Advances in Economics, Business and Management Research, volume 68
International Symposium on Social Science and Management Innovation (SSMI 2018)

Combined Throughput Prediction of Fujian Coastal Ports based on Grey Model and Markov Chain

Yu Wang* and Zhiming Wang
School of Shanghai Maritime University, Shanghai, China
* 961151333@qq.com

Abstract. The rationality and reliability of port throughput forecasting results directly affect the rational arrangement of port resources, which is of great significance to the formulation of future port development strategy. This paper collects decades of historical data of Fujian coastal ports, combines the Grey model GM (1,1) prediction with Markov chain, uses Markov chain to reflect the randomness of variable fluctuations, corrects the predicted value, and forecasts the port cargo throughput from 2015 to 2016. The results show that the average absolute error of the Grey model modified by Markov process decreases from 5.4% to 0.04%. Through comparison, we can find that the result of Grey Markov chain is more accurate than that of single grey prediction.

Keywords: Markov chain; Grey model; port throughput forecasting results; Grey Markov chain.

1. Introduction

Port cargo transport is one of the main modes of modern transportation system. In the foreseeable future, North America, Europe and Asia-Pacific are the main markets of global cargo transport. Asia-Pacific region is expected to be the most important source of commercial trade and activities. Fujian port is facing the challenge of comparative advantage. The cost of port construction is sunk cost. Once the port is built and the equipment is installed, its application will be difficult to change. To avoid waste, port managers must be able to predict container throughput. At present, there are two kinds of methods to predict port throughput: qualitative and quantitative. Among them, qualitative forecasting methods mainly include subjective probability method, scenario forecasting method, source investigation method, etc. Quantitative forecasting methods mainly include regression analysis method, time series forecasting method, grey forecasting method, artificial neural network method, etc.

Peng [1] proposed six univariate models to predict container throughput of three major ports in Taiwan. It was found that the classical decomposition model is usually the best model to predict the seasonal variation of container throughput. Fang [2] proposed a back propagation neural network model based on genetic algorithm. Experiments show that the GA-BP neural network model has better accuracy, but it consumes more time than the traditional BP network model. In order to overcome the shortcomings of railway passenger volume prediction methods, Wang [3] based on the railway passenger volume from 1980 to 1998, used the improved neural network method to forecast the railway passenger volume. The simulation results show that the improved neural network prediction results are more accurate and reliable. Lin Qiang [4] established three grey multivariate regression models by using Grey model and multivariate regression model, and concluded that the series model and embedded model are the improvement of multivariate regression model by using grey theory, which can weaken the random type of original data and improve the prediction accuracy of the model. Parallel model is essentially a combination model, which can synthesize multiple information and has certain practical value. Godfrey [5] applied exponential smoothing method to establish passenger volume prediction model. It was found that this method is relatively simple in practice and has relatively small prediction error.

Based on previous studies, this paper proposes a combined forecasting method combining grey forecasting analysis with cloud model forecasting, which mainly aims at the throughput forecasting of Fujian ports. Firstly, the data of cargo throughput in Fujian Province are collected and analyzed. The grey prediction algorithm is used to get the prediction data. On this basis, the prediction value is revised by combining Markov chain.

2. Prediction Model Selection

2.1 Grey Theory

The grey system theory holds that the prediction of the existing information and the system with unknown or uncertain information is the prediction of the time-related grey process which changes in a certain direction. Although the phenomena shown in the process are random and disorderly, they are orderly and bounded after all, so this data set has potential regularity. Grey prediction is to use this rule to establish a Grey model to predict the grey system.

Grey prediction model is the core of grey theory put forward by Professor Deng Julong. It is especially suitable for the prediction of areas where incomplete information or uncertain behavior are common problems. Grey theory has three basic operations: (1) cumulative generation, (2) inverse cumulative generation, (3) Grey modeling. The characteristic and advantage of Grey model is that it needs less data to predict. The characteristics of Grey model are the order of differential equation and the number of variables involved. For example, the first-order single variable Grey model is often expressed as GM (1,1). Here we briefly describe the steps of calculating GM (1,1) prediction process.

2.2 Modeling Steps of Grey Prediction

Step1: Let the sequence $x^{(0)}$ have n observations.

$$x^{(0)}(1), x^{(0)}(2), x^{(0)}(3), \cdots, x^{(0)}(n) \tag{1}$$

Step2: The first order cumulative generation of $x^{(0)}$ (1-AGO) yields a new sequence $x^{(1)}$.

$$x^{(1)}(i) = \sum_{m=1}^{i} x^{(0)}(m) \qquad (i = 1,2,\cdots,n) \tag{2}$$

$$\begin{aligned}
x^{(1)}(1) &= x^{(0)}(1) \\
x^{(1)}(2) &= x^{(0)}(1) + x^{(0)}(2) \\
&= x^{(1)}(1) + x^{(0)}(2) \\
x^{(1)}(3) &= x^{(0)}(1) + x^{(0)}(2) + x^{(0)}(3) \\
&= x^{(1)}(2) + x^{(0)}(3) \\
&\cdots\cdots \\
x^{(1)}(n) &= x^{(1)}(n-1) + x^{(0)}(n)
\end{aligned}$$

Step3: For the logarithmic sequence $x^{(1)}$, the whitening formal equation of the prediction model can be established.

$$\frac{\mathrm{d}x^{(1)}}{\mathrm{d}t} + ax^{(1)} = u \tag{3}$$

In the formula, a and u are the parameters to be estimated, which are called development grey number and endogenous control grey number respectively. Let $\hat{a}$ be the vector of parameters to be estimated.

$$\hat{a} = \begin{bmatrix} a \\ u \end{bmatrix}$$

Step4: Solving by least squares method.

$$\hat{a} = (B^T B)^{-1} B^T y_n \tag{4}$$

Formula:

$$B = \begin{pmatrix} -\frac{1}{2}(x^{(1)}(1) + x^{(1)}(2) & 1 \\ -\frac{1}{2}(x^{(1)}(2) + x^{(1)}(3) & 1 \\ \cdots\cdots & \\ -\frac{1}{2}(x^{(1)}(n-1) + x^{(1)}(n) & 1 \end{pmatrix} \tag{5}$$

$$y_n = [x^{(0)}(2), x^{(0)}(3), \cdots, x^{(0)}(n)]^T \tag{6}$$

Step5: Turn to Step3 and substitute $\hat{a}$ into its differential equation and solve the differential equation. There is a GM (1,1) prediction model.

$$\hat{x}^{(1)}(i+1) = \left(x^{(0)}(1) - \frac{u}{a} \right) e^{-ai} + \frac{u}{a} \qquad i = 1,2,3 \cdots n-1 \tag{7}$$

Thus, the corresponding predicted value can be obtained.

$$\hat{x}^{(0)}(i+1) = \hat{x}^{(1)}(i+1) - \hat{x}^{(1)}(i) \qquad i = 1,2,3 \cdots n-1 \tag{8}$$

2.3 Grey Markov Theory

State division. According to the prediction results of GM (1,1) model, the relative

errors of the original sequence and the predicted sequence are calculated, and the state partition interval is determined according to the concentration degree of the error range. According to the characteristics of the data, the data are divided into several states. If the degree of deviation of the predicted value from the actual value is expressed, then I can divide all the deviations into M state spaces.

2) Establishment of state transition matrix. After partitioning the state space, we make Mij the probability of transition from state I to state j through M steps. Then we establish the transition probability matrix of the starting M-step state.

3) Calculate the predicted value. The intermediate value of the state interval is selected, and the grey prediction value is corrected by using the Markov state transition probability matrix, and then the Grey Markov prediction value is obtained.

$$y = \frac{X(k)}{1 \pm 0.5 * (L_{ij} + U_{ij})} \tag{9}$$

When the forecast is overestimated, it should be positive, and when it is underestimated, it should be negative.

3. Data Processing

In this paper, the time series data of port cargo throughput in Fujian Province from 2001 to 2016 are used. The data mainly come from Fujian Statistical Yearbook. The data from 2001 to 2014 are used as the data for establishing the model, and the data from 2015 to 2016 are used as the data for forecasting and testing. from 2001 to 2016, The cargo throughput of Fujian's ports from 2001 to 2016 is shown in Table 1, and the trend chart of cargo throughput of ports is shown in Figure 1.

Table 1. Cargo throughput of Fujian coastal ports.

Years	2001	2002	2003	2004	2005	2006	2007	2008
Actual value (10,000 tons)	8278.42	10200.62	12495.48	15834.76	19605.25	23687.61	23602.90	27070.06
Years	2009	2010	2011	2012	2013	2014	2015	2016
Actual value (10,000 tons)	30541.8	32687.01	37278.95	41359.23	45475.19	49166.24	50282.09	50776.09

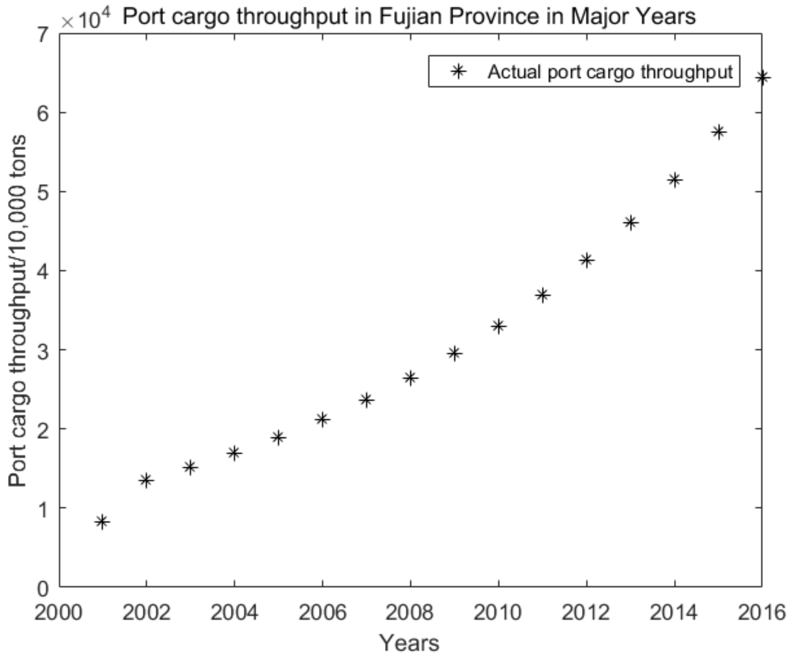

Figure 1. Trends in cargo throughput

4. Building Model

4.1 Gray Model

According to the data of Fujian port cargo throughput scale 1, a GM (1,1) model is established by using grey theory, and the port cargo throughput from 2015 to 2016 is predicted by the data from 2001 to 2014.According to the historical statistical data, we can see that under normal circumstances, the annual average reflects the changing law of the relevant indicators.

From the data table 1, the observed values and the first-order cumulative values calculated by formulas (1) and (2) are respectively:

$$x^{(0)}=(7567,9210,10841,12130,15193,14271,16803,18872,21100,23162)$$

$$x^{(1)}=(7567,16777,27618,39748,54941,69212,86015,104887,125987,149149)$$

From the least square's formula (4), we can get a=-0.1082, u=8474.81. The parameter values are substituted into formulas (3) and (7) to obtain:

$$\frac{dx^{(1)}}{dt} - 0.1082x^{(1)} = 8474.81$$

$$\hat{x}^{(1)}(i+1) = 85892.4159e^{0.1082i} - 78325.4159$$

According to formula (8), the forecast value of GM (1,1) of port cargo throughput in Fujian Province in each year can be obtained, as shown in table 2.

Table 2. GM (1,1) Prediction Value of Port Cargo Throughput in Fujian Province

Years	2001	2002	2003	2004	2005	2006	2007	2008
Actual value (10,000 tons)	8278.42	10200.62	12495.48	15834.76	19605.25	23687.61	23602.90	27070.06
Forecast value (10,000 tons)	8278.42	13572.53	15167.43	16949.74	18941.49	21167.30	23654.65	26434.29
Relative error (%)	0	-33.05	-21.38	-7.04	3.38	10.63	-0.22	2.35
Years	2009	2010	2011	2012	2013	2014	2015	2016
Actual value (10,000 tons)	30541.8	32687.01	37278.95	41359.23	45475.19	49166.24	50282.09	50776.09
Forecast value (10,000 tons)	29540.57	33011.86	36891.06	41226.10	46070.55	51484.27	57534.15	64294.95
Relative error (%)	3.28	-0.99	1.04	0.32	-1.31	-4.71	14.42	26.62

The forecast trend of cargo throughput in port can be obtained by programming with MATLAB software as shown in figure 2.

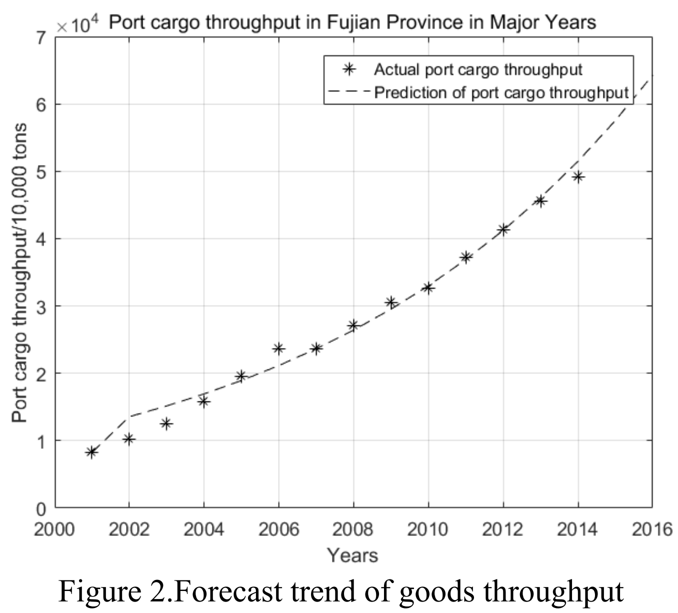

Figure 2.Forecast trend of goods throughput

4.2 Markov Modeling

According to the application experience and actual situation of Markov process, according to the relative error between the actual value of cargo throughput and the Grey prediction value of Fujian port from 2002 to 2014, the threshold of state division is determined by the relative concentration of error range, and the final state division interval is as follows:

E1: $-33<\theta\leq-10$

E2: $-10<\theta\leq1$

E3: $1<\theta\leq11$

The annual status is shown in Table 3.

Table 3. State table.

Years	State	Years	State
2002	E1	2009	E3
2003	E1	2010	E2
2004	E2	2011	E3
2005	E3	2012	E2
2006	E3	2013	E2
2007	E2	2014	E2
2008	E3		

4.3 Constructing State Matrix

According to the state of relative error of each year's prediction results, the state transition matrix is constructed according to the state transition frequency summary table (Table 4).

Table 4. Summary of State Transition Frequencies.

State	E1	E2	E3
E1	1	1	0
E2	0	2	3
E3	0	3	2

One-step transition probability matrix and two-step transition probability matrix are calculated.

$$\mathrm{P}=\begin{Bmatrix} P_{11} & P_{12} & P_{13} \\ P_{21} & P_{22} & P_{23} \\ P_{31} & P_{32} & P_{33} \end{Bmatrix}=\begin{Bmatrix} 0.5 & 0.5 & 0 \\ 0 & 0.4 & 0.6 \\ 0 & 0.6 & 0.4 \end{Bmatrix} \qquad \mathrm{P}^2=\begin{Bmatrix} 0.25 & 0.45 & 0.30 \\ 0 & 0.52 & 0.48 \\ 0 & 0.48 & 0.52 \end{Bmatrix}$$

4.4 Calculate the Predicted Value

The Markov prediction value of port cargo throughput in Fujian Province can be obtained from the state transition matrix and the state of each year. Taking 2003 as an example, because the relative error in 2002 is in the state of E1, after one year's transformation, the probability of conversion to E1 is 1, then it is considered that the relative error in 2003 is most likely in the state of E1, and the forecast value in 2003 is obtained through GM(1,1) model for 15167.43, the Grey Markov model is used to calculate the predicted value for 2003, which is 124,834,800 tons. Using the same one-step state transition matrix method, the Markov predictions for other years can be obtained (Table 5).

Table 5. Prediction comparison table.

Years	Actual value/10,000 tons	Gm (1,1) Model		Grey markov model	
		Predicted value/10,000 tons	Relative error%	Predicted value/10,000 tons	Relative error/%
2003	12495.48	15167.43	-21.38	12483.48	0.00
2004	15834.76	16949.74	-7.04	16219.85	-0.02
2005	19605.25	18941.49	3.38	20150.52	-0.03
2006	23687.61	21167.30	10.63	22518.40	0.05
2007	23602.90	23654.65	-0.22	22636.03	0.04
2008	27070.06	26434.29	2.35	28121.59	-0.04
2009	30541.81	29540.57	3.28	31426.14	-0.03
2010	32687.01	33011.86	-0.99	31590.30	0.03
2011	37278.95	36891.06	1.04	39245.81	-0.05
2012	41359.23	41226.10	0.32	39450.81	0.05
2013	45475.19	46070.55	-1.31	44086.65	0.03
2014	49166.24	51484.27	-4.71	49267.24	0.00

Table 5 gives a comparison of the relative errors between GM (1,1) model and Grey Markov model. In terms of the average absolute error between the actual value and the predicted value, the average absolute error of GM (1,1) model is 5.4%, and the average absolute error of Grey Markov model is 0.04%. the prediction effect of Grey Markov model is better than that of GM (1,1). Therefore, the Grey Markov model is chosen. The model predicts the cargo throughput from 2015 to 2016.

Table 3 shows that 2014 is in E2 state. Then the state probability distribution in 2015 is as follows:

$$P(1)=P^{0}*P=(0,1,0)*\left\{\begin{matrix} 0.5 & 0.5 & 0 \\ 0 & 0.4 & 0.6 \\ 0 & 0.6 & 0.4 \end{matrix}\right\}=(0,0.4,0.6)$$

According to the state probability distribution in 2015, its state is most likely to be in E3, and the average relative error of event throughput in 2015 can be calculated.

$$\theta=0.06$$

According to formula (9), the revised throughput in 2015 is 5427.75 million tons. The forecast results are shown in table 6:

Table 6. Prediction table.

Forecast year	2015	2016
Grey Markov Prediction Value/10000 Tons	54277.5	60655.61

5. Conclusion

According to the grey GM (1,1) preliminary forecast of the port cargo throughput data of Fujian Province from 2002 to 2014, the Markov model is applied to revise it, and the average absolute error of the revised model is reduced from 5.4% to 0.04%. It shows that the Grey Markov prediction model is more accurate and effective than the Grey model, and it can accurately reflect the actual situation, to provide a reference basis for determining the future development trend of port cargo throughput in

Fujian Province. However, because the state interval division is randomly selected, the prediction results will change with the different state intervals, so the latter is the case. In the follow-up study, it is necessary to consider the diversity of state intervals and other indicators to better improve the prediction methods and improve the prediction accuracy.

References

[1]. Wen-Yi Peng, Ching-Wu Chu. A comparison of univariate methods for forecasting container throughput volumes, J. Mathematical and Computer Modelling, 2009, 50: 1045-1057.

[2]. Fang-Feng Ping, Fang-Xue Fei. Multivariant forecasting mode of Guangdong province port throughput with genetic algorithms and Back Propagation neural network, J. Procedia - Social and Behavioural Sciences, 2013, 96: 1165-1174.

[3]. Wang Zhuo, Jia Li-Min, Qin Yong, Wang Yan-Hui. Railway passenger traffic volume prediction based on neural network, J. Applied Artificial Intelligence, 2007, 21(1): 1-10.

[4]. Qiang Lin, Yi-Mei Chen. Grey Multiple Regression Model Applied for Port Throughput Forecasting, J. Port Waterway Engineering, 2008(07): 77-80.

[5]. Godfrey G A, Powell W B. An adaptive dynamic programming algorithm for fleet management, Ⅱ:multi-period travel times, J. Transportation Science, 2002, 36(1): 40-54.