Forecasting Container Throughput Volume of the Dar es Salaam Seaport based on Several Forecasting Models

Msabaha Juma Mwendapole*, Jin Zhihong**

* PhD Candidate, Dalian Maritime University, Transportation and Management College, 116026 No. 1 Linghai Road, Dalian, China
** Professor, Dalian Maritime University, Transportation and Management College, 116026 No. 1 Linghai Road, Dalian, China

DOI: 10.29322/IJSRP.11.10.2021.p11806
http://dx.doi.org/10.29322/IJSRP.11.10.2021.p11806

Abstract - The seaport of Dar es Salaam is currently in the third position uppermost container throughput volume in East Africa. It aspires to become the entire maritime hub for the regions of East and Central Africa, and to be one of the most developed seaports in Africa in the future. In this paper various computer software’s such as MATLAB, SPSS, STARTA and Excel were applied together with several prediction models to precisely forecasting the Dar es Salaam seaport container throughput volume. Those models are exponential smoothing model, grey model, quadratic model and a combined forecast model. Comparison was made based on their Mean Absolute Percentage Error (MAPE) after the computation of each model. Finally, models results depicted that the combined forecast model has got the smallest number of errors. Hence the study confirmed that it has the best data for this research work, making it the best predicting model for the Dar es Salaam seaport container throughput volume.

Index Terms - Container Throughput, Seaport, MAPE, Combined Forecasting Model

I. INTRODUCTION

Tanzania is a maritime state with estimates population of 60.1 million people [1]. The principal seaport in Tanzania is Dar es Salaam seaport and is set up by the Act of 2004 of Tanzania Ports Authorities. It serves the Tanzania hinterland and other nations from Central and East Africa [2]. Like other major seaport in Africa, Dar es Salaam seaport at present is facing a challenge of very big container traffic, this is because demand exceeds the supply. This situation accelerated by the increased population, trade volume growth, GDP growth of Tanzania and country’s political stability. These making the increase of demand level which resulting to high congestion and poor efficiency of the Dar es Salaam seaport [3]. Forecasting its container throughput is significant as it will help in developing, deciding, managing, planning, and advancing seaport facilities and commenced the preparation for constructing adequate seaport infrastructure. Definitely, the success of any seaport policy depends on precisely forecasted container throughput [4]. Seaport projects designs and investment planning depends on the forecasted traffic, traffic to be handled deciding the quantity and size of the investment [5]. A proper forecasting model is necessary when forecasting seaport container throughput. The grey model, a combined forecast model, quadratic and exponential smoothing models are appropriate and they are used in this article to identify the model with the highest prediction accuracy so that to be used in anticipating container traffic at the Dar es Salaam seaport.

II. LITERATURE REVIEW

Tanzania is an East African country; its economic busy city is Dar es Salaam while the capital city is Dodoma. Tanzania is very close to the Indian Ocean as the reason the Dar es Salaam seaport was constructed as depicted in Figure 1 below.
International trade depends much on maritime transport, more than 85% of world trade volume is moving through maritime transport and is handled by seaports globally [6]. Topics associated with the prediction of seaport container throughput volume have attracted the attention of many researchers and experts due to the rapid growth of numerous nations’ economies globally, also a major increase in demand for maritime transportation over the last 20 years due to the speedy of globalization [7]. As a result, many of them have conducted extensive research and published numerous articles conducted either quantitatively or qualitatively and various forecasting models were applied to predict future container throughput volume of the entire seaport. Despite the fact that no prior study conducted concerning container throughput volume forecasting at the Dar es Salaam seaport. They generally emphasized on various aspects and criteria that must be considered and can directly influence or affect a country’s seaports container throughput volume.

Gamassa and Yan (2017) they conducted forecasting study using several models to predict the Abidjan port container throughput volume. They used MAPE to choose the best model to be applied in predicting the Abidjan port container throughput volume, their study finding confirmed that model of double exponential smoothing is the accurate one for anticipating the Abidjan port container throughput volume after having the lowest MAPE [8]. Lam et al 2004, they forecasted and analysed the reliability of cargo throughput in Hong Kong port. They used input data from 1983-2000 to develop a new forecasting model to fit for Hong Kong port, their paper proposed the neural network models for predicting Hong Kong port cargo throughput. Reliability was assessed using simulation done by Montle Carlo; finally study findings confirm neural network models are more consistent, more conservative, and more analogous to genuineness [9]. Tsioumas et al 2017, they forecasted bulk freight market using novel approach. They aimed to develop VARX so that to improve accuracy in forecasting. Their findings showed that the ARIMA approach outperformed by the VARX model. The study is much interested to maritime experts, as it offers valuable understandings into the merchandise market's trajectory and enables them to make well-informed judgments [10].

Zhand et al 2013, conducted study for the aim of developing a grey-logistic growth curve hybrid forecasting model, so that to enhance cargo throughput forecasting for the seaport. Existing seaport data were used to authenticate the validity of the developed hybrid model. The outcomes of a test using perception data of cargo flow in a physical seaport showed that the hybrid model may produce significantly higher forecast correctness when more information is difficult to collect. Furthermore, the hybrid model's forecast is more precise than any of the individual ones [11]. Wang and Ghalih 2017, they wrote an article focused on predicting the
domestic consumption of coffee in Indonesia using grey differential model. By considering the calculated results, the Grey forecast model’s average residual error was greater than 5%. The model indicated the rise of aggregate consumption year after year. Grounded on the testing results, the suggested method appears to not only increase the predicting correctness of the genuine Grey models, but also offered a helpful benchmark for Indonesia coffee enterprises and farmers in developing their future action plans [12].

Ansari Saleh Ahmar, 2018 he analysed the Sutte gauge to check for its suitability in forecasting stock movement. He used data from Indonesia stock exchange for ten years period from 2006 to 2016. Simple moving average and moving average divergence/convergence were used as other technical analysis to check the Sutte Indicator performance. To examine the amount of predictability MAPE, MSE and MAD, were used to compare the stock data. According to the findings of this study, the Sutte Indicator applied as a benchmark in forecasting stock trends, and when matched to other gauge means (MACD and SMA) via MAPE, MSE and MAD, the Sutte Indicator found to have higher reliability level [13]. Phan Van Thanh 2016 he suggested two effective hybrid grey models, namely the Fourier Nonlinear grey Bernoulli Model and Fourier grey Model (1, 1) to be used in forecasting the demand for international tourism as well as to defeat the limitations of previous grey models. In terms of situations, both out-of-sample and in-sample, the empirical outcomes showed the correctness of both the NGBM (1, 1) and GM (1, 1) predicting models were higher than the previous ones after updating their residual error using Fourier series. Furthermore, his article found that the F-GM (1, 1) model outperformed various forecasting models used in predicting international visitors to Vietnam, with average MAPEs of 0.013 percent and 5.19 percent in-sample and out-of-sample, respectively [14].

Bogna et al 2012 they forecasted sales demand or volume for transport service so as to determine the size and quality of transport needs. In their article they compared three predicting methods namely the artificial immune system, a multiplicative version, and harmonic analysis. MAPE was used to determine the effectiveness of forecasting method [15]. Dejan et al, 2014, they addressed three methods for predicting container throughput which have been adopted by port authorities in Ravenna, Trieste and Venice (Italy), North Adriatic ports Koper (Slovenia), Forecasting methods include the autoregressive integrated moving-average (ARIMA), the classical decomposition and Holt-Winters exponential smoothing models. Ten years period data of container throughput were used to test all the models. Models were compared based on the MAPE, MAD and RMSE. ARIMA model showed the best performance over the others and decomposition model outperformed by Holt-Winters model [16]. SHU et al 2014, they aimed to find the effective predicting models for the estimation of seaport cargo throughput. Residual Fourier-modified models, Grey models and traditional ARIMA models were comprehensively compared. Finally a residual fourier-modified model was revealed to be meaningfully boosted [17].

Van-Thanh Phan and Chia-Nan Wang 2014, they conducted cargo throughput forecasting from 2013-2015 based on previous data of Kaohsiung seaport. They collected data from Taiwan ministry of transport and communication. Study confirmed that both grey models are useful but FRMGM (1, 1) is the outstanding model in forecasting Kaohsiung seaport cargo because it possesses 100% predicting average accuracy, therefore was strongly recommended [18]. Shouwei XIE and Yadong YANG, 2019, they adopted wavelet analysis, neural network and genetic algorithm to design the forecasting models namely BP neural network and genetic wavelet neural network. They used Jujiang seaport ship volume as data for the experiment for simulation and analysis. In terms of prediction accuracy, the data showed that the genetic wavelet neural network forecasting model performed better also it showed to have vast application potential in the Yangtze River port for predicting its ship traffic flow. In terms of practical application, this method is critical [19]. Jiaqi Hou et al 2015, they proposed an anticipating model established on the Pearl Curve, the Exponential Smoothing and the GM (1, 1). Dalian seaport was used as a case study; data were obtained from the website of the Ministry of communication in China archived from 2002 to 2013. The finding obtained from the hybrid model was compared with the first outcomes obtained from single models. The findings showed that a hybrid model is efficient and feasible because it possessed lower fraction errors and more smother than single model [20].

III. METHODOLOGY OF THE SELECTED FORECASTING MODELS

For the aim of getting the best precisely model for the prediction, a combined, the exponential smoothing, grey, and quadratic prediction models are presented.

A. The Exponential Smoothing Model

This is the method of predicting time series data. Robert Goodell Brown proposed exponential smoothing in the arithmetic works without referencing earlier work in 1956, and in 1957 Charles C. Holt built on it. Exponential smoothing is a general-purpose method which uses the exponential window function to smooth out time series data. The smoothing factor or smoothing constant adjusting the exponential smoothing calculations input [21]. Exponential Smoothing is applied globally to forecast and smoothing time series data, also it has found achievement as a predicting technique since it is easy to apply, extensive in adaptability, simple to understand, low in cost, excellent performance and, in many cases, quite accurate [22]. The exponential smoothing model expression given as:

\[ s_t = \alpha x_t + (1 - \alpha)s_{t-1} = s_{t-1} + \alpha(x_t - s_{t-1}) \]  

(eqn. 1)

Whereby,

- \( s_t \) = smoothed figure, it is the current observation's simple weighted average. \( x_t \)
- \( t \) = time period
- \( s_{t-1} \) = previous smoothed figure

This publication is licensed under Creative Commons Attribution CC BY.  
http://dx.doi.org/10.29322/IJSRP.11.10.2021.p11806  
www.ijsrp.org
\[ \alpha = \text{smoothing factor of data; } 0 < \alpha < 1 \]

**B. The Grey Model:**

Traditional Grey Forecasting Model with numerous benefits and a straightforward calculated ground is GM (1, 1) \[ [23, 24, 25] \]. It is broadly prediction model used to establish a dynamic model for the creation of abstract systems after getting differential equations which are converted directly from time series data \[ [26] \]. The Grey Model is an essential method for forecasting small data sets \[ [27, 28] \]. The Grey system philosophy is a significant branch in research, and has been applied in financial, industrial, economic, social, agricultural and other areas \[ [29, 30] \]. The Grey model's foundation is outfitted with the original sequence \[ [31] \].

Let \[ X^{(0)} = (x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(n)) \], \( x^{(0)}(k) \geq 0 \),

(eqn. 2)

\[ X^{(1)} = (x^{(1)}(1), x^{(1)}(2), \cdots, x^{(1)}(n)) \],

(eqn. 3)

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

(eqn. 4)

Then

\[ x^{(0)}(k) + ax^{(1)}(k) = b \]

(eqn. 5)

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

(eqn. 6)

Then the GM (1, 1) formula will be

\[ X^{(0)}(k) = x^{(1)}(k) - x^{(1)}(k-1) \]

(eqn. 7)

Sequence \( X^{(0)}(k) \) is the computed figure of the genuine sequence, for this case, it is the original data for the container throughput of the Dar es Salaam seaport.

**C. The Quadratic Model**

This is mathematical model expressed by a quadratic equation or a series of quadratic equations, such as \( Y = aX^2 + bX + c \). When displayed on a graph, the connection between the variables in a quadratic equation is a parabola. Quadratic time series analysis is used to forecast data with a trend but no seasonal component. Furthermore, the trend exhibits curvature, indicating that it is not a linear trend. The quadratic model is as follows:

\[ Y_t = b_0 + b_1 t + b_2 t^2 \]

(eqn. 8)

Whereby,

- \( Y_t \) = the value of the data at time \( t \)
- \( b_0 \) = a constant
- \( b_1 \) and \( b_2 \) = the coefficients for time

**D. The Combined Model**

At the moment, this model has made significant progress also is being used broadly over the world \[ [31] \]. Assume “\( m \), “\( n \)” is constant, “\( m \)” is the number of approaches to predict cargo throughput, and “\( n \)” represents the number of time periods we employ. \( y_{ij} \) \( i = 1,2,3,\ldots, \quad n, \quad j = 1,2,3,\ldots, \quad m \) represents the predicted figure in the "\( i \)” approach and the 1 "\( i \)” time period, and “\( r_j \)” represents the weight of the various prediction ways, and it should satisfy the criteria \[ [31] \]. Presume “\( i \)” as time period and \( Y_{i0} \) is the figure for a combined forecasting model, hence

\[ Y_{i0} = \sum_{j=1}^{m} r_j y_{ij} \]

(eqn. 9)

**IV. DATA ANALYSIS**

The data for this study were gathered from the World Bank's official website \[ [32] \] and span the years 2010 to 2019.
Table 1. Dar es Salaam Seaport Container Throughput, annual, 2010-2019 Measure: TEU (Twenty-foot Equivalent Unit) [32].

<table>
<thead>
<tr>
<th>Year</th>
<th>Dar es Salaam Seaport Container Throughput</th>
</tr>
</thead>
<tbody>
<tr>
<td>2010</td>
<td>409,517</td>
</tr>
<tr>
<td>2011</td>
<td>476,733</td>
</tr>
<tr>
<td>2012</td>
<td>547,364</td>
</tr>
<tr>
<td>2013</td>
<td>553,900</td>
</tr>
<tr>
<td>2014</td>
<td>612,600</td>
</tr>
<tr>
<td>2015</td>
<td>683,600</td>
</tr>
<tr>
<td>2016</td>
<td>648,100</td>
</tr>
<tr>
<td>2017</td>
<td>771,000</td>
</tr>
<tr>
<td>2018</td>
<td>903,000</td>
</tr>
<tr>
<td>2019</td>
<td>1,000,775</td>
</tr>
</tbody>
</table>

Table 1 and Figure 2 above both demonstrate the container throughput (TEU) of Dar es Salaam Seaport (2010-2019). It depicts that the annual average increase of by 11% from 409,517 TEUs in 2010 to 1,000,775 TEUs in 2019. This volume increase once remaining constant in the coming years, around 2022-2025 the seaport will grasp its existing volume of 1.2 million TEUs. It is important to bear in mind that, the most current data accessible in the World Bank database at the time of preparing this article was the container throughput statistics of the year 2019. The Dar es Salaam Seaport must be modernized in order to be well positioned to manage this significant growth in container throughput volume. For the aim of obtaining more precise data, numbers were used to represent years when using each of the following models, as shown in Table 2 below.

Table 2. Numbers adjusted to be applied for the exponential smoothing, grey, quadratic and combined forecasting models

<table>
<thead>
<tr>
<th>Numbers</th>
<th>Years</th>
<th>Numbers</th>
<th>Years</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2010</td>
<td>6</td>
<td>2015</td>
</tr>
<tr>
<td>2</td>
<td>2011</td>
<td>7</td>
<td>2016</td>
</tr>
<tr>
<td>3</td>
<td>2012</td>
<td>8</td>
<td>2017</td>
</tr>
<tr>
<td>4</td>
<td>2013</td>
<td>9</td>
<td>2018</td>
</tr>
<tr>
<td>5</td>
<td>2014</td>
<td>10</td>
<td>2019</td>
</tr>
</tbody>
</table>

A. The Exponential Smoothing Model

The Dar es Salaam Seaport Container Throughput original data, as shown in Table 1, illustrates that the trend is always changing. Hence, the Exponential Smoothing model is ideal for usage in these conditions. By applying figures from Table 1, the calculated smoothing coefficient figure was $\alpha = 0.9$. Also it discovered $x=781330$, $s=53989$ and $t = 10$ so $t-1 = 10-1$. 

[32]
B. The Grey Model

The eqn. (6) was used to compute grey model results using CHS based on the data of Dar es Salaam Seaport Container Throughput as follows; \( a = -0.01141, b = 694387.683 \) and \( p = 0.5 \). Therefore, the equation is:

\[
\hat{X}^{(1)}(k) = (X^{(0)(1)} - \frac{b}{a}) e^{-\frac{b}{a}(k-1)} + \frac{b}{a}, \quad k = 1, 2, \ldots n = 63989740.273e0.21418k—63576319.293
\]

C. The Quadratic Model

The Quadratic Model was analysed using Excel software, and the results showed there was correlation coefficient between the Container Throughput and Years.

D. The Combined Model

In defining the value of exponential smoothing model, grey model, quadratic model and a combined forecast model \( z_1, z_2, z_3 \) the eqn. (9) are calculated as: \( p_1 = 0.5 \); \( p_2 = 0.3 \); \( p_3 = 0.2 \). Hence a combined forecast model is expressed as:

\[
y = 0.5 p_1 + 0.3 p_2 + 0.2 p_3
\]

V. COMPARISON OF MODELS RESULTS

In Table 3 below, results of each forecasting model is well presented along with MAPE.

<table>
<thead>
<tr>
<th>Year</th>
<th>Dar es Salaam Seaport Container Throughput</th>
<th>Exponential Smoothing Model</th>
<th>Grey Model</th>
<th>Quadratic Model</th>
<th>Combined Forecast Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>2010</td>
<td>409,517</td>
<td>430,501</td>
<td>409,517</td>
<td>436,922</td>
<td>425,490</td>
</tr>
<tr>
<td>2011</td>
<td>476,733</td>
<td>469,111</td>
<td>509,011</td>
<td>470,119</td>
<td>468,946</td>
</tr>
<tr>
<td>2012</td>
<td>547,364</td>
<td>511,184</td>
<td>553,600</td>
<td>509,114</td>
<td>563,885</td>
</tr>
<tr>
<td>2013</td>
<td>553,900</td>
<td>557,030</td>
<td>582,164</td>
<td>537,806</td>
<td>512,164</td>
</tr>
<tr>
<td>2014</td>
<td>612,600</td>
<td>606,989</td>
<td>603,833</td>
<td>606,989</td>
<td>613,299</td>
</tr>
<tr>
<td>2015</td>
<td>683,600</td>
<td>661,427</td>
<td>659,863</td>
<td>674,748</td>
<td></td>
</tr>
<tr>
<td>2016</td>
<td>648,100</td>
<td>720,749</td>
<td>721,538</td>
<td>722,686</td>
<td></td>
</tr>
<tr>
<td>2017</td>
<td>771,000</td>
<td>785,390</td>
<td>788,909</td>
<td>803,178</td>
<td></td>
</tr>
<tr>
<td>2018</td>
<td>903,000</td>
<td>855,829</td>
<td>861,976</td>
<td>859,378</td>
<td></td>
</tr>
<tr>
<td>2019</td>
<td>1,000,775</td>
<td>1,047,947</td>
<td>1,041,800</td>
<td>1,044,398</td>
<td></td>
</tr>
<tr>
<td>MAPE</td>
<td>-</td>
<td>4.1%</td>
<td>4.3%</td>
<td>4.2%</td>
<td>3.8%</td>
</tr>
</tbody>
</table>

Forecasting Model with the most accurate results is Combined Forecasting Model. Indeed, it found to have the smallest mean absolute percentage error (MAPE) as it can be noticed in the Table 3 above. Therefore a Combined Forecasting Model is the best model for forecasting the Dar es Salaam Container Throughput in the next years. Forecasting container throughput using the correct model is significant for the strategic planning and development of country’s seaport and a country’s economy at large.

VI. CONCLUSION

The major goal of Tanzania is to become an emerging country by 2025. In order to attain this goal, the government recognizes the critical role of the Dar es Salaam Seaport in strengthening the country’s economy. Regarding the findings of this paper, the Tanzania Port Authority (TPA) can begin preparations to improve its seaport infrastructures in order to effectively serve the anticipated great demand. Dar es Salaam Seaport, which is presently the third most important seaport in Eastern Africa, has the potential to become the region’s leading seaport in the future. This amazing growth, however, will be dependent on the country’s political stability. A Combined Model was determined to be the most accurate of the forecasting models utilized in this study; however the findings of the Exponential Smoothing Model were very good. This is simply explained by the fact that Container Throughput of the Dar es Salaam Seaport is extremely volatile.

ACKNOWLEDGMENT

Authors give deeply thanks to the Chinese Government for providing a full scholarship to the first author under CSC to pursue a 4-years PhD program at Dalian Maritime University in China.
REFERENCES

[17] Ming-Hung Shu, Thanh-Lam Nguyen, Bi-Min Hsu, Chunwei Lu And Jui-Chan Huang, Forecasting Cargo Throughput with Modified Seasonal ARIMA Models, WSEAS TRANSACTIONS on MATHEMATICS, 2014. Volume 13, p.1710-181

AUTHORS

First Author – Msabaha Juma Mwendapole, PhD Candidate Logistics Engineering and Management, Dalian Maritime University and mwendapole.msabaha@gmail.com.

This publication is licensed under Creative Commons Attribution CC BY.

ISSN 2250-3153


42
Second Author – Jin Zhihong, Professor, Dalian Maritime University and jinzhihong@dlmu.edu.cn

Correspondence Author – Msabaha Juma Mwendapole, mwendapole.msabaha@gmail.com, mwendapole.msabaha@yahoo.com, +8615524811359.