FORECASTING THE DEMAND FOR TRANSPORT SERVICES
ON THE EXAMPLE OF A SELECTED LOGISTIC OPERATOR

Małgorzata GRZELAK¹, Anna BORUCKA², Zbigniew BUCZYŃSKI³
¹, ², ³ Military University of Technology, Faculty of Logistics, Warsaw Poland

Abstract:
The number of shipments is growing every year, and as a result, new transport companies arise. The increase in competition requires from entrepreneurs to apply solutions increasing the level of services provided in order to best satisfy the needs of the customers. In this aspect, minimizing the time of deliveries is extremely important, and it can be achieved, for example, by implementing the cross-docking method. It consists in consolidation of cargo from different shipment locations that is delivered in the same direction. The main feature of the above method is to keep the labor intensity of operations and the interference in the cargo to the minimum.

The purpose of this article is to present a research on a logistic operator working based on a cross-docking warehouse with a capacity significantly lower than the average daily quantity of shipments handled. This requires both effective management of the available space and minimizing the time spent on manipulation activities. Therefore, it is important to know the expected number of parcels that are planned to be received and shipped on a given day in order to coordinate the work in the warehouse. It is possible to estimate it by using mathematical methods of forecasting. One of them - the multiple regression - is presented in this article. The calculations were made on the basis of collected empirical observations concerning orders for pallet spaces placed by customers. Such a forecast allows for improvement of the processes of planning and management of the possessed resources. It allows to adjust the number of warehouse workers or vehicles necessary for internal transport to the expected needs. Ultimately, it may translate into more efficient functioning not only of the surveyed branch, but also of the whole network.

Keywords: multiple regression, forecasting, cross-docking

To cite this article:
Grzelak, M., Borucka, A., Buczyński, Z., 2019. Forecasting the demand for transport services on the example of a selected logistic operator. Archives of Transport, 52(4), 81-93. DOI: https://doi.org/10.5604/01.3001.0014.0210

Contact:
1) malgorzata.grzelak@wat.edu.pl [https://orcid.org/0000-0001-6296-7098], 2) anna.borucka@wat.edu.pl [https://orcid.org/0000-0002-7892-9640]
1. Introduction

In recent years, the market of transport services has been growing steadily and the so-called "one delivery day" is already a standard. Reports and calculations prepared by experts indicate that 395 million shipments could have been delivered in Poland in 2017 and that the number of parcels handled will increase every year by approx. 11% (Kawa, 2017). Dynamics and requirements of the market force suppliers to introduce better and better logistic solutions, mainly those ensuring the delivery of ordered goods within no more than 24 hours.

In times of strong competition on the TSL market, logistic operators have to meet the requirements of customers by delivering diversified and fragmented shipments. The supply chain should guarantee short delivery time while maintaining low prices of services provided. The support may be provided by the implementation of a system based on the cross-docking method, which allows to increase the effectiveness and efficiency of the processes (Zdunek, 2017). This technique permits to reduce the cost of supplies that are stored in a warehouse in less than 24 hours. In addition, it leads to increase the flow of goods and to reduction of delivery time. (Ladier, Alpan, 2016).

Cross-docking requires precise synchronization of the processes of goods receipt and release. It can be implemented on three levels (Nikolopoulou et al., 2017), (Michałowska et al. 2010):

− pallet cross-docking understood as the transshipment of whole pallets, homogeneous in terms of pallet content, from one means of transport to another,

− casket cross-docking, where all homogeneous pallets are delivered directly to dispatch warehouses, where their contents are divided into components forming "caskets" containing various goods according to the received orders and then delivered directly to recipients,

− cross-docking of orders completed by suppliers, used least frequently, where the manufacturer is responsible for preparing and completing a pallet which is then transported directly to the recipient.

The specificity of transport processes carried out with the use of the presented method and the logistic challenges faced by the companies that use it have become the origin of the article. Its aim is to indicate the possibility of developing a model of forecasting the demand for pallet space on the basis of historical data on customer orders and to present a method of multiple regression as a tool to predict the demand for pallet space, supporting decision-making in this area.

2. The state of the problem - literature analysis

Literature analysis indicates that there is no clear definition of cross-docking. It can be understood as a process of consolidation of cargo originating from different shipment locations and intended to be delivered to the same destination with minimum labor intensity of the activities performed and minimum interference in the cargo, and without storage of materials between the loading and unloading phase or with storage in a warehouse for a very short period of time (Lewczuk, 2013). According to Bozarth and Handfield (Bozarth, Handfield, 2007), cross-docking is a kind of warehouse management in which warehouse, accepting large incoming transports, divides them into smaller ones and then ships them to local customers. According to the dictionary of transport and logistics (Kozarkiewicz, 2009), cross-docking consists of activities related to unloading, loading, segregation and integration of cargoes, which are performed in a warehouse, without the storage phase, directly between means of external transport waiting at receipt and release docks.

Available in the literature cross-docking models allow to support decision making in the enterprise management process. Conducted researches can be classified based on its decision level and horizon of planning. There are available solution on operational, tactical and strategic level. For the needs of ongoing business operations, authors usually prepare picking schedules. These include issues such as determination of truck scheduling (Dwi, Lee, Rajesh, 2014; Mohammad, Mohtsen, 2016; Keshtzari et al., 2016), dock door assignment (Nassief et al., 2016; Enderer et al., 2017; Nassief, Contareras, Juamard, 2018), vehicle routing (Dwi, Lee, Rajesh, 2016), transshipment problem (Ladier, Alpan, 2016) and product allocation (Soleimaninadegeny et al., 2017). Various models were developed for two variants: with a separate temporary storage area and in the absence thereof (Wooyeon, Egbelu, 2008).

At the tactical level, the available solutions present the problems of designing and organizing cross-docking terminals (Lewczuk, 2013). Some of the researches cover the dock door assignment problem, because this issue affects the flow of materials inside
At strategic level, researches were conducted to support long term decision making. They focus mainly on the problems of determination of the number and location of cross-docking terminal and the number of vehicles in the network (Dwi, Lee, Rajesh, 2014). They enable the management of distribution networks, leading to minimization of delivery costs or minimization of total of delays (both in the warehousing and transport phases). The research was conducted on the basis of linear programming methods (including mixed integer programming model), as well as using heuristic methods based on genetic algorithms, among others simulated annealing, tabu and their combination (Kucukoglu Ozturk, 2014; Mousavi et al., 2014; Hosseini et al., 2014).

3. Research subject and company characteristics

This article analyses the transport process realized in one of the branches of a large logistics operator dealing with the handling of forwarding and transport orders within the supply chain, providing storage services, intermodal, road, rail, air and sea transport services together with customs service. The most frequent form of transport is domestic transport of general cargo shipments with the use of cross-docking system (Ladier, Alpan, 2016). The company has 19 branches and warehouse facilities adapted for cross-docking, which ensures flexibility and comprehensive customer service (Fig. 1).

The management of the assortment in the surveyed company consists in unloading the cargo in a special zone and preparing it for further dispatch, without storage (Dydkowski, 2018). This allows handling large volumes, standardized in terms of cargo unit (in the case in question - a EURO pallet and its multiplicity). Limiting the storage process to the necessary minimum allows the company's customers to reduce the costs of storage and realize their production in the Just In Time system (Jacyna-Golda et al., 2018). In addition, it reduces storage space, which, however, entails the need for precise synchronisation of activities and use of modern solutions in the field of internal transport. A simplified way of functioning of the system is presented in figure 2.

The transport process of a shipment begins with its assignment to appropriate driver and truck, who will pick it up from the customer and deliver it to the company's branch. Then, on the same day, it is planned to be loaded on a line route which will take it to the final transshipment warehouse from where it will be delivered to the customer on the following day (Comi et al., 2018). It is assumed that the total time does not exceed 24 hours.

The term "shipment" shall be understood as goods placed in packages without carriers or in packages on carriers of the following sizes:

- parcels - from 0.01 to 0.4 of pallet space,
- cargo:
  - on a EUR pallet - 1 pallet space,
  - on quarter pallets - 0.25 pallet space,
  - on half pallets - 0.5 pallet space,
  - on a disposable pallet - 1 pallet space,
  - on a max pallet - 1.5 pallet space,
  - on other carriers - double pallets, pallets of non-standard dimensions,
- shipments which cannot be palletised because of their intended use or outer dimensions.
For the purpose of further calculations, it was assumed that one shipment occupies on average 0.8 - 0.9 pallet space.

The branch under analysis serves mainly the Mazowieckie Province, reaching north to Przasnysz and Ostrołęka, east to Sokółow Podlaski and Siedlce, south to Garwolin, Kozienice, Mogielnica and Rawa Mazowiecka, and west to Łowicz, Płock and Sierpc. It has the largest transshipment warehouse in the company with a capacity of 2600 pallet spaces. The average daily turnover of shipments is 5000 pcs, which translates into about 4000 pallet spaces. The ratio of the quantity of cargo passing through the warehouse to its capacity requires effective space management and short handling time in order to minimize the possibility of warehouse blockage. Therefore, it is important to develop an appropriate model for forecasting the number of pallet spaces occupied by shipments ordered by customers to support the management of their turnover in the warehouse.

4. Test method - multiple regression model

The research covered the demand for pallet space in the warehouse in the years 2015-2018. In the first step, a database was prepared by completing the missing information in the time series resulting from the occurrence of public holidays, replacing them with the average of observations for the day of the week in a given month (Sokołowski, 2016). The development of a model for forecasting daily demand for pallet spaces with the use of multiple regression started with visual inspection of the graph (Fig. 3) (Mattias, Tamas, Csaba, 2017).

The analysis of the time series indicates the existence of a trend. The confirmation is the calculated correlation coefficient, which is 0.709 and is statistically significant. In addition, cyclical and seasonal fluctuations are visible. They are evidenced by the basic statistics contained in Table 1 (mean, median and coefficient of variation) and the box plot of observation variability presented for individual days of the week and months (Fig. 4 and 5).

The average demand for pallet spaces in the whole research period was 2213 pcs. and is close to the median value. As far as days of the week are concerned, the biggest needs are on Thursdays (2352 pcs.), while the lowest turnover is recorded on Mondays (2037 pcs.). On a monthly basis, the highest average indications are in November (2423 pcs.) and the lowest in December (1934 pcs.).

![Fig. 3. Pallet space demand during the research period](image-url)
The article proposes a model of forecasting the demand for pallet spaces using the multiple regression method based on explanatory variables resulting from the calendar date. Since these are quality variables, it was necessary to re-code them into binary variables. Model estimation is only possible if one of them is eliminated for each identified category and attribute. The model thus obtained consists of an absolute term, a linear trend, a sum of products of variables representing individual months and days of the week, variables related to a sudden increase or decrease in demand and a random component. The estimated parameters are presented in table 2. The variable for Monday is omitted for the days of the week, and the variable for January is omitted for the months.

![Box plot showing the weekly variability](image)
The regression equation takes the form (1):

\[ y_i = 1694.64 + 445.33 \cdot x_{\text{growth\ in}} - 547.16 \cdot x_{\text{decline\ in}} + 0.46 \cdot x_{\text{linear\ trend}} + 110.61 \cdot x_{\text{Tuesday}} + 108.01 \cdot x_{\text{Wednesday}} + 152.76 \cdot x_{\text{Thursday}} + 64.53 \cdot x_{\text{February}} + 114.74 \cdot x_{\text{March}} + 130.12 \cdot x_{\text{April}} + 178.84 \cdot x_{\text{May}} + 126.91 \cdot x_{\text{June}} + 114.33 \cdot x_{\text{July}} + 86.38 \cdot x_{\text{August}} + 178.55 \cdot x_{\text{September}} + 106.81 \cdot x_{\text{October}} + 167.95 \cdot x_{\text{November}} \]

The last stage is diagnostics, performed on the basis of the analysis of the distribution of model residuals (Bielińska, 2007). They are characterized by a normal distribution, which is confirmed by the chi-square test and the calculated p-value equal to 0.35, which proves that there are no grounds for rejecting the H0 which speaks about the normality of distribution (Fig. 5).

However, the analysis of the graphs of autocorrelation and partial autocorrelation functions showed the existence of significant values of these functions, which means that some dependencies remained in the model that were not explained by it (Fig. 6 and 7) (Sokołowski, 2016).
Chi-squared test = 9.89, df = 9, p = 0.35

Fig. 5. Residuals distribution histogram

Fig. 6. Graph of the autocorrelation of the residuals of the model
Therefore, an attempt was made to improve the model. Since on ACF and PACF diagrams a significant value of these functions is clearly visible for a delay equal to one, it was proposed to add a variable, which is the value of empirical observation delayed by one. The estimated parameters of the new model are shown in table no. 3. The variable for Monday is omitted for the days of the week, and the variable for December is omitted for the months. The model consists of an absolute term, a linear trend, a sum of products of variables representing the monthly and weekly variability, variables related to sudden increase and decrease in demand, parameters of a delayed variable and a random component.

The regression equation takes the form (2):

\[
y_i = 1321.89 + 404.03 \cdot x_{\text{growth in}} - 498.96 \cdot x_{\text{decline in}} + 0.32 \cdot x_{\text{linear trend}} + 0.21 \cdot x_{\text{delayed}} + 139.55 \cdot x_{\text{Tuesday}} + 89.11 \cdot x_{\text{Wednesday}} + 128.09 \cdot x_{\text{Thursday}} - 41.35 \cdot x_{\text{Friday}} + 75.49 \cdot x_{\text{February}} + 99.29 \cdot x_{\text{March}} + 113.57 \cdot x_{\text{April}} + 155.72 \cdot x_{\text{May}} + 111.31 \cdot x_{\text{June}} + 105.74 \cdot x_{\text{July}} + 84.57 \cdot x_{\text{August}} + 146.55 \cdot x_{\text{September}} + 95.28 \cdot x_{\text{October}} + 138.79 \cdot x_{\text{November}}
\]

![Fig. 7. Chart of partial autocorrelation of the model residuals](image)

Table 3. Estimated parameters of the improved multiple regression model

<table>
<thead>
<tr>
<th></th>
<th>R2= 0.79</th>
<th>Adjust. R2= 0.79</th>
<th>F(19,11)=214.68 p&lt;0.00</th>
</tr>
</thead>
<tbody>
<tr>
<td>N=1089</td>
<td>b</td>
<td>Std. error</td>
<td>t(347)</td>
</tr>
<tr>
<td>Absolute term</td>
<td>1321.89</td>
<td>43.74</td>
<td>30.21</td>
</tr>
<tr>
<td>growth in demand</td>
<td>404.03</td>
<td>16.61</td>
<td>24.32</td>
</tr>
<tr>
<td>decline in demand</td>
<td>-498.96</td>
<td>31.01</td>
<td>-16.08</td>
</tr>
<tr>
<td>linear trend</td>
<td>0.32</td>
<td>0.02</td>
<td>12.75</td>
</tr>
<tr>
<td>delayed variable</td>
<td>0.21</td>
<td>0.02</td>
<td>9.64</td>
</tr>
<tr>
<td>Tuesday</td>
<td>139.55</td>
<td>18.51</td>
<td>7.53</td>
</tr>
<tr>
<td>Wednesday</td>
<td>89.11</td>
<td>18.74</td>
<td>4.75</td>
</tr>
<tr>
<td>Thursday</td>
<td>128.09</td>
<td>18.85</td>
<td>6.79</td>
</tr>
<tr>
<td>Friday</td>
<td>-41.35</td>
<td>18.76</td>
<td>-2.20</td>
</tr>
<tr>
<td>January</td>
<td>30.74</td>
<td>27.18</td>
<td>1.13</td>
</tr>
<tr>
<td>February</td>
<td>75.49</td>
<td>28.00</td>
<td>2.69</td>
</tr>
<tr>
<td>March</td>
<td>99.29</td>
<td>29.73</td>
<td>3.33</td>
</tr>
<tr>
<td>April</td>
<td>113.57</td>
<td>30.38</td>
<td>3.73</td>
</tr>
<tr>
<td>May</td>
<td>155.72</td>
<td>30.20</td>
<td>5.15</td>
</tr>
<tr>
<td>June</td>
<td>111.31</td>
<td>30.22</td>
<td>3.68</td>
</tr>
<tr>
<td>July</td>
<td>105.74</td>
<td>29.76</td>
<td>3.55</td>
</tr>
<tr>
<td>August</td>
<td>84.57</td>
<td>29.38</td>
<td>2.87</td>
</tr>
<tr>
<td>September</td>
<td>146.55</td>
<td>31.03</td>
<td>4.72</td>
</tr>
<tr>
<td>October</td>
<td>95.28</td>
<td>29.83</td>
<td>3.19</td>
</tr>
<tr>
<td>November</td>
<td>138.79</td>
<td>30.66</td>
<td>4.52</td>
</tr>
</tbody>
</table>

(2)
Again, based on the chi-square test, the distribution of the model residuals was examined, which confirmed its compliance with the normal distribution (p-value greater than the confidence level \( \alpha=0.05 \) - Fig. 8) and thus the correctness of the model built.
The above conclusions are also confirmed by the charts of autocorrelation and partial autocorrelation of the model residuals, which indicate the lack of significant dependencies between the individual values, allowing to consider their distribution as a white noise process.

The chart of empirical and forecast values shown in Figure 11 proves that the model effectively follows the time series, well reflecting the seasonality of the process and taking into account the values of sudden growth and decline in demand. Moreover, the coefficient of determination of the estimated model is 78% and is at a satisfactory level.

The quality of the proposed model was checked using test observations (from March 2019), which were not used to build the model. The comparison of empirical observations and forecast values together with the relative error of the forecast is presented in Table 4.

The biggest forecasting error is 13.77% and the smallest is 0.59%. The average relative forecast error for the test period is 6%, which should be considered a satisfactory result.

### Table 4. Comparison of forecast values and test observations

<table>
<thead>
<tr>
<th>date</th>
<th>Forecast - improved regression model</th>
<th>Empirical data</th>
<th>Relative error of the forecast Ψ [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>2019-03-11</td>
<td>2703</td>
<td>2888</td>
<td>-6.86</td>
</tr>
<tr>
<td>2019-03-12</td>
<td>2288</td>
<td>2297</td>
<td>-0.41</td>
</tr>
<tr>
<td>2019-03-13</td>
<td>2304</td>
<td>2290</td>
<td>0.59</td>
</tr>
<tr>
<td>2019-03-14</td>
<td>2733</td>
<td>2847</td>
<td>-4.18</td>
</tr>
<tr>
<td>2019-03-15</td>
<td>2669</td>
<td>2847</td>
<td>-6.70</td>
</tr>
<tr>
<td>2019-03-18</td>
<td>2617</td>
<td>2890</td>
<td>-10.45</td>
</tr>
<tr>
<td>2019-03-19</td>
<td>2642</td>
<td>2742</td>
<td>-3.81</td>
</tr>
<tr>
<td>2019-03-20</td>
<td>2441</td>
<td>2293</td>
<td>6.07</td>
</tr>
<tr>
<td>2019-03-21</td>
<td>2505</td>
<td>2849</td>
<td>-13.77</td>
</tr>
<tr>
<td>2019-03-22</td>
<td>2754</td>
<td>2850</td>
<td>-3.45</td>
</tr>
<tr>
<td>2019-03-25</td>
<td>2843</td>
<td>2893</td>
<td>-1.73</td>
</tr>
<tr>
<td>2019-03-26</td>
<td>2568</td>
<td>2745</td>
<td>-6.88</td>
</tr>
<tr>
<td>2019-03-27</td>
<td>2683</td>
<td>2738</td>
<td>-2.04</td>
</tr>
<tr>
<td>2019-03-28</td>
<td>2751</td>
<td>2852</td>
<td>-3.67</td>
</tr>
<tr>
<td>2019-03-29</td>
<td>3038</td>
<td>2852</td>
<td>6.11</td>
</tr>
</tbody>
</table>
5. Summary
The aim of the article was to present a method that could be used in forecasting the demand for pallet spaces. The research was conducted on the example of historical data on customer orders acquired from a company that bases its warehouse and transport processes on the cross-docking system, i.e. on picking transshipment, and the time of realization of the above tasks for individual parcels does not exceed 24 hours.

The article uses multiple regression method, with explanatory variables resulting from days of the week, months and periods of rapid growth and decline in demand. It also takes into account the fact that development trends occur in a series. Two models were proposed. In the first one the coefficient of determination was 77%, but the diagnostics of the distribution of residuals showed that although it is consistent with the normal distribution, there are significant dependencies in it that were not explained by the model, which was proven by the autocorrelation and partial autocorrelation charts. Therefore, a second model was developed, taking into account the delayed value of the analysed variable. The calculated coefficient of determination did not improve, but the autocorrelation and partial autocorrelation charts indicated the lack of significant dependencies between the individual model residuals, which indicates their randomness and allows to consider the developed model as the better one.

The analysis of the literature indicated that there are few studies aimed at forecasting the demand for pallet spaces in cross-docking terminals. Most of the research (at the operational level) focuses on the issues of developing transhipment work schedules to optimize the efficiency of ongoing processes based on current data. The presented tool may complement the available models, providing information on how and at what level the studied phenomenon will shape in the future. It can be a useful tool to support the planning of warehouse processes in the subject company.

References


[24] NASSIEF W., CONTERRAS I., JUAMARD B., 2018, A comparison of formulations and re-


