Fuzzy Logic-Based Truck Demand Computational Model

Anugrah Mauludin and Ditdit Nugeraha Utama

Department of Industrial Engineering, Bina Graduate Program-Master of Industrial Engineering, Bina Nusantara University, Jakarta, Indonesia
Department of Computer Science, Bina Graduate Program-Master of Computer Science, Bina Nusantara University, Jakarta, Indonesia

Abstract: Trucking is one of the essential aspects linking suppliers and customers in many countries. Services at a trucking service provider will be based on a first-come-first-served basis. When a customer contacts a trucking service provider, there is a possibility that all the truck units owned are in fact being used to serve other customers causing potential delays in their delivery arrangement. For the trucking service provider, the truck demand prediction is not well observed resulting in a loss of business opportunities. The purpose of this research is to develop a truck demand prediction model using the Chen Fuzzy Time Series (FTS) method. The dataset used in this research was 11 years of monthly trucking demand information from a trucking service provider in Batam City, Indonesia. The proposed model achieved a Mean Absolute Percentage Error (MAPE) of 4.4%, this result indicates that the prediction was very accurate. After model deployment, expected can improve the readiness of trucking units.

Keywords: Trucking, Transportation, Forecast, Fuzzification, Time Series Data

Introduction

Nowadays, trucking is an essential link between the economies of many countries worldwide, bringing together suppliers and customers (Dorofeev et al., 2020). Regional effectiveness and economic development need dependable and vigorous road transport networks to accommodate the goods movement by truck (Madar-Vani et al., 2022). Trucks are still the primary mode of transportation of goods to transport containers (Adi et al., 2020). Trucking services at a distribution center are based on the principle of getting it quick and using time slots, but ad-hoc changes still require that sometimes orders are served in order of arrival (Carlan et al., 2022).

Based on information obtained from a trucking service provider in Batam City, Indonesia, when ad-hoc customers contact them by providing usage information such as date, route, type, and the number of trucks required, it is possible that all truck units are being used to serve other customers. If this happens, the contacting customer will be given information about when the next truck schedule will be available.

They can follow the new available schedule if there is no potential for late delivery. However, if there is a potential for late delivery, they will choose to cancel their orders and must contact other trucking service providers. Prediction of truck demand that is not well observed and the unavailability of truck units causing delays in the delivery process of goods for customers and the loss of business opportunities for trucking service providers. The trucking service provider so far predicted truck demand only based on intuition during any meeting and their prediction results cannot be scientifically justified. The purpose of this research is to develop a computational truck demand prediction model which provides more accountable prediction results to the business.

Many studies have discussed improving trucking management (Nadi et al., 2021) and developed a model able to predict any short-term fluctuation in trucking volume that happened in the primary container terminal at a seaport in Rotterdam. Research with the same theme was also carried out by Xie and Huynh (2010) at the port in Houston.

The developed model can predict trucking traffic generated at the port in one day. Hwang (2021) conducted a case study that uses data on the whole highway network in the United States and shipping information in 2007. The deployment of the proposed model makes it feasible to identify many issues, for example, improving the road infrastructure, reducing vehicle emissions, and reducing traffic congestion. These studies are relevant in helping to improve the efficiency and reliability of transportation.
networks performed by traffic management agencies, particularly around seaports. The research conducted by Chen et al. (2021) proposes an autonomous scheduling for trucks to resolve issues that happened when shipping containers between two ports in Singapore. Li et al. (2022) conducted research to balance inbound and outbound containers by transferring them from one truck to another in the logistics center to consolidate idle resources. The research conducted by Yang et al. (2021) focuses to enhance truck parking facilities, in collaboration with Washington's Department of Transportation.

From the above literature review on truck management, it can be concluded that, so far, no research has yet focused on truck demand prediction's research aims to build a truck demand prediction model. Time series is a group of data that has been arranged based on a time period or data organized in sequential arrangement referring to the days, weeks, months, and years (Afnisah and Marpaung, 2020). Analysis of data patterns is carried out on historical data to reveal trends that will occur in the future (Taufik et al., 2020). Time series analysis is essential in planning and decision-making (Telezhkin et al., 2021).

Fuzzy Time Series (FTS) is an artificial intelligence concept used to process available historical data to get a prediction (Andika and Sari, 2022). Prediction is needed to observe whenever an event can occur so that the right decision can be taken to deal with the event (Taufik et al., 2020). FTS was introduced by Song and Chissom (1993) and FTS methods that have developed in recent years include the Chen method, the Cheng method, the Markov Chain method, and other methods (Arnita et al., 2020). Compared to the Classical Time Series (CTS) method, FTS has the advantage that the predicted actual number values generated contain linguistic values so they can be easily understood (Solikhin et al., 2021). FTS avoids the basic assumptions of CTS (Panigrahi and Behera, 2020). FTS has more specific steps with each purpose that helps to calculate the accountable prediction, meanwhile, CTS only involves the identification of stationary time series, parameter estimation, and diagnostic check (Fattah et al., 2018). Also, FTS divides the interval into several ranges, this drives a more thorough prediction interval compared to CTS, which will provide more precise prediction results (Xie et al., 2021). In comparison with machine learning models, overall, the average performance of the machine learning models does not outperform FTS (Carbonneau et al., 2007). For long-term prediction, FTS found to be the most accurate model compared to the machine learning-based model, this is because FTS performs well with small numbers of data, and its fuzzy component is able to capture the uncertainty of the data (Lee et al., 2022).

Prediction using various FTS methods has been implemented widely over the years. In 2018, (Aziz et al., 2018) predicted sugar prices in Indonesia using Song and Markov Chain methods, and (Ningrum et al., 2018) predicted Jakarta Composite Stock Price Index (JCI) using the Chen method. Kafi et al. (2019) predicted the closing price of Company X shares using the Markov Chain method and (Novitasari et al., 2019) predicted the number of sunspots to identify flares using the Markov Chain method. In 2020, (Alyousifi et al., 2020) predicted the daily basis air pollution index in Malaysia using the Markov Chain method, (Arnita et al., 2020) predicted rainfall in Medan using Chen, Cheng, and Markov Chain method, (Bettiza, 2020) predicted Indonesia’s total population using Chen, Cheng and Markov Chain method, (Ramadani and Devianto, 2020) predicted bitcoin prices using Chen and Markov Chain methods and (Taufik et al., 2020), predicted gold prices in Indonesia using Markov Chain method. Ariyanto et al. (2021) predicted the sales at PT Y as an improving method to get the cheapest storage cost possible using the Chen method, (Boechel et al., 2021), predicted the apple tree phenological stages using the Chen method, (Rizqon and Jati, 2021), predicted train passengers’ numbers in Indonesia using Chen method and (Wirawan et al., 2021), predicted the chili prices in Indonesia using Chen and Cheng methods.

From the above literature review on FTS, it can be concluded that, so far, no FTS prediction has been implemented on truck demand prediction. The above literature review as well concluded that the Chen method has been the method being used more in the last two years; therefore, the Chen method is selected to build the truck demand prediction model in this research. Additionally, this research assesses the performance of the generated predictions using Mean Absolute Percentage Error (MAPE), a popular metric for assessing prediction outcomes. Based on literature reviews, MAPE has been used in various sectors namely finance, economics, and environments. For instance, in finance, MAPE can be used to evaluate the accuracy of stock price predictions. In economics, MAPE can be utilized to assess the effectiveness of sugar price predictions. In environments, MAPE can be employed to measure the accuracy of air pollution index predictions.

Materials

The training data for this research were obtained from a trucking service provider in Batam City, Indonesia, and are displayed in Table 1. Contain 11 years of monthly truck demand, ranging from January 2012 to December 2022. There are 132 lines of total monthly data. Table 1 summarizes and displays the data by 3 first months and 3 last months from the entire duration. Full dataset can be accessed at the link provided at the data availability statement section. No further data pre-processing needs to be done since the data has been in good shape with no abnormalities found.
method, the Markov Chain method and other methods (Arnita et al., 2020). The FTS method used the most in studies in the last two years is the Chen method. Therefore, the Chen method will be used to build a trucking demand prediction model in this research.

The prediction results’ performance measurement was carried out using MAPE. MAPE aims to measure the percentage error of the resulting predictions. Thus, the MAPE measurement can serve as a valuable tool for evaluating the proposed model’s effectiveness. It is expected that the prediction model built has a relatively good level of accuracy. The result of the MAPE measurement is in the form of a percentage (%). The lower the percentage, the more accurate the prediction model that has been built.

**Fuzzy Time Series Chen Method**

Rizqon and Jati (2021) described the FTS Chen method in five main steps, as shown in Fig. 1. The very first step is establishing the universe of discourse (U) using Eq. 1, where $D_1$ and $D_2$ are constants determined during the research. $D_{min}$ represents the smallest number available in the data series and $D_{max}$ represents the largest number available in the data series. The second step is calculating the number of intervals (i) by dividing the (U) into separate intervals with equal length using Eq. 2, where $n$ is the total row of the available data, and then calculating each of the interval’s length (l) using Eq. 3:

$$U = \left[D_{min} \cdot D_1; D_{max} + D_2\right]$$  \hspace{1cm} (1)

$$i = 1 + 3.3\log(n)$$  \hspace{1cm} (2)

$$l = \left(D_{max} - D_{min}\right)/i$$  \hspace{1cm} (3)

The third step is creating a Fuzzy Logic Relationship (FLR) through fuzzification. The fuzzification process itself refers to categorizing every data to fall under which interval according to the interval data that has been formed earlier. The FLR indicates the relationship between the fuzzifications. The fourth step is creating a Fuzzy Logic Relations Group (FLRG). Establishing FLRG is performed by combining every result in FLRs into various grouping depending on their relation’s initial state. The purpose of forming FLRGS is to identify patterns and trends in the data that can help forecast the number of decision-making processes. The fifth step is defuzzification and forecasting. Defuzzification and forecasting are important steps in fuzzy logic-based systems, as they provide a way to convert the output into a numerical value that can be used for decision-making and further analysis. The defuzzification process determines the predictive value and refers to the FLRG results based on three rules.
Table 2: MAPE results interpretation

<table>
<thead>
<tr>
<th>MAPE results (%)</th>
<th>Predicting accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Below 10</td>
<td>Very accurate</td>
</tr>
<tr>
<td>Between 10-20</td>
<td>Good</td>
</tr>
<tr>
<td>Between 20-50</td>
<td>Moderate</td>
</tr>
<tr>
<td>Above 51</td>
<td>Not accurate</td>
</tr>
</tbody>
</table>

The first rule is where there is a single fuzzy logical relationship within a particular series of fuzzy logical relationships, for instance, \( Ai \rightarrow Aj \), the prediction value will be defined as \( F(t) = Aj \). The second rule is where there is more than a single relationship available within a particular series of fuzzy logic relationships, for instance, \( Ai \rightarrow Aj, Aj2, \ldots Ajn \), the prediction value will be the average value of the median of \( Aj1, Aj2, \ldots Ajn \), or defined as \( F(t) = \text{Average from Median of } Aj1, Aj2, \ldots Ajn \). The third rule is where there is no relationship at all on \( j \), there is none like \( (Aj \rightarrow #) \), the prediction value will be defined as \( F(t) = Aj \).

Performance Measurement

The method used to measure the prediction result in this research is MAPE. (Tricahya and Rustam, 2019) Using Eq. 4, where \( X_i \) is the actual data value, the predicted data is \( \hat{X}_i \) and \( N \) is the amount of data. MAPE works by measuring the average error percentage from all the predictions generated. The result of the MAPE measurement is in the form of a percentage (%). The interpretation of the MAPE value on the prediction performance is described in Table 2:

\[
MPAE = \frac{100\%}{N} \sum_{i=1}^{N} \left( \frac{x_i - \hat{x}_i}{x_i} \right)
\]  

Fig. 1: Proposed prediction method flow chart

Results

Calculating the Universe of Discourse (U)

U needs to be determined to create a limit of a set of data that will be involved in the prediction generation. From the monthly truck demand data, the minimum number is 430 and the maximum number is 997. Both data were used in Eq. (1) to determine \( U \). The values of \( D_1 = 0 \) and \( D_2 = 0 \), were arbitrarily chosen to obtain the appropriate values determination of the \( D_{\text{min}} \) and \( D_{\text{max}} \). As a result, [430; 997] was calculated as the \( U \) in this research.

Determining the Interval

The interval needs to be determined to create a data segmentation which will be used later in the fuzzification process. First, the number of intervals was determined by applying Eq. 2, and next, the length of each interval was determined by applying Eq. 3. Calculated, the quantity of intervals is 8 and the length of each of the intervals is 81. Each interval has its minimum and maximum data, which will then be used to determine the median data. The median data of each interval can be determined by taking the average of its minimum and maximum data. Table 3 presents the complete outcome from the interval determination stage.

Fuzzification and Fuzzy Logic Relationship (FLR)

The fuzzification stage was performed to refer to the interval information obtained earlier from the previous step. The fuzzification step will convert precise language values into fuzzy linguistic values. In this step, each
month’s demand quantity was nominated to fall under which interval refer to Table 3. For example, the first-month demand quantity is 560, which according to the formed intervals in Table 3, is fuzzified as A2. The second-month demand quantity is 596, which is fuzzified as A2 as well. Repeat the step until the end of the data sets to have all months fuzzified.

The next step here is to determine the relationship between each fuzzification. The FLR was created by using the fuzzification nomination for each month that had been done previously. For instance, the first and second months were fuzzified as A2. This resulted in the first created FLR as $A_2 \rightarrow A_2$ and repeating that step to the next month until the end of the datasets. This process of creating the FLR allowed for a clear representation of how each month’s fuzzification relates to the others in the dataset. This approach facilitated the analysis and comparison of the data for identifying any patterns or trends available within the system being modeled. Table 4 shows us the formed fuzzification, as well as the FLRs results, summarizes and displays the data by 3 first months and 3 last months from the entire duration.

Fuzzification and Fuzzy Logic Relationship (FLR)

The fuzzification stage was performed to refer to the interval information obtained earlier from the previous step. The fuzzification step will convert precise language values into fuzzy linguistic values. In this step, each month’s demand quantity was nominated to fall under which interval refer to Table 3. For example, the first-month demand quantity is 560, which according to the formed intervals in Table 3, is fuzzified as A2. The second-month demand quantity is 596, which is fuzzified as A2 as well. Repeat the step until the end of the data sets to have all months fuzzified.

Fuzzy Logic Relationship Group (FLRG)

The further step is forming the FLRG where this step is to have all established FLR relations from Table 4 grouped based on the relations starting fuzzification point. From the FLR, take the relations starting fuzzification as the baseline and then followed by to which fuzzification it has relation. For example, fuzzification A2 has 3 FLRs; $A_2 \rightarrow A_1, A_2 \rightarrow A_2$ and $A_2 \rightarrow A_3$. Then A2 group will be A1, A2 and A3. The complete information on the formed FLRG is shown in Table 5.

Defuzzification Process and Prediction

This process of calculating prediction values from FLRG involves taking the average median values from intervals and using them to predict future outcomes. The prediction values of each fuzzification were obtained by calculating the average values of the medians available in the FLRG. For example, the $A_2 \rightarrow A_1, A_2, A_3$, where the median of A1 is 511, the median of A2 is 592 and the median of A3 is 673. The average of 511, 592, and 673 is 592. Therefore, the prediction value obtained from the group of A2 is 592.

The complete defuzzification and prediction values are shown in Table 6. Now these prediction values can be used to make decisions or predictions about trucking future demand. The prediction results of the data train are obtained and refer to the prediction value for each fuzzification, as shown in Table 6. For example, January 2012 fuzzification is A2; therefore, the prediction value for that month is 592.
Table 7: Prediction results

<table>
<thead>
<tr>
<th>Month</th>
<th>Quantity</th>
<th>Fuzzification</th>
<th>Prediction</th>
</tr>
</thead>
<tbody>
<tr>
<td>January 2012</td>
<td>560</td>
<td>A2</td>
<td>592</td>
</tr>
<tr>
<td>February 2012</td>
<td>596</td>
<td>A2</td>
<td>592</td>
</tr>
<tr>
<td>March 2012</td>
<td>664</td>
<td>A3</td>
<td>673</td>
</tr>
<tr>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
</tr>
<tr>
<td>October 2022</td>
<td>548</td>
<td>A1</td>
<td>552</td>
</tr>
<tr>
<td>November 2022</td>
<td>521</td>
<td>A1</td>
<td>552</td>
</tr>
<tr>
<td>December 2022</td>
<td>544</td>
<td>A1</td>
<td>552</td>
</tr>
<tr>
<td>January 2023</td>
<td>-</td>
<td>-</td>
<td>511</td>
</tr>
<tr>
<td>February 2023</td>
<td>-</td>
<td>-</td>
<td>511</td>
</tr>
<tr>
<td>March 2023</td>
<td>-</td>
<td>-</td>
<td>511</td>
</tr>
</tbody>
</table>

Table 8: MAPE measurement

<table>
<thead>
<tr>
<th>Month</th>
<th>Actual quantity</th>
<th>Predicted</th>
<th>Absolute percentage error %</th>
</tr>
</thead>
<tbody>
<tr>
<td>January 2012</td>
<td>560</td>
<td>592</td>
<td>5.7</td>
</tr>
<tr>
<td>February 2012</td>
<td>596</td>
<td>592</td>
<td>0.7</td>
</tr>
<tr>
<td>March 2012</td>
<td>664</td>
<td>673</td>
<td>1.4</td>
</tr>
<tr>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
</tr>
<tr>
<td>October 2022</td>
<td>548</td>
<td>552</td>
<td>0.6</td>
</tr>
<tr>
<td>November 2022</td>
<td>521</td>
<td>552</td>
<td>5.9</td>
</tr>
<tr>
<td>December 2022</td>
<td>544</td>
<td>552</td>
<td>1.4</td>
</tr>
<tr>
<td>MAPE</td>
<td></td>
<td></td>
<td>4.4</td>
</tr>
</tbody>
</table>

Fig. 2: Time series graph of the training data (2012-2022) and the testing data (2023)

The prediction value for the data test January 2023-March 2023 obtained by using constructed model is 511. The prediction results are shown in Table 7, which summarizes and displays the data by 3 first months and 3 last months from the entire duration. The time series graph of the training data (2012-2022) and the testing data (2023) are visualized in Fig. 2.

Performance Measurement

The prediction results obtained from the data train are measured by using MAPE. Table 8 shows the individual absolute percentage error of each month's trucking order quantity as well as the overall MAPE for the whole available months, summarizes and displays the data by 3 first months and 3 last months from the entire duration. The constructed model has a MAPE value of 4.4%, which is considered a very accurate prediction.

Discussion

Coming together with the solution this model offers, there are challenges that seem to be visible during implementation as well. First are potential inaccuracies in data inputs, it is good for the trucking service provider to have validation in place prior to the data input. Second is the Covid-19 event that happened, or any other disruptive changes in the global marketplace. A proper analysis on a regular basis held by the trucking service provider will be a good step to understanding the situation and taking further action related to it. Changing the intuition-based mindset referring to excel data when deciding the future demand is another challenge faced by the stakeholders in the trucking service provider, it is better to establish a deployment team, which is responsible to monitor and report the accuracy of the prediction as well. This deployment team can also lead the extent of this model to account for complex transportation networks in the trucking industry.

Conclusion

This study proposed monthly trucking demand prediction using the FTS Chen method. Furthermore, the study found out that the FTS Chen method which is based on FTS is effective to predict the monthly trucking demand. The proposed model was trained using monthly trucking demand historical data ranging from January 1\textsuperscript{st}, 2012, to December 31\textsuperscript{st}, 2022. Then, the model was evaluated using MAPE which measures the accuracy of the predicted values, with a 4.4% result. The prediction value earned for January 2023-March 2023 is around 511 trips per month which can be used by trucking companies to plan their operations and resources. In conclusion, with a MAPE value of 4.4%, the proposed model is considered a very accurate prediction model to predict the monthly trucking demand, which can provide valuable insights for trucking companies to make decisions.

Acknowledgment

In this opportunity, we would really love to express our gratitude to Bina Nusantara University for providing us
with the chance to conduct the research, particularly the Binus graduate program, master of industrial engineering.

**Funding Information**

The authors should acknowledge the funders of this manuscript and provide all necessary funding information.

**Author’s Contributions**

**Anugrah Mauludin**: Performing all the research and analysis and written the paper.

**Ditdit Nugeraha Utama**: Supervising the research and reviewed the paper.

**Ethics**

This study is purely developed from the research that the authors went through. The contents are unpublished.

**Data Availability Statement**

The data that support the findings of this study are openly available in Kaggle at https://www.kaggle.com/datasets/anukonly/dataset-20t, reference number (Mauludin, 2023).

**References**


