Advances in Social Science, Education and Humanities Research, volume 177
International Conference on Education Reform and Management Science (ERMS 2018)

Material Demand Combination Forecasting Model Based on EMD-PSO-LSSVR

Hua MO[1], Lin XIONG[2,*] and Ruo-yu LU[1]

[1]School of Management and Economics, University of Electronic Science and Technology of China, Chengdu, 610054, China

[2]Institute of Electronic and Information Engineering of Guangdong, University of Electronic Science and Technology of China, 523808, Dongguan

*Corresponding author

Keywords: EMD, IMF, LSSVR.

Abstract: The time series data of material demand of manufacturing companies are often non-stationary. Paper uses empirical mode decomposition (EMD) to convert non-stationary time series into a series of intrinsic mode function (IMF) and a residual term (RES), and then digged out more information combined with least squares support vector machine regression (LSSVR) to forecast the model. Finally, the empirical results show that the EMD-LSSVR combination forecast can effectively predict non-stationary material demand time series, and the prediction accuracy is high. It has a certain degree of promotion and practical value.

Introduction

In recent years, how to accurately predict changes in demand information to avoid the generation of bullwhip effect has become a hot topic in academia. At the same time, customers demand for timely response to the urgent attention of the market environment. Companies improve their ability to forecast material requirements that is a powerful guarantee for them to gain competitive advantages. In a customer-oriented market environment, on the one hand companies need to respond quickly to customer needs; on the other hand, companies also want to be able to limit output and minimize procurement and inventory costs [1]. However, due to the existence of bullwhip effect in the supply chain system, changes in demand will increase as the magnitude of the supply chain level increases, resulting in problems such as the lack of timely response of the demand or the surge of inventory. Therefore, for each enterprise in the supply chain, it is of great significance to accurately predict the actual demand of materials.

With regard to research on demand forecasting methods, domestic and foreign scholars have obtained certain research results and classified the forecasting methods into qualitative analysis methods and quantitative analysis methods. Qualitative analysis mainly includes research methods, judgment analysis methods, Delphi expert survey methods and customer comprehensive judgment method; Quantitative analysis method is mainly based on historical data using modern mathematical methods to analyze and process historical data by establishing corresponding forecasting models to analyze, which mainly including arithmetic average method, moving weight method, exponential smoothing method, regression analysis method, grey prediction method and so on. For the manufacturing industry, the material demand is affected by multi-dimensional factors. For the general forecasting model, it is necessary to know in advance the model of the influencing factor. This model mainly depends on subjective analysis and feature discrimination, so the model is not ideal [2].

In 1998[3], Dr. Huang E proposed a data-based adaptive prediction method called Empirical Mode Decomposition (EMD). This method can reveal the hidden patterns and trends of time series, simplify the prediction task into several simple prediction subtasks, and can be applied to various complex prediction problems. The advantage of the EMD method is that it is well-suited for decomposing nonlinear and non-stationary time series, and performs better than wavelet and Fourier decomposition in describing local time instantaneous frequencies.

Initially used in the engineering field of EMD in recent years gradually been used in time series analysis. Xun Zhang, K.K. Lai, and Shou-Yang Wang [4] (2008) used the white noise-added EMD method (EEMD) to analyze the spot price of West Texas crude oil in the United States. Wang Wenbo, Fei Pusheng, and Yi Xuming [5] (2010) combined EMD and neural networks to forecast the Chinese stock market. Liu Haifei, Li Xindan [6] (2011) forecast stock prices based on EMD method and wavelet analysis. Zhu Jianping, Zhang Nanxi, and Zhu Wanchuang [7] (2013) used a combination of EMD and neural network methods to predict gold prices. Gong Chengzhu, Li Lanlan, Yang Juan, Zhu Kejun[8] (2014) used EMD-PSR-LSSVM to predict the short-term load of China's urban gas pipeline network.

Due to multi-dimensional factors, material requirements often exhibit non-stationary and non-linear time series characteristics. This paper intends to combine EMD (empirical mode decomposition) and LSSVR (least squares support vector regression) for material demand forecasting.

EMD-LSSVR Combination Forecasting Model

Empirical Mode Decomposition

EMD is an adaptive data processing method based on the latent features of data. Its essence is to obtain the intrinsic wave pattern through the characteristic time scale of the data, decompose the complex time series into a finite number of eigenmode functions and residuals that can be directly analyzed. Because of its superiority in dealing with non-stationary, non-linear data, this method can be applied to the decomposition of any type of time series (signals), and has more obvious advantages than the previous processing methods. The decomposition process of the EMD method is as follows:

(1)Determine the local extrema of the data series x(t). Find all the local maxima points of the data and use a cubic spline function to fit the upper envelope of the original data sequence U(t). Similarly, find all the local minima points of the data sequence to form a packet drop line under the data sequence L(t). Calculate the average of the upper and lower envelopes m1(t) as follows:

$$\mathrm{m}_1(t) = \frac{[U(t) + L(t)]}{2} \tag{1}$$

Then it $\mathrm{m}_1(t)$ will be removed from the original sequence $\mathrm{x}(t)$ to form a new sequence $\mathrm{h}_{11}(t)$. If it satisfies the symmetry and the local maxima are all positive, all local minima are negative, then the resulting component $\mathrm{h}_1(t)$ is the IMF. Otherwise, Otherwise, replace $\mathrm{x}(t)$ by $\mathrm{h}_{11}(t)$ and repeat the above process until you find the IMF that meets the requirements. The formula is as follows:

$$h_{12}(t) = h_{11}(t) - m_{12}(t) \tag{2}$$

(2)The time series h1n(t) that meets the requirements is IMF1 and is defined as $\mathrm{c}_1(t)$, and subtract $\mathrm{c}_1(t)$ with the original sequence $\mathrm{x}(t)$ to get $\mathrm{r}_1(t)$ as follows:

$$\mathrm{r}_1(t) = x(t) - c_1(t) \tag{3}$$

Take $\mathrm{r}_1(t)$ as the original sequence and repeat the above steps until it rn(t) can no longer be decomposed. The final raw time series is broken down into:

$$x(t) = \sum_{i=1}^{n} c_i(t) + r(t) \tag{4}$$

Where $\mathrm{x}(t)$ represents the original time series of data, $\mathrm{c}_i(t)$ represents each IMF value, n is the number of IMF functions, and $\mathrm{r}(t)$ represents the residual. The use of EMD requires the following assumptions:

(1)At least one maximum and one minimum value of data

(2)The local temporal characteristics of the data are uniquely determined by the time scale between the extreme points;

(3)If the data has no extremum points but inflection points, the extremum can be obtained by performing one or more differentiations on the data, and then the decomposition results can be obtained through integration.

The material demand time series conforms to these assumptions and can therefore be decomposed using the EMD method.

Least Square Support Vector Regression

EMD decomposed time series contains several imf sequences and one residual item. Each of the decomposed sequences has its own characteristics and can be further analyzed based on these feature sequences. Xun Zhang, KK Lai, and Shou-Yang Wang [6] will perform empirical mode decomposition after white noise is added to the original sequence. Finally, according to the average of different IMFs, the result is synthesized into two different CIMFs (high-precision imfs) with high frequency and low frequency. . The idea of building a combined forecasting model in this paper is as follows: First, the EMD decomposition results are clustered and superposed to obtain two parts: low frequency and high frequency. The low-frequency part shows the part of the mutation under the disturbance, and the fitting function is constructed according to the low-frequency form; the high-frequency shows the change of the trend, and the support vector regression is used to predict the fluctuation of the time series around the trend based on the high frequency part, and then the two parts are combined. The precision of the model after combination is high and can predict future demand.

(1) EMD-based time series trend forecast

The EMD decomposition results are superimposed by clustering to obtain low frequency and high frequency components. Low frequency describes the part of the mutation under the disturbance. In actual production, the demand for critical parts may increase rapidly with the urgent insertion of orders. In order to increase the forecasting accuracy of key parts demand, it is necessary to consider the influence of disturbances. This paper uses Origin software to fit the low-frequency part with a function with similar morphology. The low frequency part and the probability density distribution function are very similar. This paper uses Extreme function to fit.

The result obtained after the fitting is a function of time, which can predict the future value, but this prediction error is larger because only the mutation factor is considered and the long-term trend is not taken into account. The following specific examples will be combined with the prediction model for further expansion.

(2) Time Series Fluctuation Prediction Based on LSSVR

Least Square Support Vector Regression (LSSVR) is a nonlinear regression prediction method based on kernel functions. It can find the best regression hyperplane with the lowest structural risk in the high-dimensional feature space. Least-squares support vector regression prediction algorithm is characterized by strong adaptability to the sample, and full mining of potential features of the data. Compared with other support vector regression algorithms, it has a small demand for sample size and simple operation.

In a given sample space set $S = \{(x_i, y_i) \mid x_i \in R^n, y_i \in R, i = 1, \cdots, N\}$, x_i is the input value of the function, and generate the corresponding value of the high-latitude space function through the non-linear mapping, that is, the output value of the function y_i. The optimization problem is as follows:

$$\text{Min} \quad J(\omega, e) = \frac{1}{2}\omega^T\omega + c\sum_{i=1}^{l} e_i^2 \tag{5}$$

$$\text{s.t.} \quad y_i = \omega^T\varphi(x_i) + b + e_i \tag{6}$$

Where C is the penalty coefficient, $\varphi(x_i)$ is the nonlinear mapping function, e_i is the error term, and i is the number of error terms. The optimization problem was not solved and the Lagrangian function was introduced and defined as follows:

$$L := J(\omega, e) - \sum_{i=1}^{l} a_i(\omega^T \varphi(x_i) + b + e_i - y_i) \tag{7}$$

According to the Karush-Kuhn-Tucker condition, the solution is:

$$\begin{cases} \frac{\partial L}{\partial W} = 0 \rightarrow w = \sum_{i=1}^{l} a_i \varphi(x_i) \\ \frac{\partial L}{\partial b} = 0 \rightarrow -\sum_{i=1}^{l} a_i = 0 \\ \frac{\partial L}{\partial e_i} = 0 \rightarrow a_i = Ce_i \\ \frac{\partial L}{\partial a_i} = 0 \rightarrow \omega^T \varphi(x_i) + b + e_i - y_i = 0 \\ (i = 1,2 \cdots, l) \end{cases} \tag{8}$$

After partial derivatives of w, b, e_i, a_i respectively, eliminated e_i and w, the formula is transformed into:

$$\begin{bmatrix} 0 & E^T \\ E & \varphi(x_i)^T \varphi(x_i) + c^{-1} I \end{bmatrix} \begin{bmatrix} b \\ a \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix} \tag{9}$$

Sovled into:

$$f(x) = \sum_{i=1}^{l} a_i \psi(x, x_i) + b \tag{10}$$

Among them, $\psi(x, x_i)$ is a kernel function. This paper selects the Gaussian function with high fitting accuracy and suitable for small samples as the kernel function.

$$K(x, x_i) = \exp\left\{ - \frac{\|x - x_i\|^2}{2\sigma^2} \right\} \tag{11}$$

Where: Q is the width parameter of the Gaussian kernel function. It implicitly defines the nonlinear mapping from the demand function input space to the high-dimensional feature space to control the complexity of the final solution.

(3) In the first stage, for different types of parameters, the raw material demand data is subjected to empirical mode decomposition to obtain several IMFs and one residual trend item Res. Then use the Hilbert Yellow Transform (HHT) to find the instantaneous frequency of each IMF. Then use the squared Euclidean distance as the discriminant distance of the IMF, and select the representative average values, standard deviations, skewness, mean square values, autocorrelation, accumulation for data cleaning. The IMFs and Res as training samples and establish LSSVR for prediction.

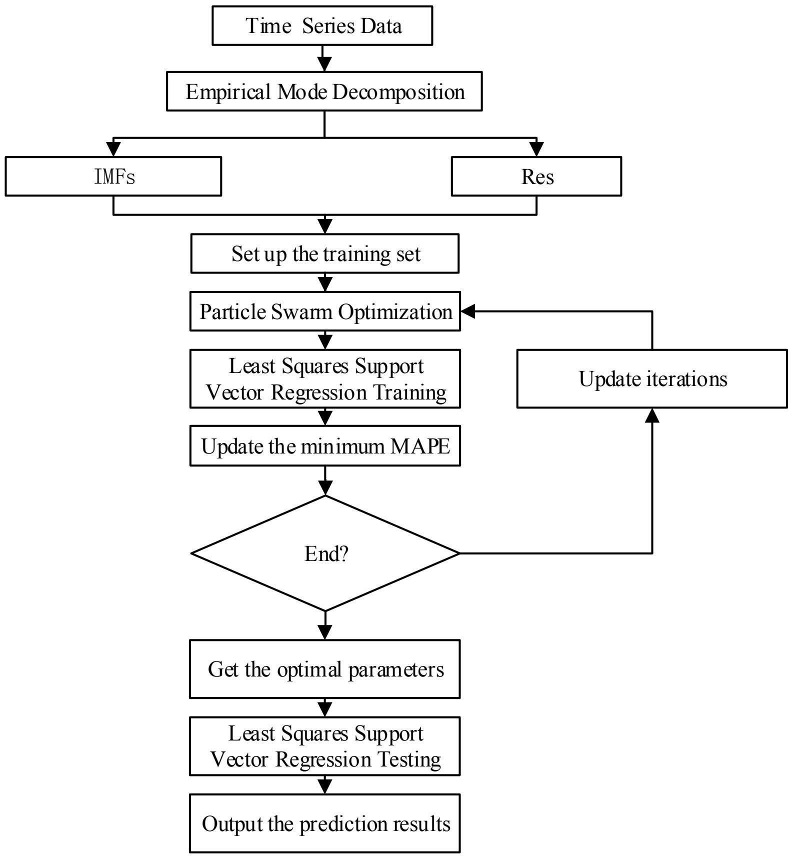

Figure 1. Combination forecasting method

In the second stage, the three training samples LSSVR obtained in the above steps are input to obtain the optimal parameters, thereby obtaining the logistics demand forecast results. Considering that the key problem of the least squares support phasor regression prediction model is the parameter determination, this paper uses the PSO algorithm to optimize the LSSVR material demand forecasting model. The specific steps are as follows:

(1) Enter sample data. The original material demand data is subjected to EMD decomposition to obtain three training sample sets.

(2) Initialize the parameters of the particle swarm algorithm. The penalty coefficient C and Q in the Gaussian kernel function are mapped to particles for initialization.

(3) Select some data from the material demand sample set as the verification set, and other as the training set.

(4) Evaluate the fitness of the particle and compare it with the optimal value and decide whether to update the optimal value.

(5) Update iterative particle velocity and position to generate new population

(6) Check whether the termination condition is satisfied. If it is satisfied, the optimal particle is defined as the width parameter Q and the penalty coefficient C to obtain the forecast model of the material demand; otherwise, it returns (3) until the parameter that meets the fitness threshold is found or reached the maximum number of iterations.

(7) Enter the test set into the material demand forecasting model optimized by parameters, and finally obtain the material demand value in the time period to be measured.

Material Requirements combination forecasting model based on EMD-LSSVR is based on time series prediction based on past potential by acquiring a regular pattern of time series, thus extrapolated to predict the sequence. Under normal circumstances, the time series of material requirements in an enterprise is affected by multi-dimensional factors and it is difficult to directly obtain its regular pattern.

The methodology of time series forecasting is to obtain information through historical data, and then obtain certain rules through information to predict and analyze. In general, there are many factors that affect the time series, especially the material demand time series. If you want to analyze them one by one, the workload will be very large, and some factors are more complex and difficult to analyze, so it is difficult to extract features directly from the time series.

The idea of the combined forecasting model proposed in this paper is similar to that of EMD. It is to screen the time series and regularize it to make the time series more regular. Combinatorial forecasting has some problems in practical applications. On the one hand, it needs the span of time series to be long enough, so that least squares support vector regression has enough sample training patterns. On the other hand, the combined forecasting model is suitable for more complex time series forecasting, and the regularly-defined time series can be directly predicted, and there is no need to use combined forecasting. This paper proposes that the combined forecasting model has certain pertinence and is more suitable for the time series with predictable time series, and the rules are difficult to mine, such as material demand. If the data does not fluctuate and there are no peaks under disturbance, there is no point in the low-frequency part based on the trend predicted by EMD, and there is no need to use the EMD-LSSVR combination forecasting model.

Material Demand Forecast EMD-LSSVR Combination Forecasting Method

Data Sources

This article selects the demand data series of A-type UV sensors for key components of A financial equipment manufacturing enterprises from January 2011 to December 2014 as a sample of the study. Because the material requirements of the industry are affected by orders, the data series will show periodicity. .The material demand data for the A-type UV sensor comes from the company's ERP database. The first 70% of the data is used as a training set, and the last 30% of the data is used as a test set. In order to evaluate the performance of the combined forecasting model, the three indicators of goodness of fit R2 and mean squared error MSE are compared with ARIMA and BP neural network methods. This shows the effectiveness of the EMD-LSSVR combination forecasting method in material demand forecasting applications.

EMD Results and Synthesis Analysis

According to the EMD algorithm, the earlier the IMF frequency is, the higher the volatility is, the more the short-term characteristics of the time series can be reflected, and the final residual item reflects the long-term trend of the time series. After the material demand sequence is decomposed by EMD, four IMF components and one residual item are obtained. The decomposition results are shown in Fig. 2. Among them, the IMF1, IMF2, IMF3, and IMF4 are arranged in descending order of frequency, and each IMF component curve exhibits an oscillating form that is substantially symmetrical around the zero-mean line, the local maximum value, and the local minimum value. It can be found that as the frequency decreases, the regularity of the sequence becomes more and more obvious. Finally, the trend of the residual item is seen to increase for a long time, reflecting the long-term trend of material demand is gradually rising.

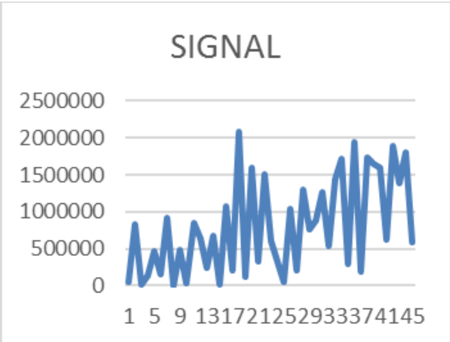

Figure 2. Original sequence line chart

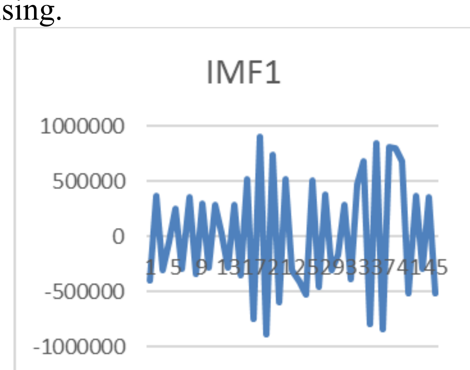

Figure 3. IMF1 sequence line chart

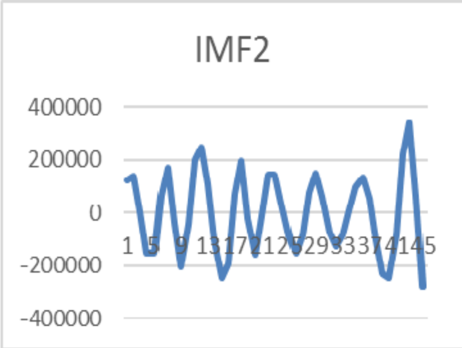

Figure 4. IMF2 sequence line chart

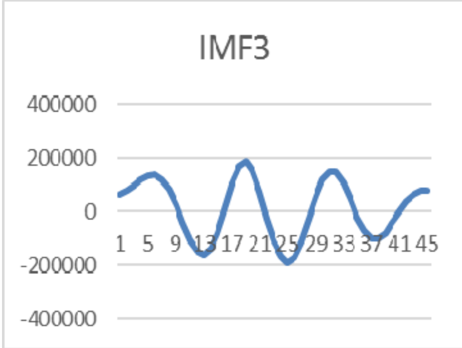

Figure 5. IMF3 sequence line chart

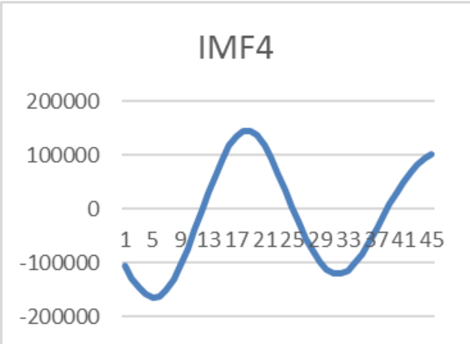

Figure 6. IMF4 sequence line chart

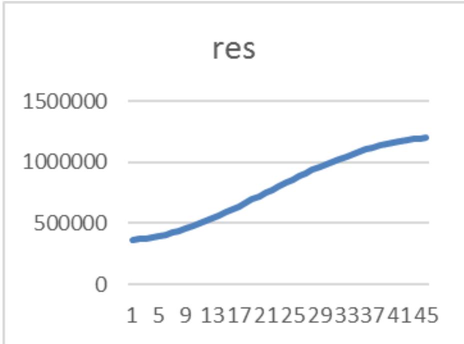

Figure 7. Residual series line chart

Descriptive statistics are made for each IMF and residual item, and representative indicators such as average value, standard deviation, skewness, mean square value, autocorrelation, and accumulativeness in the descriptive statistical results are selected for analysis, as shown in the following table1.

Table 1. Descriptive statistics

Sequence	Average value	Variance	Kurtosis	Skewness	Minimum	Maximum	Confidence (95.0%)
IMF1	38555.12	2.65705E+11	-1.22989	-0.01683	-884989	903685.8	149675.7
IMF2	-21872.5	28641708392	-0.57758	-0.15961	-388286	343401.4	49141.77
IMF3	16519.11	11347908224	-1.08114	-0.36651	-189236	190285.8	30932.09
IMF4	-9728.19	10094393448	-1.46551	0.011919	-163856	145880.1	29173.7
res	819696.8	89275585611	-1.49584	-0.14536	363868.8	1211263	86759.63
SIGNAL	843170.3	4.11245E+11	-1.25173	0.319226	6000	2082640	186209.3

In the LSSVR model with the radial basis function as the kernel function, the factors influencing the performance of the LSSVR model mainly include the values of the regularization parameter C and the kernel function parameter Q. To improve the learning and generalization ability of LSSVR, regularization parameters C and kernel function parameters must be optimized. In order to obtain the best recognition rate, the multi-parameter parallel optimization of particle swarm optimization is used to optimize the kernel function parameter Q and regularization parameter C of LSSVR. By optimizing the parameters of the particle swarm optimization algorithm, the initial parameters of LSSVR are optimized, and the prediction efficiency and prediction accuracy of LSSVR under small sample conditions are improved. This paper takes the particle swarm algorithm acceleration factors C1 and C2 to take the empirical value C1=C2=1.5, the population size is set to 20, the initial population is randomly generated, the search termination condition is to reach the maximum number of iterations 100, inertia weight Wmax=0.9 Wmin=0.4.

Based on this method, by using least square support vector machine, the high frequency part, the low frequency part and the residual term are respectively regressed and predicted, and the particle swarm optimization algorithm PSO is used to optimize the regression parameters. The first 45 items of the selected data are training sets, and the last 3 items of data are verification sets. The results are as follows:

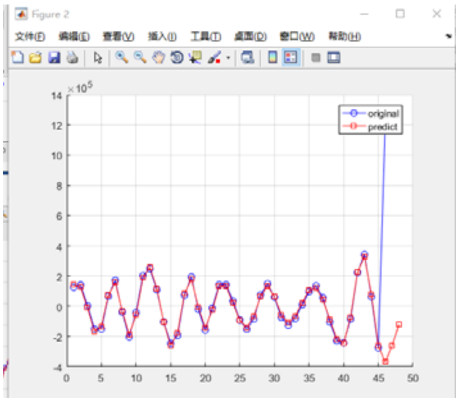

Figure 8. IMF1-PSO-LSSVR fitting results

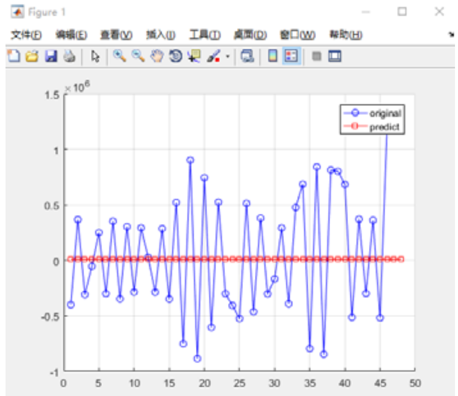

Figure 9. IMF2-PSO-LSSVR fitting results

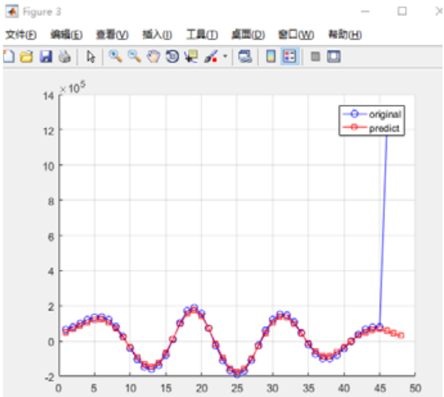

Figure 10. IMF3-PSO-LSSVR fitting results

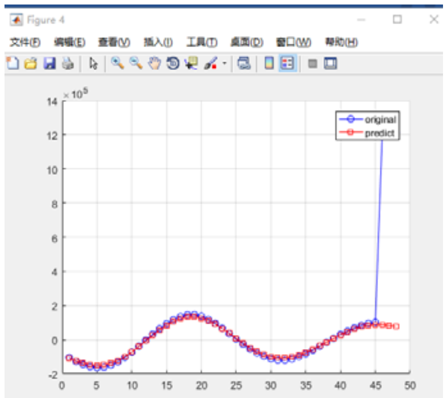

Figure 11. IMF4-PSO-LSSVR fitting results

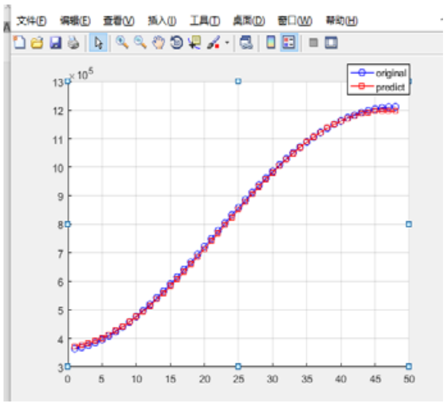

Figure 12. RES-PSO-LSSVR fitting results

From the above processing results, the EMD-LSSVR model after particle swarm optimization has a good fitting effect. From the long-term trend, the demand for materials becomes an upward trend, but as time increases, the demand growth rate gradually slows down. This is also in line with the laws of the market economy.

Model Evaluation and Comparison

In this paper, BP neural network and ARIMA are selected to perform regression forecasting on material requirements. The original data features and the EMD-processed eigenmode functions are processed separately and compared with EMD-PSO-LSSVR to verify the effectiveness of the EMD-PSO-LLSVR method. The obtained regression results are shown in the following figure:

After 100 iterations of particle swarm optimization, the optimal parameter combination Q=0.65,C=126 is obtained and the optimized parameters are used for LSSVR training and testing. In this paper, the goodness of fit (R2) indicators are selected to evaluate the prediction accuracy of the model. In this paper, the prediction model of abnormal parameters of manufacturing process based on the instantaneous frequency EMD clustering feature and the amplitude prediction of the periodic mode using LSSVR directly are shown in Table 1.

In order to evaluate the prediction effect of the model, this paper selects two parameters of MSE and R2. Both of these parameters can be used to measure the numerical accuracy of the model predictions. The prediction results and prediction accuracy of the validation set are obtained. The specific results are shown in the following table:

Table 2. Forecast Results

	Model	MSE	R^2
EMD clustering features	ARIMA	1201657	0.8723
	BP Neural Network	732504	0.9142
	PSO-LSSVR	410137	0.9583

This article will mainly compare the prediction results using autoregressive integral sliding average (ARIMA), BP neural network, and PSO-LSSVR based on instantaneous frequency EMD clustering features and original features. For the BP neural network comparison of the model proposed in this article, this paper sets the activation function as an S-function, the number of hidden layers is set to one layer, the maximum number of iterations is set to 1000, and the training accuracy is 0.01, and the learning rate is 0.1 and the adaptive gradient descent method is used as the training algorithm.

From the prediction results in Table 2, according to the MSE value of the Least Square Support Vector Regression model predicted by the EMD clustering algorithm and the particle swarm optimization algorithm, the MSE value of the LSSVR prediction model optimized by the particle swarm optimization algorithm based on the EMD clustering feature is the smallest. The R2 value is the largest and Its prediction accuracy is significantly higher than other methods. This proves that the material demand anomaly model parameter prediction model based on the instantaneous frequency EMD clustering feature can accurately predict the offset amplitude of the actual material demand mean variable and it has certain practicality.

Conclusion

This paper analyzes non-stationary time series data based on Least Squares Support Vector Machines. In order to reduce the non-stationarity of data, the original sequence is decomposed into several IMF components and residual trend terms by using Empirical Mode Decomposition (EMD). According to the statistical results of IMF components and their frequency, they are clustered into high frequencies and low frequency. Use high-frequency components to predict short-term fluctuations in material demand, and use low-frequency components to predict the long-term trend of material demand, and use particle swarm optimization algorithm (PSO) to optimize the parameters of the model, and finally the high frequency, low frequency, and residual items are summarized to obtain the final prediction result. In order to verify the validity of EMD-LSSVR, this paper selects the material demand of a company for 45 mouths as the research object, and compares the fitting and forecasting effects of ARIMA, BP neural network, and EMD-LSSVR. The results show that, after EMD decomposition, the fitting effect of LSSVR for material demand regression prediction is optimal, which verifies the effectiveness of EMD-LSSVR combined forecasting, and provides reference for enterprises to do material demand forecasting.

References

[1] Feng Yun. Research on Nonlinear Method of Supply Chain Demand Forecasting [J]. Journal of Beijing Institute of Technology (Social Sciences Edition).2008, (5).

[2] Fishman, Schandl, Tanikawa. Stochastic Analysis and Forecasts of the Patterns of Speed, Acceleration, and Levels of Material Stock Accumulation in Society [J]. Environmental Science & Technology, 2016, (50).

[3] Norden E. Huang, Zheng Shen, Steven R. Long, et al. The Empirical Mode Decomposition and the Hilbert Spectrum for Nonlinear and Non□stationary Time Series Analysis [J]. Proceedings of the Royal Society of Land A,1998, 454.

[4] Xun Zhang, K.K. Lai, Shou-Yang Wang. A New Approach for Crude oil Price Analysis Based on Empirical Mod Decomposition [J]. Energy Economics, 2008, (30)

[5] Wang Wenbo, Fei Pusheng, Ruan Xuming. The Prediction of China's Stock Market Based on EMD and Neural Network[J].Systems Engineering —Theory & Practice,2010,(6).

[6] Liu Haifei, Li Xindan. Stock Price Prediction Based on EMD Method [J].Statistics & Decision,2011,(10).

[7] Zhu Jianping, Zhang Nanxi, Zhu Wanjun. Emender Combination Prediction of Nonstationary Time Series and Its Application [J].Statistics & Decision, 2013, (15).

[8] Gong Chengzhu, Li Lanlan, Yang Juan, Zhu Kejun. Prediction of Short-term Load of Urban Gas Pipelines Based on EMD-PSR-LSSVM[J].Systems Engineering —Theory & Practice,2014,(11).