Forecast of Guangzhou Port Logistics Demand Based on Back Propagation Neural Network

Xiu Chen[1], Lianhua Liu[2], Lifen Zheng[3]*
School of Management, Guangzhou Huashang College, Guangzhou 511300, China [1, 3]
Faculty of Humanities and Social Sciences, Macao Polytechnic University, Macau 999078, China [2]

***Abstract*—In recent years, the port economy of our country has developed rapidly. Guangzhou Port is an important node of the maritime transportation of the Belt and Road, connecting the hinterland economy of our country with the countries along the Belt and Road, which is of great significance in promoting the economic development of the hinterland of our country. It is of great significance to predict the freight development demand of Guangzhou port scientifically and reasonably, which is beneficial to optimize the infrastructure construction and logistics system planning of Guangzhou port. This study selects the port cargo throughput, foreign trade cargo throughput, and container cargo throughput as three index values to measure the freight development of Guangzhou port. Firstly, the GM(1,1) model and the BP neural network model are constructed to predict the freight demand of Guangzhou port. Then, the GM(1,1) model and the BP neural network model are combined to predict again. By comparing the three models, the results show that the accuracy of the combined model is better than that of the single model. The combined model of BP neural network and GM(1,1) can be effectively applied in the prediction of Guangzhou port logistics demand. Finally, the combined model of BP neural network and GM(1,1) is used to forecast the freight development demand of Guangzhou Port in 2022-2024, which provides a reference for the development planning of Guangzhou Port. The results further indicate that the BP–GM(1,1) combination model significantly outperforms single forecasting models in terms of prediction accuracy, highlighting its effectiveness and robustness in port logistics demand forecasting.**

***Keywords*—BP neural network; GM(1,1); combination model; port logistics demand**

I. Introduction

Ports are the intersection of sea and land transport, providing a place for industrial activity and, more importantly, an integrated logistics center that promotes urban and socio-economic development. As a transportation hub linking waterways and landways, the port not only connects the inland hinterland to the sea, but also builds a platform for China to communicate with external economic materials and provides a node for interfacing with the international community. In 2015, the National Development and Reform Commission and other departments jointly released the "Vision and Action for Promoting the Construction of the Silk Road Economic Belt and the 21st Century Maritime Silk Road. Guangdong is one of the designated locations of the 21st Century Maritime Silk Road. As the largest comprehensive hub port in South China, Guangzhou Port also occupies a very important position in the southern route of the new Silk Road. Moreover, the port of Guangzhou makes an important contribution to the economic development of China's hinterland regions, with strong point and line radiation capabilities that can radiate the economies of several regions. Zhuang et al. confirm that port-throughput forecasting can directly support hinterland industrial planning and green-port development, offering a methodological reference for the sustainable multi-regional economic spillover of Guangzhou Port [1]. Guangzhou Port is an important export channel in the Pearl River Delta of Guangdong Province and one of the most important foreign trade ports in China. On February 18, 2019, the Central Committee of the Communist Party of China and the State Council issued the Outline of the Development Plan for the Guangdong-Hong Kong-Macao Greater Bay Area, with the four central cities of Guangdong, Hong Kong, Macao and Shenzhen as important forces to guide the economic development of other regions. In addition, the Guangdong-Hong Kong-Macao Greater Bay Area contributes to the construction of the Belt and Road Initiative in the planning outline. Guangzhou Port is located at the mouth of the Pearl River or the center of the Pearl River Delta, adjacent to Hong Kong and the South China Sea. Therefore, it is very important to predict the logistics demand of Guangzhou port for the construction of the Guangdong-Hong Kong-Macao Greater Bay Area. Europe and the United States attaches great importance to the shipping and port economy, but also because of early industrialization their port economic level is now located in the world, even our country every year for ports, shipping cost is higher, the present economic level is still in the harbor is inferior to European and American countries, thus to promote the development of port economy in China is urgent. Therefore, China should make relevant strategies according to local conditions to promote the development of port logistics and improve the economic level of port logistics to reverse the disadvantage in the world. In order to optimize the port level in our country with high efficiency, it is necessary to make a reasonable prediction of the port logistics demand, reduce the prediction error and improve the efficiency of port logistics operation. During the "13th Five-Year Plan" period, the smart port demonstration list was announced, marking the port development in our country has entered a new period, through a series of new technologies such as the Internet of Things, mobile Internet, big data, cloud computing, and so on. It comprehensively includes the production and operation, service and management of the port, reflecting the open sharing, connectivity, highly digitalized, and intelligent port. Therefore, the analysis of the logistics needs of Guangzhou port is conducive to optimizing the logistics system planning and layout, and improving the scientific planning and layout of the "wisdom degree" of Guangzhou Port.

*Corresponding author.

Despite the extensive attention paid to port logistics development, existing studies on port logistics demand forecasting are still dominated by single-model approaches, including time-series analysis, regression models, and standalone artificial intelligence methods. These approaches often face limitations such as insufficient adaptability to small samples, sensitivity to data fluctuations, or overfitting under complex multi-factor logistics environments. Xu et al. demonstrate that traditional time-series and single-machine-learning models produce significantly higher MAPE when sample sizes are limited or volatility is high, underscoring the need for a hybrid learning framework to improve robustness [2]. Moreover, many existing AI-based or hybrid forecasting models focus on general freight systems. Xiao et al. propose a VMD-CNN-LSTM-Attention decomposition-ensemble model for four major Asian hub ports, reducing forecast error by 27 % during the COVID-19 shock and proving that the decomposition-ensemble strategy is more resilient for port-logistics-demand forecasting under extreme events [3]. With limited empirical validation in large comprehensive hub ports under strategic initiatives such as the Belt and Road and the Guangdong–Hong Kong–Macao Greater Bay Area.

To address these limitations, this study constructs a BP neural network–GM(1,1) combination forecasting model for port logistics demand. The main contributions of this study are threefold: first, it integrates the advantages of grey system theory and neural networks to improve forecasting accuracy under limited and nonlinear data conditions; second, it applies the proposed combination model to Guangzhou Port using multiple cargo throughput indicators, providing an empirical reference for major hub ports; third, it offers quantitative support for logistics resource allocation, smart port planning, and infrastructure decision-making. This study thus fills an important gap in port logistics demand forecasting research by providing a robust and interpretable hybrid modeling framework.

II. Literature Review

A. *Significance of Port Logistics Development*

Scholars' research on the significance of the development of port logistics mainly focuses on the promotion mechanism of the development of port logistics on the regional economy and foreign trade. The geographical location of a port is important for port development. The results of the study showed that the regional advantages of the port's location can contribute to the development of port logistics itself [4]. The development of port logistics promotes the economic development of the region in which it is located, and studies from the economic structure, employment of talents, and the interaction between ports and urban development have found that port logistics has a great promotional impact on the economy of the region in which it is located [5]. Studies on the relationship between port infrastructure quality and logistics performance have shown that improving the quality of port logistics development plays an important role in a country's economic development [6].

B. *Progress of Research on Port Logistics Development Forecasting*

In recent years, port logistics development forecasting has become a hot issue studied by scholars. Port logistics demand forecasting is to forecast the indicators related to the development trend of market demand based on the relationship between the past and present demand of the port logistics market and the factors affecting the change of logistics market demand, and to utilize appropriate empirical judgment techniques, methods, and forecasting models. According to the accurate port logistics demand forecast, port enterprises can timely and accurately grasp the changing law of demand, seize the favorable opportunity to arrange port construction and production plans. In the research method of forecasting model of port logistics, scholars already have rich theoretical results and practical experience, and the current forecasting methods mainly include time series analysis, support vector machine and gray forecasting model [7-14]. Recent studies have increasingly focused on hybrid and AI-driven forecasting models for port logistics demand, emphasizing prediction accuracy improvement and operational applicability under complex logistics environments. In addition to conventional forecasting methods, scholars have innovated forecasting methods, based on the fact that the process of forecasting methods is highly dynamic, and a simulation-based method for forecasting demand load profiles has emerged [15]. Multiple regression and AW-BP prediction methods based on system order parameters have also emerged [16]. Practice has proved that the new prediction method not only has great improvement in convergence speed, prediction accuracy, and avoidance of local extremes, but also has the characteristics of adaptability, simplicity, practicability, and high efficiency, which is worthy of popularization and application. Scholars' research on port logistics demand forecasting methods is mainly to construct a single model. There are mainly regression equation forecasting methods, time series forecasting methods, artificial intelligence algorithm and so on. And the single models have their own limitations. The range of sample values in the regression equation forecasting method affects the scope of application of the regression equation. Generally, the time series forecasting method is suitable for short-term and medium-term forecasting, because it shows that the economic development has more uncertainty and is easily affected by a variety of emergencies, and the time series forecasting model does not take into account other factors besides the time factor, so it is easy to come up with the forecasting results that do not conform to the actual economic development, and it has the unavoidable forecasting error. The reality is that there are many unexpected situations, and if we ignore the influence of external factors on the predicted object, the prediction results will be seriously inconsistent with the actual situation. In recent years, the research on the prediction of port container throughput has shown the evolution of "from single source to multi-source, from shallow to deep". Liang et al. took the lead in proposing a deep learning hybrid model that takes into account hinterland two-way economic indicators, incorporating the data of port hinterland and forward hinterland into the feature space at the same time. The empirical results show that RMSE is 18% lower than the traditional Arima, verifying the marginal contribution of "spatial interactive information" to the prediction accuracy, and providing a quantifiable decision-making basis for port operators [17]. Zeng and Xu further introduced the "decomposition before integration" strategy: Using Particle Swarm Optimization (PSO) to adaptively determine the parameters of variational mode decomposition (VMD), the original throughput sequence is decomposed into several intrinsic mode functions (IMF), and then the gated cyclic unit (Gru) is used to predict mode by mode, and finally aggregate the output. In the Medium-term Forecast Experiment of China's three major hub ports (Guangzhou, Qingdao, and Shanghai), the MAPE of this framework is 27% lower than that of the single Gru model, indicating that the vmd-pso-gru combination can effectively capture the nonlinear and seasonal fluctuations of shipping demand [18]. Zhang et al., for the first time, transformed the throughput time series into a complex network from the perspective of system science, used 13 similarity indicators to describe the node Association, and then used the maximum correlation and minimum redundancy (mrmr) to screen the features. Finally, support vector machine (SVM), deep neural network (DNN), and long-term and short-term memory network (LSTM) were used for link prediction and regression estimation. This method explicitly incorporates "network topology dynamics" into the prediction framework, which provides a structural perspective for explaining throughput mutations and verifies the stability of ensemble learning in ultra-high-dimensional feature space [19]. To sum up, the current research has formed a general technical route of "multi-source data fusion - Feature Engineering - hybrid deep learning": the hinterland economic indicators are used to characterize the external impact, the noise is filtered by modal decomposition, the structural features are extracted by a complex network, and then the nonlinear mapping is realized by depth model. Because single models have their own limitations, the use of combined models to optimize prediction and improve prediction accuracy has become the focus of research.

C. Application of the GM(1,1) Model and the BP Neural Network Prediction Model in Freight Transportation Forecasting

Gray system theory is an effective method for studying and modeling systems composed of small samples, which contain a limited amount of information and have a wide range of applications in many fields [20]. Valuable information is extracted by processing the known information. The method is further utilized to explore the evolutionary laws of the system, leading to the development of predictive models As there are many factors affecting the development of port logistics, such as transportation and logistics environmental factors, regional economic environmental factors, government policy factors, scientific and technological environmental factors, etc., it can be regarded as a gray system [21-26]. BP neural network prediction can predict the demand of port logistics through machine learning training, which is scientific and feasible to be applied in the prediction of port logistics. In summary, a single model prediction has greater limitations, and there are fewer studies using the gray prediction model combined with the BP neural network model to predict the logistics demand of ports, so this study constructs a BP neural network-GM(1,1) combination model to carry out a study on logistics demand prediction of the Port of Guangzhou, to provide data support for the rational allocation of the logistics resources of the Port of Guangzhou and the construction of a smart port.

III. Selection of Research Indicators and Data Sources

Selection of research indicators: Indicators selected for port research should comprehensively reflect port operation and production. Referring to relevant literature, this study chooses throughput to analyze port output, which mainly includes the following three indicators: Port cargo throughput, foreign trade cargo throughput, and container cargo throughput. The influencing factors of port logistics are mainly considered from the perspective of demand capacity, mainly from the three indicators of regional GDP, fixed asset investment, and the total value of commodity import and export. Data sources: Official website of the Ministry of Transport of China, Guangzhou Statistical Yearbook. This study mainly selects the relevant data of cargo throughput, foreign trade cargo throughput, and container cargo throughput of Guangzhou port from 2011 to 2020 for research. By constructing the 2011-2020 prediction model, the predicted value of 2021 is obtained and then compared with the actual data of 2021 to judge the accuracy of the prediction model. In order to facilitate expression, the port cargo throughput (A), foreign trade cargo throughput (B) and container cargo throughput (C) are represented by A, B, and C, respectively.

IV. Modeling of Guangzhou Port Logistics Demand Forecasts

A. GM(1,1) Model

1) Rank ratio test: First of all, the prerequisite for the construction of the GM(1,1) model is that the original sequence can pass the rank ratio test, which is as follows [see Eq. (1) to Eq. (3)]:

- Calculate the grade ratio λ(k):

$$\lambda(k)A=(0.992\ 0.956\ 0.944\ 0.963\ 0.958\ 0.917\ 0.96\ 0.981\ 0.988) \tag{1}$$

$$\lambda(k)B=(0.904\ 0.968\ 0.947\ 1.007\ 0.942\ 0.981\ 0.932\ 0.969\ 1) \tag{2}$$

$$\lambda(k)C=(0.978\ 0.951\ 0.933\ 0.943\ 0.935\ 0.938\ 0.93\ 0.947\ 0.985) \tag{3}$$

- Judge the level ratio

When $\lambda(k)\epsilon(e^{\frac{2}{n+1}},e^{\frac{2}{n+1}})$, when n=10, $\lambda(k)\epsilon(0.834,1.199)$, and when λ(k) falls into the above range, the GM(1,1) model is suitable.

$\lambda(k)\epsilon(0.917,0.988)$,K=2,3,4,5,6,7,8,9,10, therefore, it can be modeled by GM(1,1).

$\lambda(k)\epsilon(0.904,1.007)$,K=2,3,4,5,6,7,8,9,10, therefore, it can be modeled by GM(1,1).

$\lambda(k)\epsilon(0.93,0.985)$,K=2,3,4,5,6,7,8,9,10, therefore, it can be modeled by GM(1,1).

2) Construct GM (1,1) model: Using the data from 2011 to 2020 as the original series, the GM(1,1) prediction model is obtained according to the following five steps, as shown in Eq. (6) to Eq. (8), and the predicted value is calculated according to the GM(1,1) prediction model, as shown in Table I and Eq. (4) to Eq. (7):

Step 1: Set the original sequence as follows:

$$X^{(0)}=\{X^{(0)}(1),X^{(0)}(2),\ldots,X^{(0)}(n)\} \tag{4}$$

Step 2: Generate a cumulative sequence

$$X^{(1)}=\{X^{(1)}(1),X^{(1)}(2),\ldots,X^{(1)}(n)\} \tag{5}$$

Step 3: Create differential equations based on the cumulative series:

$$\frac{dX^{(1)}}{dt}+aX^{(1)}=b \tag{6}$$

where, a,b are the parameters to be estimated for the differential equation.

Step 4: Least squares method to solve the equation:

$$X^{(1)}(k+1)=\left(X^{(0)}(1)-\frac{b}{a}\right)e^{-ak}+\frac{b}{a} \tag{7}$$

Finally, the prediction equations for the GM(1,1) model were derived by calculating the coefficients a and b through MATLAB as follows:

Port cargo throughput (A): $955544.3367e^{0.045k}-912395.6667$.

Foreign trade cargo throughput (B): $307749.5686e^{0.035k}-297836.0286$.

Container Cargo Throughput (C): $23944.67667e^{0.035k}-22502.56667$.

The actual value from 2011 to 2020 is plotted against the predicted value of GM(1,1), as shown in Fig. 1 to Fig. 3. It is found that the actual value of port cargo throughput, foreign trade cargo throughput and container cargo throughput have a consistent growth trend with the predicted value, which is a relatively stable positive growth. The previous data from 2011 to 2013 almost overlap, while the actual and predicted values from 2014 to 2020 are also close, proving that GM(1,1) is suitable for port logistics forecasting.

TABLE I. COMPARISON BETWEEN ACTUAL VALUE AND GM(1,1) PREDICTED VALUE FROM 2011 TO 2020 (UNIT: TEN THOUSAND TONS AND TEN THOUSAND TEU)

Year	Port cargo throughput		Foreign trade cargo throughput		Container cargo throughput	
	Actual value	Predicted value	Actual value	Predicted value	Actual value	Predicted value
2011	43148.67	43148.67	9913.54	9913.54	1442.11	1442.11
2012	43517.39	43996.035	10968.42	10957.839	1474.36	1479.862
2013	45516.57	46031.058	11329.27	11345.164	1550.45	1570.813
2014	48217.25	48160.21	11957.1	11746.18	1662.62	1667.354
2015	50053.01	50387.846	11868.51	12161.371	1762.49	1769.828
2016	52253.82	52718.52	12596.03	12591.237	1885.77	1878.6
2017	57003.47	55156.999	12843.79	13036.298	2010	1994.057
2018	59396.2	57708.268	13782.16	13497.09	2162.27	2116.61
2019	60520	60377.546	14221	13974.17	2284	2246.695
2020	61240	63170.291	14224	14468.113	2319	2384.774

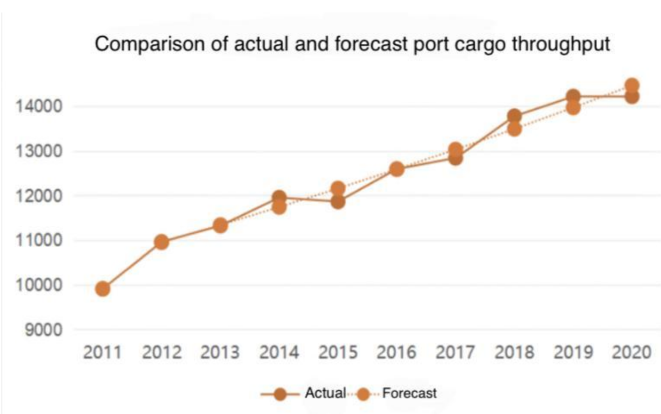

Fig. 1. Comparison of actual and forecast port cargo throughput 2011-2022.

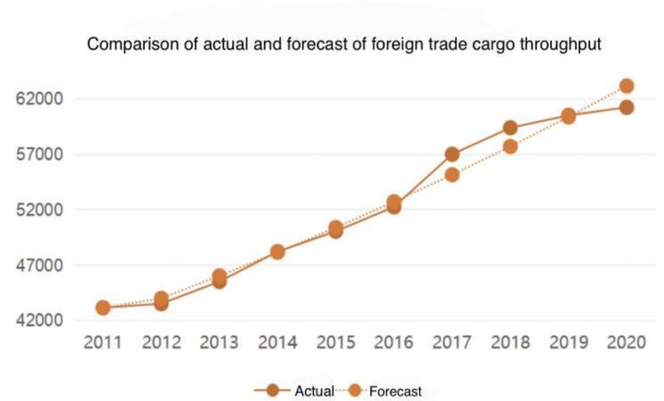

Fig. 2. Comparison of actual and forecast of foreign trade cargo throughput 2011-2022.

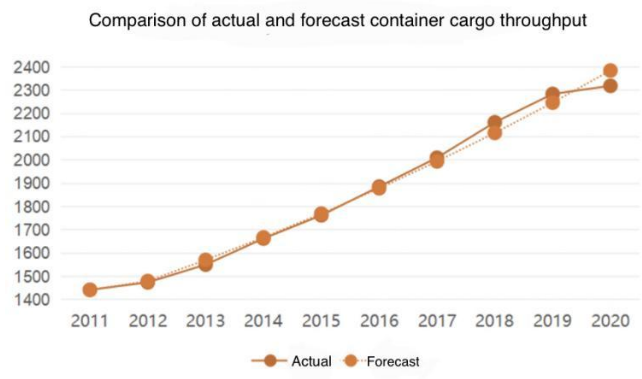

Fig. 3. Comparison of actual and forecast container cargo throughput 2011-2022.

3) Post-test difference ratio test and relative error test: After calculating the prediction equations and predicted values, the accuracy of the model needs to be further tested. The posterior difference ratio can verify the accuracy of the gray prediction. When the posterior difference ratio is smaller, it indicates the high accuracy of the gray prediction model, and the relative error is the value obtained by multiplying the ratio of the absolute error caused by the measurement and the true value of the measured (convention) by 100%, which can also reflect the credibility of the measurement. Therefore, the posterior difference test and the mean relative error test are performed in this study, as follows in the Table II.

TABLE II. GM(1,1) ACCURACY TEST

	Post-test ratio	Average Relative Error %
Port cargo throughput A	0.024	1.338
Foreign trade cargo throughput B	0.02	1.153
Container Cargo Throughput C	0.009	1.014

In general, a posterior difference ratio less than 0.35 indicates a high model accuracy. The ratio less than 0.5 indicates that the model accuracy is qualified. The ratio less than 0.65 indicates that the model accuracy is basically qualified. Analysis of the above table shows that the posterior difference ratio of the GM(1,1) model consisting of cargo throughput, foreign trade cargo throughput, and container cargo throughput is less than 0.35, and all of them are between 0.009 and 0.024, which proves that the GM(1,1) model is highly accurate. Also, the lower the relative error value, the better. A relative error value of less than 20% indicates a good fit of the model. Therefore, in order to judge the model fit of the gray prediction model, the relative error was calculated for the predicted values from 2011 to 2020 and then averaged to obtain the interval (1.014%-1.338%), which implies a good model fit. After constructing the GM(1,1) model and passing the posterior difference ratio test and relative error test, it shows that the GM(1,1) model constructed in this study has good accuracy and can therefore be used to forecast the port logistics demand in Guangzhou.

4) Forecast accuracy: This study assumes that the logistics demand of Guangzhou port in 2021 is unknown and needs to be forecasted by the GM(1,1) model, and then the data obtained from the forecast is compared and analyzed with the actual data published by the Ministry of Transport of China in 2021 to determine whether the GM(1,1) model is applicable to the logistics demand forecast of Guangzhou port, as shown in Table III.

TABLE III. COMPARISON OF ACTUAL AND FORECAST VALUES IN 2021

2021	Port cargo throughput	Foreign trade cargo throughput	Container Cargo Throughput
Actual	62367	15795	2418
Forcast	66092.21313	14979.51533	2531.340499

The 2021 cargo throughput, foreign trade cargo throughput and container cargo throughput predicted by GM(1,1) model are 66,092, 14,980 and 25,310 (million tons) respectively after rounding, and the actual values are 62,367, 15,795 and 24,180 (million tons) respectively. Calculating the error between the forecasted and actual values of the above three throughputs in 2021 yields 3725, 816, and 113, and it can be found that the error is relatively large for cargo throughput. But it is understandable due to the large base. Analysis of the table shows that the GM(1,1) model is higher than the actual value for both cargo throughput and container cargo throughput, and lower than the actual value for the forecast of foreign trade cargo throughput. The accuracy of the GM(1,1) model for foreign trade cargo throughput and container cargo throughput is excellent, but the predicted value of cargo throughput differs significantly from the actual value. Therefore, this study hopes to compare the prediction by combining a BP neural network and GM(1,1) model to investigate whether the prediction accuracy can be improved to reduce the error.

B. BP Neural Network Model

1) Basic concepts of the BP neural network model: The BP neural network is a member of a large family of artificial neural networks. BP neural network prediction method is a multi-level neural network with reverse flow according to error, which can not only distinguish different modes, but also has a good mapping function for multi-dimensional functions [15]. In this way, the method solves a problem that other perceptrons cannot understand. The principle is to achieve the optimal value of the network target by gradient reduction. In general, it is a multidimensional function and takes the square error of the network as the variable, and calculates the minimum value of this function by the gradient descent method. And it is mainly composed of input layer, hidden layer and output layer. The BP algorithm consists of forward and reverse calculations. Because the BP neural network has been relatively mature in network theory and performance and has been used in many aspects, such as artificial intelligence, it will continue to be optimized and used in the future. It is a very popular model.

2) Construction of BP neural network logistics demand forecasting model

a) Data selection and source: Since the Guangdong Statistical Information Network does not publish the gross regional product, fixed asset investment, and total value of merchandise import and export in Guangzhou in 2021. Therefore, this study selects the data from 2011 to 2020 for modeling the BP neural network and constructs BP neural networks for cargo throughput, foreign trade cargo throughput and container cargo throughput, respectively. The data from 2011 to 2019 were used as the training set and 2020 as the validation set.

b) Number of layers in the network: A network unit of a BP neural network can correspond to multiple inputs but only one output, and the input layer corresponds to neurons composed of influence factors. Gross regional product, fixed asset investment, and gross merchandise import and export are the input layers of the three neurons chosen for this model. Gross regional product, fixed asset investment, and gross merchandise import and export are the input layers of the three neurons chosen for this model. The output layers of the model are cargo throughput, foreign trade cargo throughput and container cargo throughput, respectively. After the calculation, it was finally decided to use 5 hidden layers. The structure of the BP neural network is shown in Fig. 4.

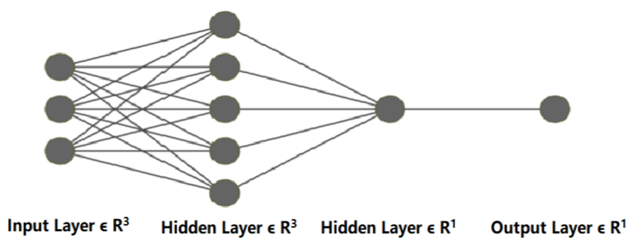

Fig. 4. BP neural network structure diagram.

c) BP neural network activation function: The sigmoid function and tanh function are more typical functions that can be applied to BP neural networks. Because the step function is not derivable, there is a more complicated problem of calculation, so this study chooses the sigmoid function as the network layer activation function. Because the sigmoid function is not only derivable but also can be found by the gradient descent method to find the extreme value, the computational difficulty is much reduced. In summary, in this study, we chose to use the S-type function (Sigmoid) as the activation function of the network layer as follows:

$$f(x)=\frac{1}{1+e^{-2n}}-1 \tag{8}$$

3) BP neural network prediction accuracy: A BP neural network model is constructed using MATLAB, and three cargo throughputs from 2011-2019 are selected for the training set, and the 2020 data is the validation set. First, the BP neural network is set to display the results once every 10,000 rounds. Then the learning speed is set to 0.001 and the maximum training rounds are 50,000. After the BP neural network model is trained, the errors for 2011-2019 are below 1%, and R^2 is very close to 1. This proves that the predicted values of the BP neural network after training are very close to the actual values, and the predicted values for 2011-2019 can be approximated to be equal to the actual values. The absolute errors of cargo throughput, foreign trade cargo throughput, and container cargo throughput for 2020, as the validation set, are 3.4215%, 2.4999%, and 2.5275%, respectively. The data show excellent prediction accuracy with small errors. The detailed training results are shown in Fig. 5 to 7.

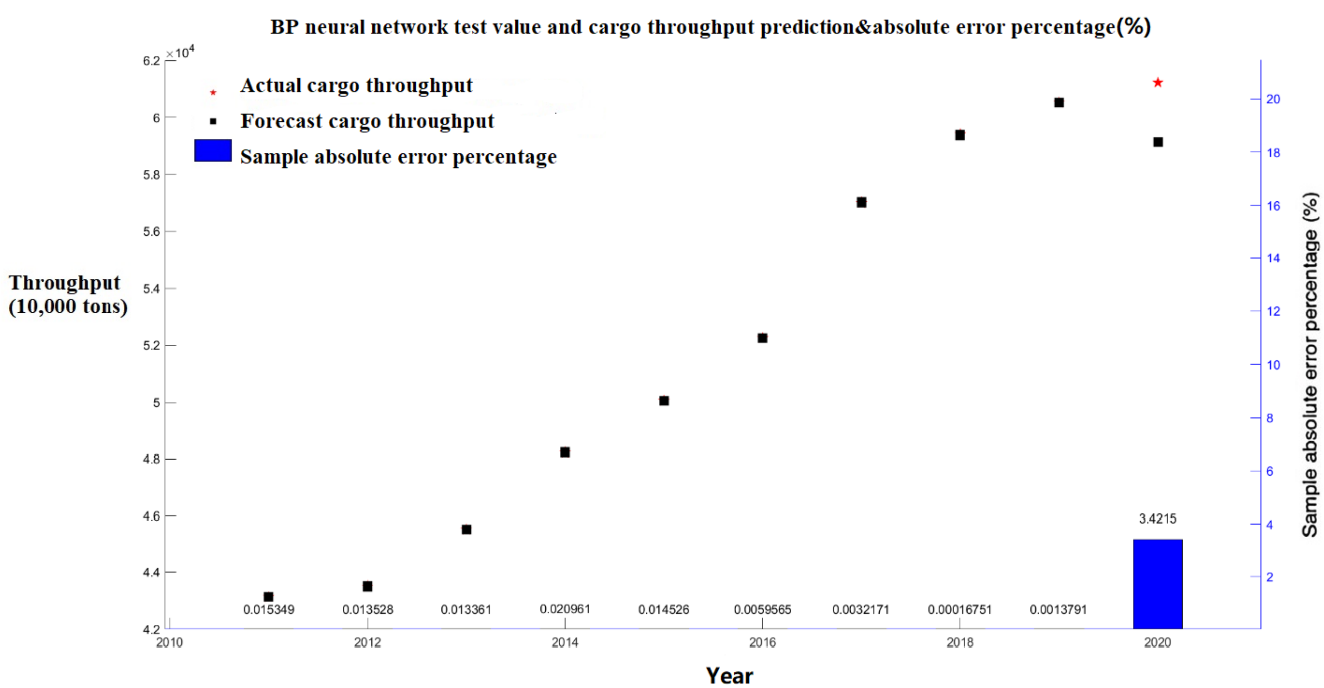

Fig. 5. BP neural network test values and 2020 cargo throughput forecast and absolute error percentage.

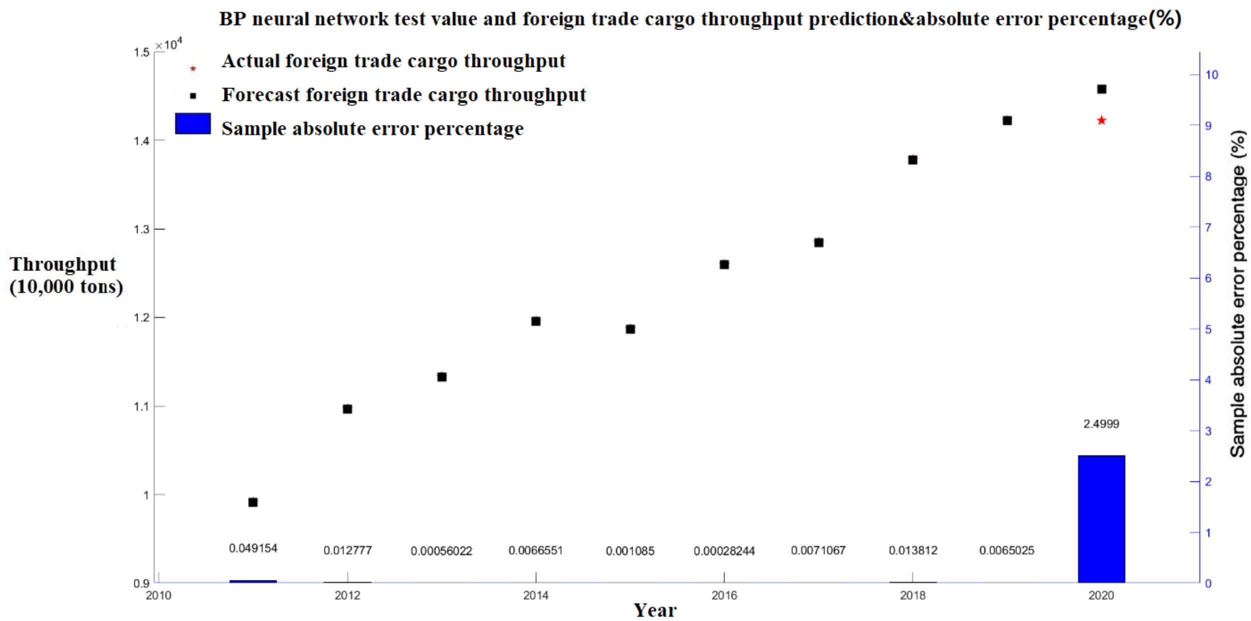

Fig. 6. BP neural network test values and 2020 foreign trade cargo throughput forecast and absolute error percentage.

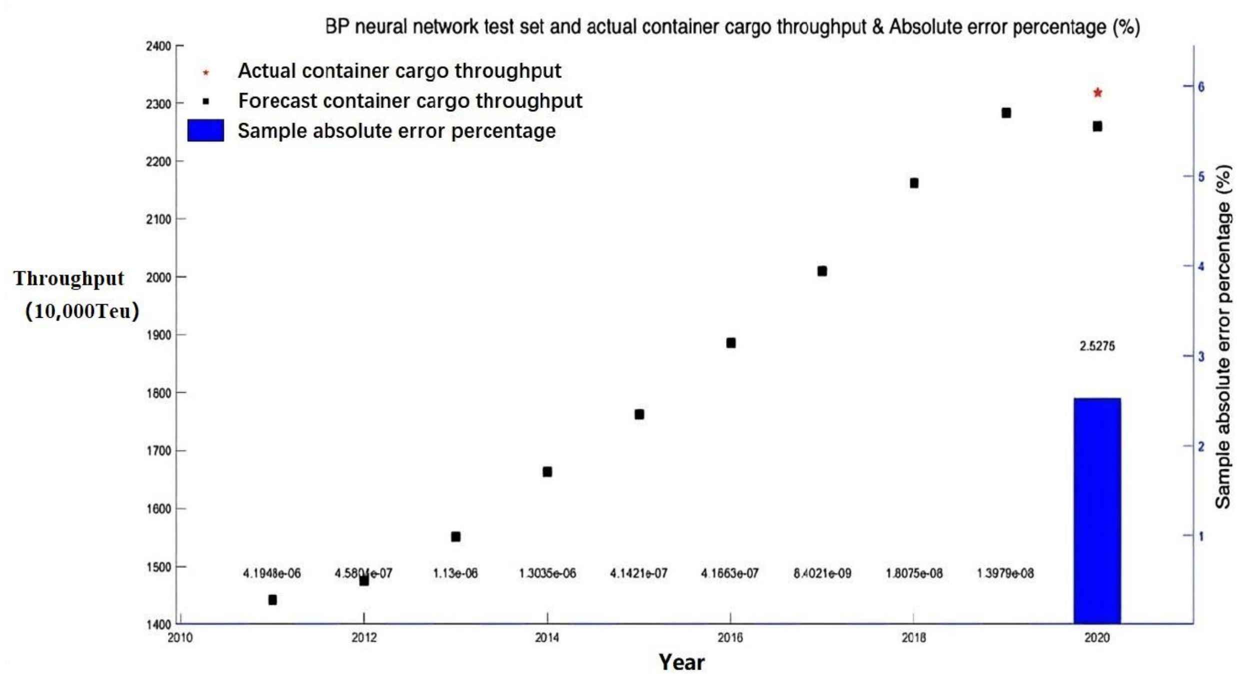

Fig. 7. BP neural network test values and 2020 container cargo throughput forecast and absolute error percentage.

C. BP Gray Prediction Combined Model

The GM(1, 1) model shows a large error in the prediction for cargo throughput, while the error of cargo throughput in the BP neural network prediction is only 3.4215%, which is a better prediction result. Therefore, this study decides to construct a combined model with a BP neural network model and a gray prediction model to investigate whether the combined model is better than a single model in terms of accuracy.

1) Determination of weight coefficients: Combined model y_t is shown in Eq. (9) and Eq. (10). There are many ways to combine individual models, among which the most critical is the establishment of weight coefficient ω_i. Considering the convenience of calculation and referring to most literature, this study finally chooses the equal-weight average method as Eq. (10):

$$y_t=\sum_{i=1}^{k}\omega_i f_{it}\ ,\ \sum_{i=1}^{k}\omega_i=1 \text{ and } \omega_i\geq 0 \quad (9)$$

where, ω_i is the weight, f_{it} is the predicted value, and k=1，2…n represents the number of models.

$$\omega_i=\frac{1}{k} \quad i=1,2,\ldots,k \tag{10}$$

After determining the weight composition, it is necessary to calculate the predicted value and the actual value, and then carry out the combination forecast and analyze the forecast results.

2) Combination prediction: The logistics demand of Guangzhou Port in 2020 is forecast, compared and analyzed, and the calculation results are shown in Table IV below:

TABLE IV. COMPARISON BETWEEN PREDICTED VALUES AND ACTUAL VALUES OF EACH FORECAST MODEL IN 2020 (TEN THOUSAND TONS AND TEN THOUSAND TEU)

	Actual	BP	GM(1,1)	Combination model
Port cargo throughput A	61240	59144.6734	63170.291	61157.4822
Foreign trade cargo throughput B	14224	13868.41422	14468.113	14168.26361
Container cargo throughput C	2319	2260.387275	2384.774	2322.580638

The table is drawn according to the output results of the BP neural network model, GM(1,1) model, and the BP neural network grey prediction combination model. The analysis results show that the predicted values of the BP neural network model are all smaller than the actual values, and the predicted values of the GM(1,1) model are all larger than the actual values. The combination model of the two is closer to the actual value than the single item model, weakening the limitations of the two single item models, and improving the accuracy of prediction to a certain extent. It proves that the combination model of BP neural network and GM(1,1) is more effective and conducive to reasonably predicting the future port logistics demand in the case of predicting Guangzhou port logistics.

3) Prediction accuracy: In order to further analyze the output results of the individual model and the combined model, the absolute percentage error is calculated, respectively, and the following table is obtained (see Table V):

TABLE V. ABSOLUTE PERCENTAGE ERROR BETWEEN EACH MODEL AND THE ACTUAL VALUE (%)

	BP	GM(1,1)	Combination model
Port cargo throughput A	3.422	3.152	0.135
Foreign trade cargo throughput B	2.500	1.716	0.392
Container cargo throughput C	2.528	2.836	0.154

It can be clearly seen that the combination model constructed by BP neural network model and GM(1,1) improves the prediction accuracy: the absolute percentage error of cargo throughput is reduced from 3.422% to 0.135%. The absolute percentage error of foreign trade cargo throughput decreased from 2.5% to 0.392%, and the absolute percentage error of container cargo throughput decreased from 2.528% to 0.154%. Through the combination model, the forecast results are optimized to a large extent, and it is an effective combination model, which is suitable for the Guangzhou port logistics demand forecast.

To sum up, the combination model constructed by the BP neural network model and GM(1,1) can effectively improve the prediction accuracy of the model, which can be applied to the demand prediction of Guangzhou port logistics.

4) Forecast of logistics demand mix of Guangzhou Port in 2022-2024: After comparing and analyzing the three models, the prediction accuracy of the combined model constructed by the BP neural network model and GM(1,1) is better. In order to understand the logistics demand of Guangzhou Port in the following years, the combination model is used to further forecast the logistics demand of Guangzhou Port in 2022-2024. The forecast results are shown in the Tables VI to VIII:

TABLE VI. CARGO THROUGHPUT FORECAST FOR 2022-2024

Year	GM(1,1)	BP	Combination
2022	69149.28783	59923.61814	64536.45299
2023	72347.76658	59016.2997	65682.03314
2024	75694.18995	59415.8469	67555.01842

TABLE VII. FOREIGN TRADE CARGO THROUGHPUT FORECAST FOR 2022-2024

Year	GM(1,1)	BP	Combination
2022	15508.99416	13678.9264	14593.96028
2023	16057.18841	13791.26738	14924.2279
2024	16624.75961	13576.82063	15100.79012

TABLE VIII. CONTAINER CARGO THROUGHPUT FORECAST FOR 2022-2024

Year	GM(1,1)	BP	Combination
2022	2686.914378	2034.919902	2360.91714
2023	2852.049687	2122.543896	2487.296791
2024	3027.334061	2082.888479	2555.11127

The three throughputs all show an increasing state in 2022-2024 in the prediction of the GM(1,1) model, while the BP neural network model is slightly different from the GM(1,1) model. However, in general, it is also an increase compared with 2020. After the combination of the two individual models through the equal weight average method, it is predicted that the demand for logistics of Guangzhou port in 2022-2024 will show an increasing trend. Extrapolate according to the actual situation, this is in line with our country's national condition and economic development trend. The combined model reduces the limitations of the single model, and the forecast result is closer to reality, which can be applied to the forecast of Guangzhou port logistics demand.

V. DISCUSSION AND CONCLUSION

A. Conclusion

This study chooses port throughput, foreign trade cargo throughput, and container cargo throughput to study the development trend of Guangzhou port cargo. The combined model is used to forecast the logistics demand of Guangzhou port in 2022-2024. The freight volume demand of Guangzhou port shows a rapid growth trend. The comparative results demonstrate that the BP–GM(1,1) combination model consistently achieves lower absolute percentage errors than the single GM(1,1) and BP neural network models across all three throughput indicators. This indicates that the hybrid approach effectively balances the advantages of small-sample adaptability and nonlinear learning capability. From a practical perspective, higher forecasting accuracy can support more rational allocation of port logistics resources, reduce congestion risks, and improve operational efficiency. Furthermore, the results provide quantitative evidence for smart port planning and infrastructure optimization, offering decision support for port authorities under complex logistics environments.

B. Discussion

From a strategic perspective, accurate logistics demand forecasting can assist port authorities in prioritizing infrastructure investment, optimizing capacity expansion, and coordinating port–city development. Under the Belt and Road Initiative and the Guangdong–Hong Kong–Macao Greater Bay Area framework, such predictive insights are crucial for enhancing the international competitiveness of hub ports and supporting regional economic integration.

C. Future Research

Future research may extend the proposed combination model to other ports or regional logistics systems to test its generalizability. In addition, integrating real-time operational data and intelligent decision-support systems could further enhance its application in smart port governance and policy formulation.

Conflicts of Interest

The authors have declared that no competing interests exist.

Research Involving Human Participants and/or Animals

No applicable.

Funding

This research was supported by Guangzhou Huashang College (Grant No. 2023HSQX091; 2022HSXK02; HS2024SFZY10; HS2024ZLGC16; HS2024ZLGC07; HSRHKC202437; HS2025AIJG04) and the Zhaoqing Philosophy and Social Sciences Discipline Co-construction Project (25GJ-365), Guangzhou Social Sciences "14th Five-Year Plan" Project (2025GZGJ264).

References

[1] Zhuang X., Li W., Xu Y. Port planning and sustainable development based on prediction modelling of port throughput: a case study of the deep-water dongjiakou port. Sustainability, 2022, 14(7): 4276-4276.

[2] Xu S., Zou S., Huang J., et al. Comparison of different approaches of machine learning methods with conventional approaches on container throughput forecasting. Applied Sciences, 2022, 12(19): 9730-9730.

[3] Xiao Y., Xue X., Hu Y., et al. Novel decomposition and ensemble model with attention mechanism for container throughput forecasting at four ports in Asia. Transportation Research Record, 2023, 2677(6): 530-547.

[4] Ramos S.J. Planning for competitive port expansion on the U.S. Eastern Seaboard: the case of the Savannah Harbor Expansion Project. Journal of Transport Geography, 2014, 36: 32-41.

[5] Transportation Policy; Investigators at Soochow University Detail Findings in Transportation Policy (Economic impact of port sectors on South African economy: An input-output analysis). Journal of Transportation, 2014.

[6] Haque M.Z., Joachim S.H. The impacts of port infrastructure and logistics performance on economic growth: the mediating role of seaborne trade. Journal of Shipping and Trade, 2018, 3(1): 1.

[7] Moscoso-López J.A., Turias T.I.J., Come M.J., et al. Short-term forecasting of intermodal freight using ANNs and SVR: Case of the port of algeciras bay. Transportation Research Procedia, 2016, 18.

[8] Dragan D., Keshavarzsaleh A., Intihar M., et al. Throughput forecasting of different types of cargo in the adriatic seaport Koper. Maritime Policy & Management, 2020, 48(1): 19-45.

[9] Abraham E., João R., Colossetti A., et al. Neural network system to forecast the soybean exportation on Brazilian Port of Santos, 2018, IFIP Advances in Information and Communication Technology, 514, pp. 83-90. Springer, Cham.

[10] García T.R., Cancelas N.G., Soler-Flores F. The artificial neural networks to obtain port planning parameters. Procedia-Social and Behavioral Sciences, 2014, 162: 168-177.

[11] Goedhals-Gerber L.L. Predicting the throughput of grain products at the multipurpose terminal at the Port of Cape Town. Journal of Transport and Supply Chain Management, 2016, 10(1): 1-9.

[12] Yan P., Zhang L., Feng Z., et al. Research on logistics demand forecast of port based on combined model. Journal of Physics: Conference Series, 2019, 1168(3): 032116.

[13] Farhan J., Ong P.G. Forecasting seasonal container throughput at international ports using SARIMA models. Maritime Economics & Logistics, 2018, 20(1): 131-148.

[14] Huang F., Liu D., An T., et al. Port container throughput forecast based on abc optimized bp neural network. IOP Conference Series: Earth and Environmental Science, 2020, 571(1): 012068.

[15] Grundmeier N., Hahn A., Ihle N., et al. A simulation based approach to forecast a demand load curve for a container terminal using battery powered vehicles. 2014 International Joint Conference on Neural Networks (IJCNN), Beijing. IEEE. 2014.

[16] Ya B. Study of food cold chain logistics demand forecast based on multiple regression and AW-BP forecasting method on system order parameters. Journal of Computational and Theoretical Nanoscience, 2016, 13(7): 4019-4024.

[17] Liang X, Wang Y, Yang M. Systemic modeling and prediction of port container throughput using hybrid link analysis in complex networks. Systems, 2024, 12(1): 35.

[18] Zeng F, Xu S. A hybrid container throughput forecasting approach using bi-directional hinterland data of port. Scientific Reports, 2024, 14(1): 12345.

[19] Zhang D, Li X, Wan C, et al. A novel hybrid deep-learning framework for medium-term container throughput forecasting: an application to China's Guangzhou, Qingdao and Shanghai hub ports. Maritime Economics & Logistics, 2024, 26(1): 45-67.

[20] Deng J. Introduction of grey system. Journal of Grey System 1989, 1: 1-24.

[21] Zheng B. Material procedure quality forecast based on genetic BP neural network. Modern Physics Letters B, 2017, 31: 19-21.

[22] Tang S., Xu S., Gao J. An optimal model based on multifactors for container throughput forecasting. KSCE Journal of Civil Engineering, 2019, 23(9): 4124-4131.

[23] Yang D. Logistics demand forecast model for port import and export in coastal area. Journal of Coastal Research, 2020, 103(sp1): 678-681.

[24] Zhang Z. Research on the Carbon Emission Diving Factors and Forecasts of Logistics Industry in the Bohai Rim Economic Zone based on the Theory of Grey System. Journal of Management & Decision Sciences, 2020, 3(1): 34-44.

[25] Zhang H., Zhu J., Zhou L. Research on Logistics Demand Forecasting and Transportation Structure of Beijing Based on Grey Prediction Model. Science Journal of Applied Mathematics and Statistics, 2015, 3(3): 144-152.

[26] Wang C., Nguyen H., Liao R., et al. Partner Selection in Supply Chain of Vietnam's Textile and Apparel Industry: The Application of a Hybrid DEA and GM (1, 1) Approach. Mathematical Problems in Engineering, 2017.