M. C. Casey, "A Comparative Study ofAutoregressive Neural Network Hybrids", *Neural Networks*, vol. 18, no. (5–6), pp. 781–789, 2005. https://doi.org/10.1016/j.neunet.2005.06.003

[10] L. D. S. Coelho and A. A. P. Santos, "A RBF Neural Network Model with GARCH Errors: Application to Electricity Price Forecasting", *Electric Power Systems Research*, vol. 81, no.1, pp. 74–83, 2011. https://doi.org/10.1016/j.epsr.2010.07.015

[11] M. Khashei and M. Bijari, "A Novel Hybridization of Artificial Neural Networks and ARIMA Models for Time Series Forecasting", *Applied Soft Computing*, vol. 11, no. 2, pp. 2664–2675, 2011. https://doi.org/10.1016/j.asoc.2010.10.015

[12] L. Yu *et al.*, "A Multiscale Neural Network Learning Paradigm for Financial Crisis Forecasting", Neurocomputing, vol. 73, no. (4–6), pp. 716–725, 2010. https://doi.org/10.1016/j.neucom.2008.11.035

[13] F. Omer, "A Hybrid Neural Network and ARIMA Model for Water Quality Time Series Prediction", *Engineering Applications of Artificial Intelligence*, vol. 23, no.4, pp. 586–594, 2010. https://doi.org/10.1016/j.engappai.2009.09.015

[14] H. A. Cagdas *et al.*, "Forecasting Nonlinear Time Series with a Hybrid Methodology", *Applied Mathematics Letters*, vol. 22, no. 9, pp. 1467–1470, 2009. https://doi.org/10.1016/j.aml.2009.02.006

[15] J. Sheng *et al.*, "Determination of Weak Knock Characteristics for Two-Stroke Spark Ignition UAV Engines Based on Mallat Decomposition Algorithm", *Mathematical Problems in Engineering*, 2021. https://doi.org/10.1155/2021/1250327

[16] H. Zhao and N, Ansari, "Wavelet Transform-Based Network Traffic Prediction: A Fast On-line Approach", *Journal of Computing and Information Technology*, vol. 20, no. 1, pp. 15–25, 2012. https://doi.org/10.2498/cit.1001989

[17] A. Ali *et al.*, "Affine Invariant Contour Descriptors Using Independent Component Analysis and Dyadic Wavelet Transform", *Journal of Computing and Information Technology*, vol. 16, no. 3, pp. 169–181, 2008. https://doi.org/10.2498/cit.1001009

[18] X. Wang *et al.*, "Research on Short-term Load Forecasting of Units Based on PSR-LSTM", *Journal of Engineering for Thermal Energy and Power*, vol. 35, no. 5, pp. 1–7, 2021. (in Chinese) http://dx.doi.org/10.16146/j.cnki.rndlgc.2021.05.011

[19] S. Luo *et al.*, "Blast Furnace System Modeling by Multivariate Phase Space Reconstruction and Neural Networks", *Asian Journal of Control*, vol. 15, no. 2, pp. 553–561, 2013. https://doi.org/10.1002/asjc.574

Received: February 2021
Revised: October 2021
Accepted: October 2021

Contact addresses:

Xiangjun Cai
Zhejiang college of Security Technology
Wenzhou
China
e-mail: xiangjun_cai@163.com

Shihua Luo
School of Statistics
Jiangxi University of Finance and Economics
Nanchang
China
e-mail: luoshihua@jxufe.edu.cn

Xiangjun Cai received his BSc degree in mathematics from Quzhou University, China, in 2016 and the MSc degree in Applied Statistics from Jiangxi University of Finance and Economics, China, in 2018. He is currently a full-time teacher at the Zhejiang College of Security Technology. His research interests include modeling and data analysis.

Shihua Luo received the BSc degree and MSc degree in mathematics from Jiangxi Normal University, China, in 1998 and 2001, respectively, and the PhD degree in operational research and cybernetics from Zhejiang University, China, in 2007. He is currently a Professor at the Jiangxi University of Finance and Economics. His research interests include modeling and optimization of complex systems, especially on the extraction of nonlinear characteristics of chemical and financial processes.