ISSN: 2414-1895 DOI: 10.6919/ICJE.202012_6(12).0043

# Port Throughput Forecast Based on Grey Model and Neural Network Combined Model--Taking Lianyungang Cargo Throughput as an Example

Jinzhu Liu

College of Transport & Communications, Shanghai Maritime University, Shanghai 201306, China.

937107426@qq.com

## Abstract

**Forecasting the port's throughput in the coming years will play an important role in the future planning and development of the port, which can effectively avoid resource waste and make the port infrastructure construction and port throughput development better match. In this paper, the grey model, BP neural network and the combined model of the two models are used to predict the port throughput. The variance of the three models is compared to obtain a better prediction method.**

## Keywords

**Port throughput, BP neural network, Grey model, Combined model.**

## 1. Introduction

Port throughput is an important basis for determining the level of port development. Forecasting port throughput plays an important role in the future planning direction of the port, the direction that needs to be invested in the future, and future management decisions. By making reasonable and scientific strategies for port throughput and making relevant analysis on market changes in real time, it is related to the degree of development and utilization of port resources and can effectively avoid resource waste.

At present, there are many methods to predict port throughput, such as time series [1-3], system dynamics method [4], GM (1, 1) [5], and a large number of combined model methods, such as ARIMA -RBF neural network [6], GRA-TOOPSO-LSSVM [7], dual time series-RBF neural network [8]. However, among these methods, a common feature of exponential smoothing, linear regression, system dynamics, and combined model method is the need for a large amount of relevant statistical data, and these data are difficult to meet the modeling requirements in actual work, and the selection of variables Improper forecasting results will also be unsatisfactory. This article will forecast the port throughput from the gray model, BP neural network and the combination of gray model and neural network.

## 2. Theoretical analysis of port forecast model

### 2.1 Grey system theory

Gray Forecast Model (Gray Forecast Model) is a forecasting method that uses a small amount of incomplete information to establish a mathematical model and make predictions [9-10]. When we apply the thinking methods of operations research to solve practical problems, formulate development strategies and policies, and make decisions on major issues, we must make scientific predictions for the future. Forecasting is based on the past and present development laws of objective things, with the aid of scientific methods to describe and analyze its future development trends and conditions,

ISSN: 2414-1895 DOI: 10.6919/ICJE.202012_6(12).0043

and form scientific assumptions and judgments. Port throughput is just such a gray system, Instead of studying the many factors that affect port throughput and their interrelationships, we can treat throughput as a time-related gray quantity that changes within a certain range, and mine useful information from its own data to build models and reveal rules. Therefore, the gray system mainly has the following three characteristics: use gray mathematics to deal with the uncertainty and quantify it; make full use of the known information to seek the law of movement of the system; gray system theory can deal with poor information systems. But for short-term forecasting, the gray model has better prediction accuracy, but for long-term fluctuations, the gray model has poor prediction accuracy. The calculation steps of the model used in this article are as follows:

(1) Accumulate generation once

Use formula 1 to accumulate the original data once to generate a new sequence of numbers to make it into an accumulative growth state.

$$x^{(1)}(i) = \{ \sum_{j=1}^{i} x^{(0)}(j) \mid i = 1,2,\cdots,N \} \tag{1}$$

(2) The calculation formula for the calculation parameters a and b is formula 2

$$\frac{dx^{(1)}}{dt} + ax^{(1)} = b \tag{2}$$

(3) Use formula 3 to predict the result:

$$\hat{x}^{(1)}(k+1) = (x^{(0)}(1) - \frac{b}{a}) * \hat{e}^{-ak} + \frac{b}{a} \tag{3}$$

$$\hat{x}^{(0)}(k+1) = \hat{x}^{(1)}(k+1) - \hat{x}^{(1)}(k) \tag{4}$$

(4) Calculating residuals and accuracy test

The specific flowchart is shown in Fig. 1

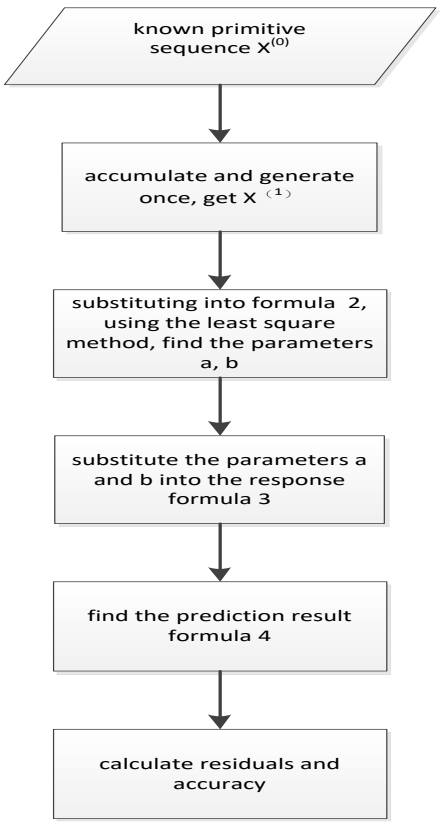


Fig. 1 GM (1, 1) calculation flow chart

ISSN: 2414-1895 DOI: 10.6919/ICJE.202012_6(12).0043

## 2.2 Neural Network Theory

BP neural network is a widely used multi-layer feedforward neural network. BP neural network training mainly relies on error back propagation. The weight and threshold of the network are continuously adjusted through the back propagation process. It can achieve many non- Solving linear problems. BP neural network has very powerful computing capabilities, and can be trained to master the mapping relationship between the input and output variables of the sample data, so that it can approximate a certain function. Applying this genetic neural network model to the prediction of coastal port throughput, experiments show that it has a better overall and faster convergence speed, and the prediction accuracy is quite high.

The learning process of BP neural network is mainly divided into two stages: forward propagation and back propagation. The first stage is the signal forward propagation process. At the beginning of training, multiple sets of sample data are input into the network at the same time, and processed through the input layer and hidden layer, and finally reach the output layer and output the result [11-13]. Each set of samples contains two parts: the input value and the expected output value; when the data is transferred between adjacent levels, it needs to be processed in combination with the transfer function, connection weight and threshold; then, the output value of each set of input samples and the corresponding The expected output values are compared, and the total error of all samples is calculated. If the error fails to reach the ideal level or the number of iterations is not enough, then the second stage is back propagation, and the information is returned according to the original connection path, so that the network can be restarted. Adjust the connection weight and threshold of each neuron, and adjust the output error to the minimum, so that the network output value is as close as possible to the value given by the actual sample.

The structure of BP network is simple and clear. The most typical model structure includes input layer (input), hidden layer (hide layer), output layer (output layer). The input layer is used for data input, and the hidden layer is used for processing data. , The output layer expresses the output of the network. When the network has an input signal, the input signal first propagates forward to the hidden layer node, and then to the next hidden layer, until it is transmitted to the output layer node. In this process, the propagation of the input signal is layer by layer. Propagation, and each time it passes through a layer, it will be processed by the corresponding characteristic function. Because the signal is transmitted layer by layer until the output layer, so the BP neural network is a parallel and multi-layer feedforward network. The BP neural network has at least one hidden layer, but it is difficult to determine the number of hidden layer nodes. In the actual neural network construction process, experience and trial and error methods are mainly used to determine the most appropriate number of nodes. Generally, refer to The model judges the approximate interval, and substitutes them into trial calculations one by one. In actual problems, the number of hidden layer nodes has different effects on the output results. The neural network structure diagram is shown in the figure:

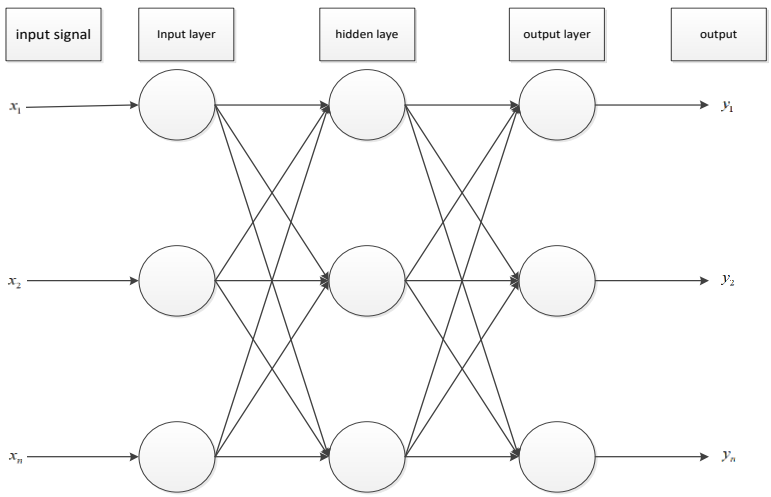


Fig. 2 Neural network structure diagram

ISSN: 2414-1895 DOI: 10.6919/ICJE.202012_6(12).0043

For the BP neural network, the number of output layer nodes is N, the number of hidden layer nodes is H, the number of output layer nodes is M, i represents the i-th input layer node, h represents the h-th hidden layer node, and the input layer node i The weight to the hidden layer node h is Whi. j represents the j-th output layer node. The weight between hidden layer node h and output layer node J is Wjh. The threshold value of hidden layer node h is bh, the threshold value of output layer node J is Whi, u represents the learning rate, and E represents the error requirement. Suppose $X_k=[x_1\ x_2\ x_3...x,]$ is the input vector, $Y_k=[Y_1\ Y_2\ Y_3...Y_m$" is the actual output vector, and $T_k=[t_1\ t_2\ t_3...t_m]$ is the desired output vector. The forward propagation process is as follows:

Input of the hth neuron in the hidden layer:

$$I_h = \Sigma_n w_{h_i} Y_i + b_h \tag{5}$$

The output of the hth neuron in the hidden layer:

$$Y_h = f(I_h) \tag{6}$$

The input of the jth neuron in the output layer:

$$I_j = \Sigma_n w_{jh} Y_h + b_j \tag{7}$$

The output of the jth neuron in the output layer:

$$Y_j = f(I_j) \tag{8}$$

Among them, f is the activation function, and the sigmoid function or linear function is generally selected. The back propagation process is as follows

Calculate sample error:

$$\mathrm{E_k} = \frac{1}{2} \times \sum_{\mathrm{j=1}}^{\mathrm{m}} (\mathrm{t_j} - \mathrm{y_i})^2 \tag{9}$$

Modify the ownership value and threshold value obtained by the output layer and the hidden layer:

$$\mathrm{w'_{jh}} = \mathrm{w_{jh}} + \Delta \mathrm{w_{jh}} = \mathrm{w_{jh}} + u_\delta \mathrm{j^M Y_h} \tag{10}$$

$$\mathrm{w'_{hi}} = \mathrm{w_{hi}} + \Delta \mathrm{w_{hi}} = \mathrm{w}_{hi} + u_\delta \mathrm{j^M Y_i} \tag{11}$$

$$b'_h = b_h + \Delta b_h = b_h + u\delta_j^m \tag{12}$$

$$b'_j = b_j + \Delta b_j = b_j + u\delta_h^{\mathrm{H}} \tag{13}$$

Each round of training uses all the records of the data set, and stops training when the stop condition is reached. The calculation process is shown in Fig. 3:

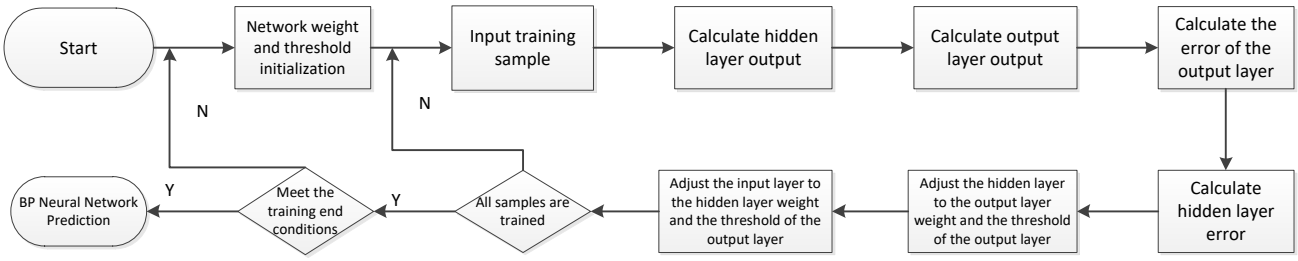


Fig. 3 Neural network calculation flowchart

The neural network calculation method has the following characteristics. For complex nonlinear systems, it has a strong nonlinear mapping ability; it can deal with regular problems hidden in a large amount of information, and it has strong adaptability and has an advantage in classification; In addition, it has a certain fault tolerance for neurons that cannot function normally. However, this method also has certain limitations. The setting of the hidden layer needs to be manually input by yourself, and the result cannot be calculated directly; the correction weight and threshold value in the network are slightly modified each time, this feature will affect the training The process enters into a

ISSN: 2414-1895 DOI: 10.6919/ICJE.202012_6(12).0043

local area and can no longer be effectively adjusted; the model is only slightly modified each time, and the mode determines the gradual reduction method, which determines its slower convergence speed.

### 2.3 Establishment of gray neural network hybrid model

Based on the above analysis of gray model and artificial neural network, this chapter proposes a method of combining the theory of gray system with artificial neural network. By using the nonlinear characteristics of neural network, the gray system requires a small amount of numbers. Solve the shortcomings of their respective prediction methods and establish a hybrid model with relatively good accuracy. The ideas adopted are as follows: (1) Input measured data; (2) Apply the gray model to make predictions to obtain predicted data; (3) Use the predicted results of the gray model as input data and the measured data as output data, and then use it The artificial neural network is trained to obtain the optimal training set. (4) Use the gray model to predict the year that needs to be predicted, then substitute the predicted data into the trained artificial neural network, and finally get the predicted data.

## 3. Case calculation

### 3.1 Sample selection

The data calculated in this experiment is the cargo throughput of Lianyungang Port from 2005 to 2018 as an experimental sample (the specific data is shown in Table 1). This experiment mainly predicts the throughput based on the throughput, without considering the influence of other factors. Under the premise of making relevant experiments to predict port cargo throughput. When processing the samples, for the gray prediction GM(1,1), it is mainly used as the initial known original sequence; for the BP neural network, it is mainly used as the training sample, if the 2005-2009 data is used as Input samples, then the 2010 data is the output data, and so on to calculate the best training results.

Table 1 Cargo throughput of Lianyungang Port

| Year | Cargo throughput (10,000 tons) |
|---|---|
| 2005 | 6016 |
| 2006 | 7232 |
| 2007 | 8507 |
| 2008 | 10060 |
| 2009 | 10843 |
| 2010 | 12739 |
| 2011 | 15627 |
| 2012 | 17367 |
| 2013 | 18898 |
| 2014 | 19638 |
| 2015 | 19756 |
| 2016 | 20082 |
| 2017 | 20605 |
| 2018 | 21443 |

### 3.2 Three algorithm prediction methods and conclusions

#### 3.2.1 GM (1, 1)

After calculation, the final calculation result of the model has good accuracy, and the posterior difference ratio is 0.32481.

#### 3.2.2 BP Neural Network

Calculation results

ISSN: 2414-1895 DOI: 10.6919/ICJE.202012_6(12).0043

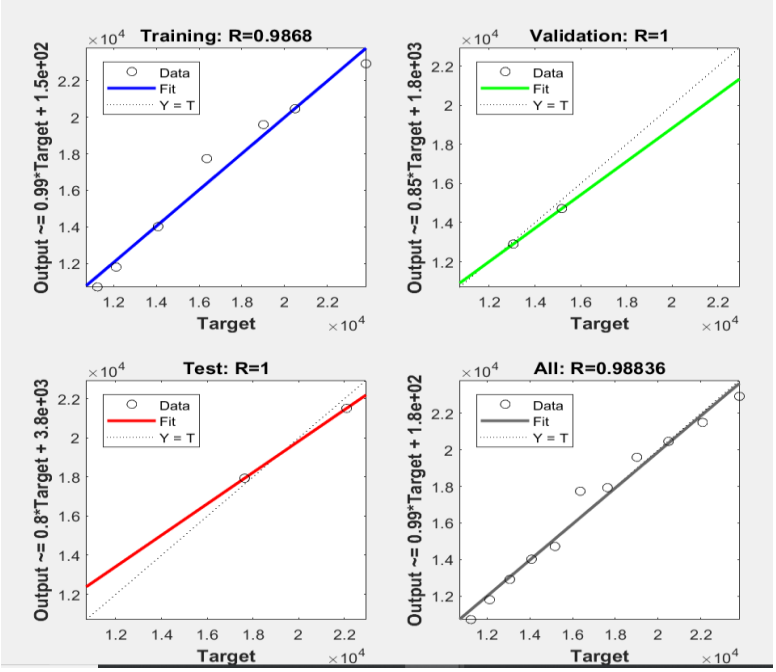


Fig. 4 BP neural network regression analysis

### 3.2.3 Two combination models

(1) Make the following changes to the input and output data in the neural network programming: load DATA2

x=M';

y=T';

The data in T is the measured data, and the data in M is the data calculated by the gray model. The two models are combined for analysis and prediction.

(2) Related accuracy results

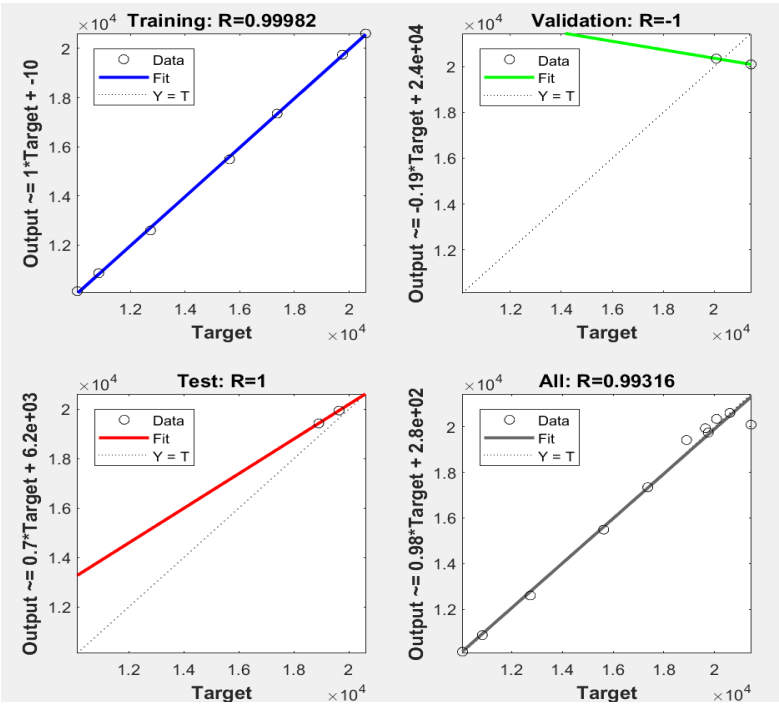


Fig. 5 Regression analysis of combined model

ISSN: 2414-1895 DOI: 10.6919/ICJE.202012_6(12).0043

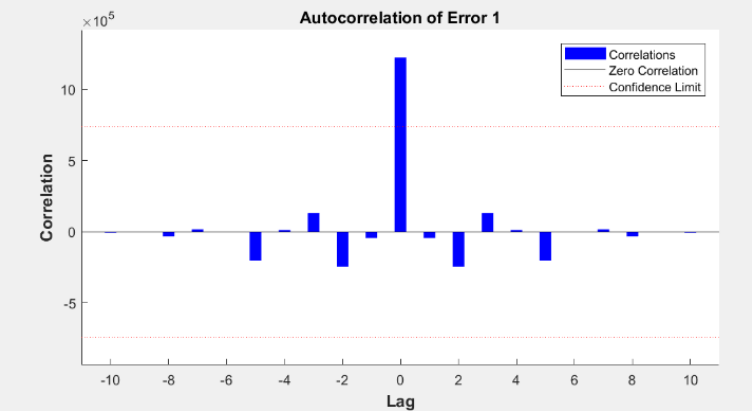


Fig. 6 Error histogram of combined model

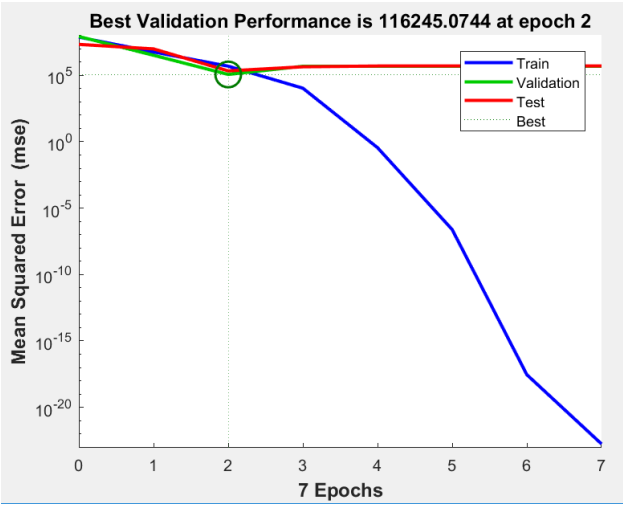


Fig. 7 Error drop line of the combined model

### 3.3 Comparative analysis of the results of the three models

After calculation and analysis of the above three program codes, the predicted and actual values shown in Table 2 can be obtained. The comparison result of the actual and predicted values can be seen in Figure 8. According to the correlation coefficient R^2, it can be known after three predictions that the best prediction is a prediction method that combines the two models. Therefore, when selecting the forecast data of less than 5 years in Table 3, the forecast data of the combined model can be used for related planning research.

Table 2 The above three model predictions and actual value units: (ten thousand tons)

| Year | Actual value | gray | Neural Networks | Combination model |
|---|---|---|---|---|
| 2008 | 10060.00 | 11236.83 | 10064.28 | 10712.85 |
| 2009 | 10843.00 | 12113.27 | 11000.04 | 11799.00 |
| 2010 | 12739.00 | 13058.07 | 12565.20 | 12910.58 |
| 2011 | 15627.00 | 14076.56 | 15628.28 | 14019.41 |
| 2012 | 17367.00 | 15174.49 | 19990.01 | 14715.43 |
| 2013 | 18898.00 | 16358.05 | 18894.45 | 17742.32 |
| 2014 | 19638.00 | 17633.93 | 19624.77 | 17935.44 |
| 2015 | 19756.00 | 19009.32 | 20370.23 | 19601.38 |
| 2016 | 20082.00 | 20491.99 | 20083.27 | 20470.21 |
| 2017 | 20605.00 | 22090.30 | 20608.00 | 21507.04 |
| 2018 | 21443.00 | 23813.28 | 21434.74 | 22936.37 |
| R^2 | | 0.89 | 0.98 | 0.99 |

ISSN: 2414-1895 DOI: 10.6919/ICJE.202012_6(12).0043

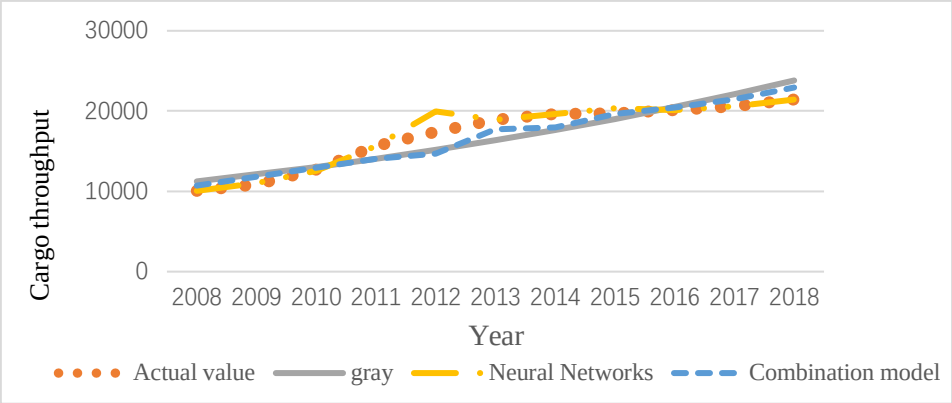


Fig. 8 Comparison of predicted values and actual values of the three prediction models

Table 3 The prediction results of the three models in the next 5 years

| Year | gray | Neural Networks | Combination model |
|---|---|---|---|
| 2019 | 25670.64 | 22842.54 | 23839.06 |
| 2020 | 27672.88 | 25168.22 | 26605.41 |
| 2021 | 29831.28 | 27669.25 | 27039.31 |
| 2022 | 32158.03 | 28657.85 | 28198.32 |
| 2023 | 34666.25 | 28562.15 | 28957.65 |

## 4. Conclusion

After related research and experiments, the model algorithm based on the combination of GM (1, 1) and BP neural network in the gray model has good prediction results for port throughput, and the prediction accuracy is relatively good. By calculating the cargo throughput of Lianyungang Port, Lianyungang Port has made relevant measures for the growth of cargo in a timely manner. According to the status of Lianyungang as a coastal port, Lianyungang port is mainly responsible for the transportation of bulk cargo, while the volume of containers is relatively small. However, according to the current green port requirements, Lianyungang Port needs to promptly provide application measures related to the construction of green ports in the face of so much bulk cargo volume.

## References

[1] Y.Q. ZHAO. Port throughput forecast based on time series analysis[J]. Market Research, 2018, (8):35-37.

[2] S.W. ZHAO, J.H. ZHOU. China Port Container Throughput Forecast: Based on Combined Time Series[J]. Systems Science and Mathematics, 2018, 38(2): 210-219.

[3] S. WU. Port container throughput forecast based on time series model[J]. Pearl River Water Transport, 2019, (5):73-74.

[4] K.K. HE, Y. CHEN. Wuhan New Port Container Throughput Forecast Based on System Dynamics[J]. Logistics Technology, 2017, Volume 36 (6):57-62.

[5] X. TIAN, D.D. WANG, Y.Y. WANG, et al. Research on Port Throughput Forecast Based on Grey Model——Taking Caofeidian Port as an Example[J]. Mathematical Practice and Knowledge,2018.

[6] M.Y. YIN. Research on Throughput Forecast of Coastal Port Based on ARIMA-RBF Neural Network[J]. Journal of Wuhan University of Technology (Transportation Science and Engineering Edition),2014, 38(1):241-244.

[7] H.C. ZHANG, Y.F. HUANG, J.K. HU. Port throughput forecast based on GRA-TOOPSO-LSSVM[J]. Journal of Shanghai Maritime University,2017, 38(1):43-46,89.

[8] S.E. XI. Research on Throughput Forecast of Coastal Ports Based on Dual Time Series-RBF Neural Network[J]. Smart City, 2016(9):157-159.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202012_6(12).0043

[9] N.M. XIE, K. ZHANG. Discrete Grey Forecasting Model and Its Application[M]. Science Press, 2016.

[10] D. LUO, B.L. WEI. The unified processing method and application of a kind of discrete grey forecasting model[J]. System Engineering Theory and Practice,2019, 39(2):451-462.

[11] C.B. SHENG. BP neural network principle and MATLAB simulation[J]. Journal of Weinan Teachers College, 2008, 23(5).

[12] Anonymous. Principles of Artificial Neural Networks [M]. 1995.

[13] J. ZHANG, J. ZHAN, J.C. ZHANG. 30 cases of MATLAB neural network[M]. Publishing House of Electronics Industry, 2014.