Analysis of Grey Model for Container Traffic Forecasting at Indian Major Ports

Vipulesh Shardeo and Anchal Patil
Department of Management Studies
Indian Institute of Technology Delhi
Hauz Khas, Delhi, India
vipuleshshardeo@gmail.com, patil.anchal@gmail.com

Ashish Dwivedi
Jindal Global Business School
O. P. Jindal Global University
Sonepat, Haryana, India
ashish0852@gmail.com

Jitender Madaan
Department of Management Studies
Indian Institute of Technology Delhi
Hauz Khas, Delhi, India
jmadaan@dms.iitd.ac.in

Abstract

In the present era of globalization, the transportation system needs to be more flexible to meet customer demand. The port system has significant contributions to the fluent operations of the freight transportation system. With the increase in transport demand, the port faces multiple complexities such as lower utilization of resources, capacity issues, etc. The forecasting of container traffic at the port would help the port planning team and managers analyze the port system's infrastructural investment and optimization. In this study, the annual data of the past 20 years (1999-2019) of container traffic in TEUs at three Indian major ports have been considered. The grey forecasting model and Auto-Regressive Integrated Moving Average (ARIMA) model is developed to analyze the container traffic data. Further, the Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE) is used to test the model's accuracy. The results of the study show that both models are fit to forecast the container traffic data. Precisely, the grey forecasting model fits better than ARIMA for two ports Tuticorin and Cochin, with MAPE of 10.52% & 4.80%, respectively. The findings of this study would guide the practitioners and planning managers in decision-making related to port optimization.

Keywords
Container Forecasting, ARIMA, Grey model, Freight transportation

1. Introduction

In the era of globalization, containerization played a vital role in freight transportation by reducing transportation time and cost. This transformation forced major ports to adopt a container-based transport system to improve port efficiency as well. Initially, the term "Intermodalism" was penetrated in the late 1960s when significant challenges of high turnaround time have been faced. Such complexities of the maritime sector resulted in containerization and then
mechanized handling of the goods. Containerization helped a lot in minimizing the turnaround time of freight in the maritime industry. After the containerization, the focus was given towards the integration of different modes, as different modes of transportation have their advantages and limitations, like for shorter distance road transportation while for average distance rail transportation whereas for long-distance maritime transport is usually preferred (Rodrique & Slack, 2017). Such a combination of two or more modes of transportation in a single loading unit is termed as Intermodal Transportation (OECD, 2003). Containerization also helped to promote Intermodal freight transportation by providing container as an Intermodal Transport Unit. With the increase in container transport demand, the complexities also evolve in reducing port performance, lower utilization rate, etc. (Mo et al., 2018). The container throughput, being an essential indicator of the port performance, has gained enormous attention among the decision-makers. The introduction of emerging Information and Communication Technologies also transformed the freight transportation system resulted in the reduction of turnaround time (Patil, Shardeo, & Madaan, 2020; Shardeo, Patil, & Madaan, 2020). However, the present uncertain demand fluctuations create a massive loss to the transport companies. Such issues urge the transportation system to be more flexible to deal with uncertainties. Hence, efficient container demand forecasting is needed to make effective and efficient decisions. Better and accurate container forecasting would provide volume flexibility to the freight transportation system and keep aware of responding quickly (Shardeo, Dwivedi, & Madaan, 2020).

In recent years, container forecasting is gaining colossal attention among data analysts and academicians. Two types of forecasting methods are widely used to forecast the container demand: causal and time series-methods. The literature suggests that the causal methods require massive geographical and economic data and therefore suitable for long term forecasting. In contrast, the time series method requires past data to predict the future and ideal for short term and medium-term forecasting (Grifoll, 2019). The time series methods have been widely used in container forecasting (Gao, Luo, & Zou, 2016; Schulzea & Prinzb, 2009; Y. Zhang, Fu, & Li, 2020; Z. Zhang, 2020). Different forecasting methods include regression-based models, neural network models, exponential smoothing, grey models, etc., have been used in the forecasting studies. The present study adopts the grey forecasting model and ARIMA model to forecast the annual container traffic at Indian major ports. The study's objective is to provide an efficient and accurate forecasting model that can be used to forecast the container traffic at ports. The rest of the paper is organized as: the next section presents the research background followed by a brief description of the methodologies. Further, the results of the study are presented and discussed. The final section presents the conclusion of the study.

2. Research Background

Container forecasting has gained enormous attention among data analysts and researchers due to the importance of port and freight transportation. Several forecasting methods have been applied in academics which include causal, qualitative and time series methods. Schulzea & Prinzb (2009) examined the container transshipments at German ports using Seasonal Auto-Regressive Integrated Moving Average (SARIMA) and Holt-Winters exponential smoothing approach. Their results showed that the SARIMA model forecasted better than the exponential smoothing approach. Peng & Chu (2009) compared different univariate forecasting methods used to forecast container throughput. The study examined the forecasting models by using Taiwan's three major ports container throughput data. The study suggested that the first step to choosing the best forecasting method is to examine the distribution or pattern of the data. The study of Gosasang, Chandraprakaikul, & Kiattisin (2011) compared the traditional and neural network forecasting methods to forecast the container throughput of Bangkok's port. The study used RMSE and MAE as parameters to check the accuracy of the forecasting models. The results found that the Multilayer Perceptron (MLP) neural network is a better forecasting technique than linear regression. Patil & Sahu (2016) developed both univariate and multivariate regression models to forecast container demand at Mumbai port. The study results suggested that the multivariate regression methods provide better results than the time series models. Yang & Yujie (2018) applied a multiple regression model and Back Propagation (BP) neural network to forecast container throughput at Shanghai Port. The results of their study found that the neural network model fits better than the linear regression model. Zhang et al. (2020) further combined ARIMA and BP neural network to enhance the forecasting model accuracy. They tested their model on the container throughput data of Qingdao port and found the integrated model more accurate than other prediction models. Mo et al. (2018) proposed a hybrid forecasting model to forecast the container throughput of China's two major ports. The developed model was based on the group method of data handling (GMDH) neural network. The results showed the proposed hybrid model was efficient enough to forecast than a traditional neural network. Zhang (2020) established two linear regression models with exponential smoothing and linear regression.
to forecast the container throughput of Shenzhen port. The combination of principle based on least square and validity is used in the study, and the results presented significant results compared to other linear regression-based models.

Based on the literature survey in container forecasting, it is observed that time series forecasting methods are found to be more attractive and accurate. However, most of the studies focused on regression, moving average, neural network, etc. The survey found fewer studies on grey forecasting model in the field of container forecasting. However, grey forecasting is widely used in other domains such as electricity consumption, natural gas demand, etc. (Kharista, Permanasari, & Hidayah, 2015; Liu, Wang, Wang, & Liu, 2016; Tsai et al., 2016; X. Yang, Zou, Kong, & Jiang, 2018). Also, studies on forecasting the container traffic that too in an Indian context are limited. This study aims to develop a grey forecasting model to forecast the container traffic at Indian major ports to fill these gaps.

3. Methodology

The section presents the methodology and data used to conduct this study. As mentioned in the earlier section, the study's main objective is to provide an efficient forecasting model to forecast the container traffic at major ports. The data of container handling of the past 20 years at three Indian major ports, namely, Kolkata, Tuticorin and Cochin, have been considered for the study. The data is extracted from the official website of the Indian Ports Association (Indian Port Association, 2020). The container handling data includes both the import and export containers handled at the port. The distribution of the data year-wise at abovenamed three major ports is shown in Figure 2. Further, the data has been categorized into training and testing data set with a ratio of 70:30. 70% of data is considered to train model, and the rest 30%, i.e. from 2013-14 to 2018-19 is considered for the testing model. The two forecasting models, namely, the Grey forecasting model, more precisely grey model with a first-order and single variable (GM (1,1)) and ARIMA model, are used in this study. To test the forecast accuracy, the two most used parameters, RMSE and MAPE, are used. In forecasting studies, these two parameters are commonly used to test the accuracy (Gosasang et al., 2011; Ho, 2012). Figure 1 depicts the flowchart of the research methodology used in this study.

3.1. Grey model

![Figure 1. Research Methodology](image-url)
Deng (1982) developed a grey system theory to solve the problem containing inaccurate or vague data. The system can be categorized into three categories: black, white and grey. The black system is those which cannot be recognized. While the white system are those which can be recognized. The grey system is that system that is neither black nor white and lies somewhere between both having vague or imprecise data. Grey forecasting is a part of the grey system theory and is denoted by G (n,m). In G (n,m), 'n' denotes differential order, and 'm' denotes the number of variables. The Grey model, with first-order and single variable or commonly known as GM (1,1) is used in this study. It is only used with positive data (Deng, 1989). The steps of GM (1,1) are shown below:

**Step 1:** Initiate the original sequence

\[ x^{(0)} = (x^{(0)}(1), x^{(0)}(2), x^{(0)}(3), ..., x^{(0)}(n)) \]

where, \( x^{(0)}(i) \) is the time series data at time \( i \), \( n \) is the length of sequence and \( n \geq 4 \).

**Step 2:** Using accumulating generating operation (AGO), the original sequence \( x^{(0)}(i) \) is transformed into \( x^{(1)}(i) \) and can be defined as:

\[ x^{(1)} = AGO [x^{(0)}] = \sum_{k=1}^{n} x^{(0)}(i), \quad k = 1,2,3, ..., n \]

obtained data sequence is \( x^{(1)} = (x^{(1)}(1), x^{(1)}(2), x^{(1)}(3), ..., x^{(1)}(n)) \)

**Step 3:** Using differential equation, the sequence is modelled as:

\[ \frac{dx^{(1)}}{dt} + ax^{(1)}(t) = b \]

It can be further rewritten as:

\[ x^{(0)}(k) + a z^{(0)}(k) = b, \quad 2 \leq k \leq n \]

where, \( z^{(1)}(k) = a x^{(1)}(k) + (1-a) x^{(1)}(k-1) \), \( 2 \leq k \leq n \)

**Step 4:** After finding the values of \( a \) and \( b \), calculate the time response sequence as:

\[ \hat{x}^{(1)}(k + 1) = \left( x^{(0)}(1) - \frac{b}{a} \right) e^{-ak} + \frac{b}{a}, \quad 2 \leq k \leq n \]

**Step 5:** Using inverse accumulated generating operation (I-AGO), find the predicted value as:

\[ \hat{x}^{(0)}(k) = \hat{x}^{(1)}(k) - \hat{x}^{(1)}(k - 1), \quad 2 \leq k \leq n \]

### 3.2. ARIMA model

Box and Jenkins popularized the Autoregressive Integrated Moving Average (ARIMA) model and was also referred to as the Box and Jenkins model. The prediction done by the model is based on the past values, mistakes of the past values and current values (Kharista et al., 2015). It only uses a stationary time data series. The ARIMA (p,d,q) model is used in this study to forecast the container traffic. In the ARIMA (p,d,q) model, the term 'p' denotes the Auto Regressive order, the term 'd' denotes the Integrated order, and the term 'q' denotes the Moving Average order. The best models for the Kolkata, Tuticorin and Chennai ports were found to be ARIMA (1,2,1), ARIMA (2,2,0) and ARIMA (1,3,1), respectively. The results with accuracy measures are presented in the next section.

### 3.3. Forecasting Accuracy Measurement

Several parameters are available in the literature that has been used to test the forecasting model's accuracy. Among these, the Root Mean Square Error (RMSE) and Moving Absolute Percentage Error (MAPE) are commonly used to measure the accuracy. These two methods are often used in studies related to forecasting. The lower values of RMSE depicts the better forecasting of the data with historical data. Similarly, the MAPE compares the accuracy of methods in percentage form. Table 1 shows the graded scale for the preference of the forecasting models in terms of MAPE.

<table>
<thead>
<tr>
<th>MAPE</th>
<th>Preference</th>
<th>10% - 20%</th>
<th>20% - 30%</th>
<th>&gt;50%</th>
</tr>
</thead>
<tbody>
<tr>
<td>≤ 10%</td>
<td>High</td>
<td>Good</td>
<td>Ration</td>
<td>Not Good</td>
</tr>
</tbody>
</table>

### 4. Results and Discussion

The annual container traffic data of the past 20 years (1999-2019) of three Indian major ports, namely, Kolkata, Tuticorin and Cochin, is considered in this study. The year-wise distribution of the container traffic data is shown in Figure 2.
The actual container traffic data of three aforementioned Indian major ports is divided into two parts with a ratio of 70%: 30%. 70% of the actual data is used to train the model, and the remaining 30% is used to test the model. The GM (1,1) model is used for forecasting as per the steps described in an earlier section. Table 2 presents the forecasted container traffic at three Indian major ports. It is evident from the RMSE & MAPE values of the forecasting model that the GM (1,1) model is accurate enough to forecast the container traffic data.

Table 2. Forecasting results of GM (1,1).

<table>
<thead>
<tr>
<th>Year</th>
<th>Actual Data</th>
<th>Forecasted Data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Kolkata</td>
<td>Tuticorin</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1999-2000</td>
<td>147</td>
<td>137</td>
</tr>
<tr>
<td></td>
<td>2000-01</td>
<td>138</td>
</tr>
<tr>
<td></td>
<td>2001-02</td>
<td>98</td>
</tr>
<tr>
<td></td>
<td>2002-03</td>
<td>106</td>
</tr>
<tr>
<td></td>
<td>2003-04</td>
<td>123</td>
</tr>
<tr>
<td></td>
<td>2004-05</td>
<td>159</td>
</tr>
<tr>
<td></td>
<td>2005-06</td>
<td>203</td>
</tr>
<tr>
<td></td>
<td>2006-07</td>
<td>239</td>
</tr>
<tr>
<td></td>
<td>2007-08</td>
<td>297</td>
</tr>
<tr>
<td></td>
<td>2008-09</td>
<td>302</td>
</tr>
<tr>
<td></td>
<td>2009-10</td>
<td>378</td>
</tr>
<tr>
<td></td>
<td>2010-11</td>
<td>377</td>
</tr>
<tr>
<td></td>
<td>2011-12</td>
<td>412</td>
</tr>
</tbody>
</table>

Figure 2. Year-wise distribution of container traffic at three Indian major ports
Figure 3 compares GM (1,1), and ARIMA models on Kolkata, Tuticorin, and Cochin port forecasted data. As it can be clearly seen in figure 3, the GM (1,1) model outperforms the ARIMA model for Tuticorin and Cochin port data. In contrast, the ARIMA model performs better for Kolkata port data. The same can also be seen from the RMSE & MAPE values shown in Table 3. The highlighted values in Table 3 shows a better result compared to other models. Based on the analysis, it can be said that the Grey model is accurate enough to forecast container traffic at ports.

![Diagram of Cochin, Tuticorin, and Kolkata ports showing comparison between GM (1,1) and ARIMA models.]

**Table 3. The forecasting accuracy measures of GM (1,1) and ARIMA.**

<table>
<thead>
<tr>
<th>Major Ports</th>
<th>RMSE</th>
<th>MAPE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GM (1,1)</td>
<td>ARIMA</td>
</tr>
<tr>
<td>Kolkata</td>
<td>125.2912</td>
<td>36.16457</td>
</tr>
<tr>
<td>Tuticorin</td>
<td>53.69760</td>
<td>122.77306</td>
</tr>
<tr>
<td>Cochin</td>
<td>22.10950</td>
<td>60.27770</td>
</tr>
</tbody>
</table>

Proceedings of the 11th Annual International Conference on Industrial Engineering and Operations Management Singapore, March 7-11, 2021

© IEOM Society International 1019
5. Conclusion
The forecasting of container traffic with improved accuracy is essential for efficient port operations and management. This study used and compared the two most widely used grey forecasting and ARIMA to forecast the container traffic at three Indian major ports. The container traffic data of the past 20 years of three Indian major ports, namely, Kolkata, Tuticorin and Cochin, is considered to conduct the study. The results show that the grey forecasting model outperforms the ARIMA model for the container data of Tuticorin and Cochin port with RMSE of 53.70 & 22.10 and MAPE of 10.51 & 4.80, respectively. The ARIMA (1,2,1) model for Kolkata port fits better than GM (1,1) with RMSE of 36.16 and MAPE of 5.24, respectively. However, both models are accurate better in terms of the MAPE table. Hence, it can be concluded that the grey forecasting model is efficient enough to forecast container traffic data at Indian major ports. The developed model would help port managers and planners to make decisions related to capacity and resource requirement, container schedules, etc. Also, the present study possesses some limitations too. The grey forecasting model is used for data of the last 20 years of container traffic (in TEUs) at three Indian major ports. The results may vary when more past data of more ports are included. Also, the rolling mechanism could be incorporated along with the grey forecasting model to increase the accuracy of the model. Also, some other forecasting methods could also be used along with the grey forecasting model to make a hybrid model. Such ideas could be taken as future research suggestions.

Acknowledgements
The authors would like to thank the Indian Institute of Technology Delhi Research Committee for financial and technical support.

References

© IEOM Society International 1020


**Biographies**

**Vipulesh Shardeo** is a Research Scholar at Department of Management Studies, Indian Institute of Technology Delhi, Delhi, India. His area of research is study on flexibility focused Intermodal Freight Transportation for Performance Improvement. He has done Master of Technology majoring Industrial Engineering and Management from Indian Institute of Technology (Indian School of Mines) Dhanbad, Dhanbad, India. He has also been involved in several projects related to performance improvement models, manpower assessment and fleet optimization projects. He has interests in topics related to Flexibility, Mathematical modelling, Optimization, Intermodal Freight Transportation, and Supply Chain Management.

**Anchal Patil** is a doctoral candidate at the Indian Institute of Technology (IIT) Delhi in the Department of Management Studies (DMS). He is working under the supervision of Prof Jitendra Madaan on improving performance aspects of Humanitarian Supply Chain. Before joining DMS he worked as a business analyst in an MNC. He received his graduation degree in Industrial Engineering and Management from NIT Kurukshetra, Haryana and post-graduation in Industrial Engineering from IIT Delhi. His current research interest includes Operation research, simulation, optimization and disaster relief work.

**Ashish Dwivedi** works as an Assistant Professor in the area of Operations Management and Decision Sciences at Jindal Global Business School, O.P. Jindal Global (Institution of Eminence Deemed to be University), Sonipat. He holds a doctorate (PhD) from Department of Management Studies, Indian Institute of Technology Delhi. He holds an MTech in Mechanical Engineering with specialization in Computer Aided Design and Manufacturing from Motilal Nehru National Institute of Technology Allahabad, Prayagraj. He is an engineering graduate in Mechanical Engineering. His research has been published in various journals of international repute including Journal of Cleaner
Jitender Madaan is an Associate Professor at Department of Management Studies, Indian Institute of Technology Delhi, Delhi, India. Dr. Jitender Madaan received his B.Tech Degree in Production and Industrial Engineering from MBM Govt. Engg College, Jodhpur (JNV University), India, and obtained his M. Tech in Manufacturing System Engg. from Department of Mechanical Engg. MREC (Now MNIT), Jaipur and PhD in Mechanical Engineering from the Indian Institute of Technology (IIT) Delhi, India. Prior joining IIT Delhi, Dr. Madaan has many years of working experience in other universities including IIT Roorkee; GGSIP University Delhi. His current research interests are Reverse Logistics and Supply Chain Management, Sustainable Operations Management, Production Management, information and governance effectiveness, Systems Modelling and Simulation, etc. To date, Dr. Madaan has published over 4 book chapters, over 25 refereed international journal papers and 48 peer reviewed international conference papers. He is a reviewer of several international journal of repute.