MODELING OF TRANSPORT PROCESSES IN TERMS OF SEASONALITY OF TRANSPORT

Andrzej Świderski

Military University of Technology, Faculty of Logistics, Urbanowicza 2 Str., 00-908 Warsaw, Poland; e-mail: andrzej.swiderski@wat.edu.pl; ORCID ID 0000-0001-7451-9161

ABSTRACT

The company's success depends on many factors. One of them is the ability to meet the customers' expectations and adjust to the needs reported by market. Mathematical methods and tools are helpful in assessing demand. Forecasts should take into account all factors shaping the demand for goods and services, but they are often difficult to define, not only because of their large number, but also due to the impact of individual variables, which is difficult to determine. In many cases, the number of placed orders strongly depends on the time at which they are submitted. The needs can vary according to the time of day, week and year. Then we are dealing with the so-called seasonality, the inclusion of which is very important in the company and enables to better adapt its activity to the customers' requirements.

This article describes the seasonality of demand in a company providing domestic and road transport services with high-tonnage vehicles. The validity of such analyses and potential benefits were indicated.

Key words:

demand forecasting, seasonality, road transport, multiple regression.

Research article

INTRODUCTION

The demand for transport services is affected by many factors that cause its significant fluctuation. These fluctuations are often subject to repeated changes, observed in short, medium and long time intervals. They are shaped by many economic, legal, social and climate factors. In addition, they are strongly conditioned by the volatility of the demand for products offered by the sectors that the TSL industry supports. This cyclical change is an opportunity for transport companies for better adaptation towards market needs. The tools and statistical and mathematical methods that help describe the occurring phenomena and draw conclusions on the basis of the appropriate wording and, if necessary, modification of the company's strategy, are useful.

In the article, based on the company offering transport services, the method of demand analysis was presented. For this purpose, the multiple regression method was used. The forecast was prepared on the basis of the observations concerning the demand for the transport service over the last three years. The presented company is only an example. An analysis of the literature indicates that the need to model the level of use of transport means is not only important for civil enterprises [3, 5], but also for state institutions. Such models can be used in forecasting the fleet's readiness, for example firefighting vehicles [3] or armed forces [4, 6]. They also work in forecasts regarding the passable roads and possible disturbances resulting, for example, from congestion [14] or road accidents [8, 17, 19].

THE METHOD OF THE ANALYSIS

The analysis of phenomena and processes is possible to model using stochastic processes. The implementation of the stochastic process in the field of time is called the time series [10, 20]. It is a sequence of information arranged in time. The individual measurements are recorded with the exact time step and constitute a set of observations representing the implementation of the phenomenon under investigation and characterizing the changes occurring in it. It is also necessary to identify all elements in a row that can be systematic components such as trend, periodic or cyclical fluctuations as well as accidental random elements. This process is called the decomposition of the time series [3, 15].

The analyzed example uses multiple regression. It belongs to the group of analytical models in which it is necessary to find mathematical dependencies — functions reflecting the implementation of the process, which also allows to extend the analysis

to arguments outside the scope of collected empirical data and to determine how the phenomenon will be shaped in the future. Estimation involves finding the right trend function, and then describing and isolating (if there are) seasonal and cyclical fluctuations. The simplest form of the trend is the linear function, however, the trend can also be estimated using another function, for example exponential, square or power, depending on the structure of the phenomenon being studied. However, the trend function is not sufficient if there are fluctuations in the process. As mentioned above, they may concern changes resulting from the calendar, determining a certain rhythm of phenomena: daily, weekly, monthly, etc. then they are seasonal fluctuations or shaped by economic factors — that is, cyclical fluctuations. The estimation of model parameters uses the classic method of least squares (MLS) or the maximum likelihood estimation (MLE) [2, 6].

One of the analytical methods of assigning the value of a dependent variable to specific values of independent variables is regression. Its simplest type is simple linear regression, describing the relationships between variables using a straight line (1).

$$\hat{y} = \beta_0 + \beta_1 x + \varepsilon, \tag{1}$$

where:
β_1 — directional factor;
β_0 — free word (intersection with the ordinate);
x — independent variable;
y — dependent variable (explained, predicted);
ε — random error.

If there are more explanatory variables, then we deal with multiple regressions, and the multiple linear regression model takes the form (2):

$$\hat{y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n + \varepsilon, \tag{2}$$

where:
β_0 — free word;
β_i — model parameters — regression coefficients;
x — independent variable;
y — dependent variable (explained, predicted);
ε — random error.

Regression coefficients describe how much the value of the dependent variable *y* changes on average, if the value of the independent variable *x* to which they refer changes to the unit, assuming a set level of other independent variables.

Independent variables in the analysis of economic phenomena are often not quantitative, but qualitative, as in the analyzed example. In this situation, it is necessary

to recode them to zero-one variables. Because their number is limited, they cannot be treated in the manner adopted for continuous variables in regression, which results from the fact that they do not have an economic sense, and the calculated model coefficients do not have an economic interpretation. Then, such a qualitative or discrete variable, having from a few or a dozen or so categories, is coded to the appropriate number of q binary variables (zero-one), which are used in the regression equation. However, in order to use the MLS method for estimating the model, q-1 artificial variables should be used, since the introduction of q variables, i.e. in an amount equal to the explanatory variable, will result in a linear relationship between the regressors and the X'X matrix will be singular. This is due to the fact that the zero-one variables add up to unity. This phenomenon is referred to in the econometric literature as a trap associated with zero-one variables (dummy variable trap) [14, 17]. Such a model is not possible to estimate and therefore the number of artificial variables must always be less by one in relation to the number of categories q (levels) identified for a given attribute (trait). Only then the estimated model is correct and consists of the free word β_0, the sum of products of structural parameters and binary variables D_k number $k = 1 \; ;; \; q$, representing seasonality, and the random component ε (3).

$$\hat{y} = \beta_0 + \beta_1 t + \delta_1 D_1 + \cdots + \delta_k D_k + \varepsilon. \tag{3}$$

CASE STUDY

The transport carried out in the enterprise over the last three years has been analyzed. It has 30 vehicles with a maximum load capacity of 24 t. Their mileage is shown in fig. 1.
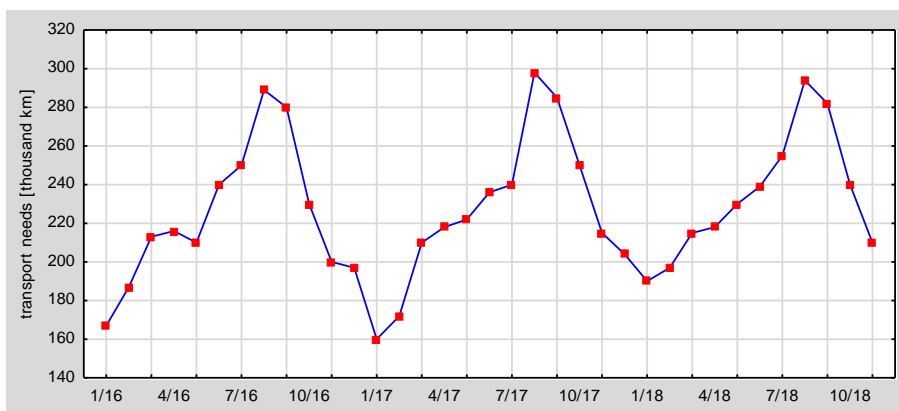

Fig. 1. The demand for transport services in 2016–2018

The table shows a high seasonality of transport. This is confirmed by selected measures of descriptive statistics, calculated for individual months, presented in tab. 1.

Tab. 1. Basic measurements of descriptive statistics

Month	Average [thous. km]	Median [thous. km]	Minimum [thous. km]	Maximum [thous. km]	Stand. dev. [thous. km]	Ceofficient of var. [%]
January	172	167	160	190	15.70	9.11
February	185	187	172	197	12.58	6.79
March	213	213	210	215	2.52	1.18
April	217	218	216	218	1.15	0.53
May	221	222	210	230	10.07	4.56
June	238	239	236	240	2.08	0.87
July	248	250	240	255	7.64	3.08
August	294	294	289	298	4.51	1.54
September	282	282	280	285	2.52	0.89
October	240	240	230	250	10.00	4.17
November	208	210	200	215	7.64	3.67
December	201	200,5	197	204	4.95	2.47

The largest transport needs occur in August and September. The lowest indications, however, were recorded in January and February. Differences in demand for particular months are well reflected in the monthly volatility graph presented in fig. 2.
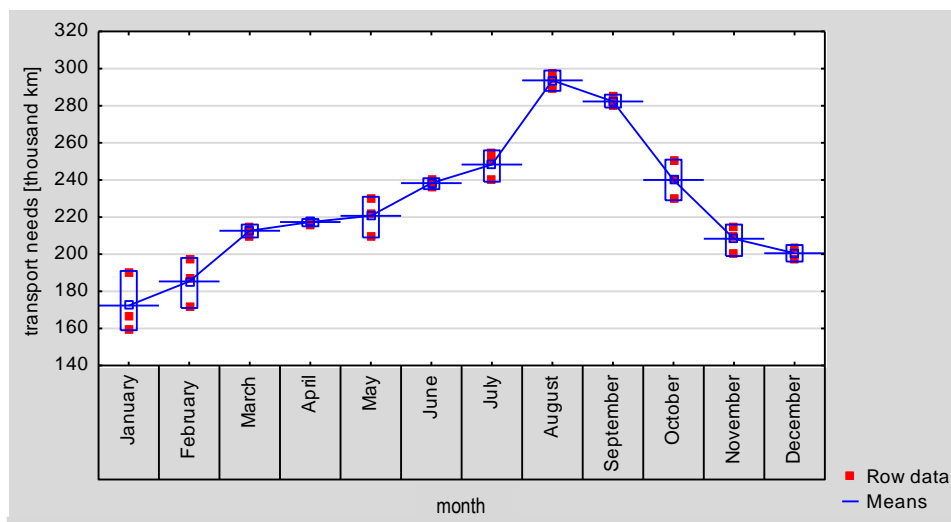

Fig. 2. Graph of volatility of average transport needs in particular months

Preparing the company for the occurring fluctuations therefore requires their inclusion in the forecasted demand. This is possible with the aforementioned multiple regression model, exactly the type of regression in which the binary variables will be used. In addition, fig. 1 suggests the existence of a development trend, however, seasonal variations make its visual separation difficult. To assess the occurrence of the trend in the average value of the series, it was proposed to use a mechanical method of its determination using moving averages. The simple moving averages were determined: 3-periodic, 6-periodic and 12-periodic, and charts were obtained as in fig. 3–5.
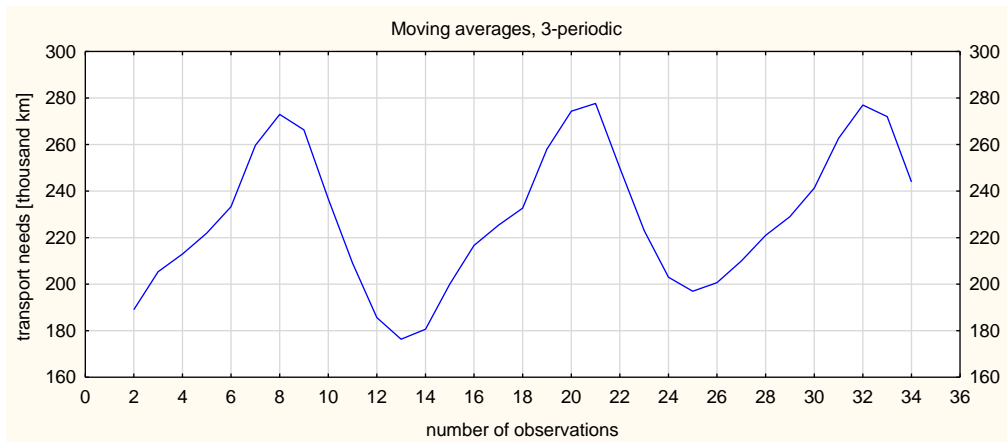

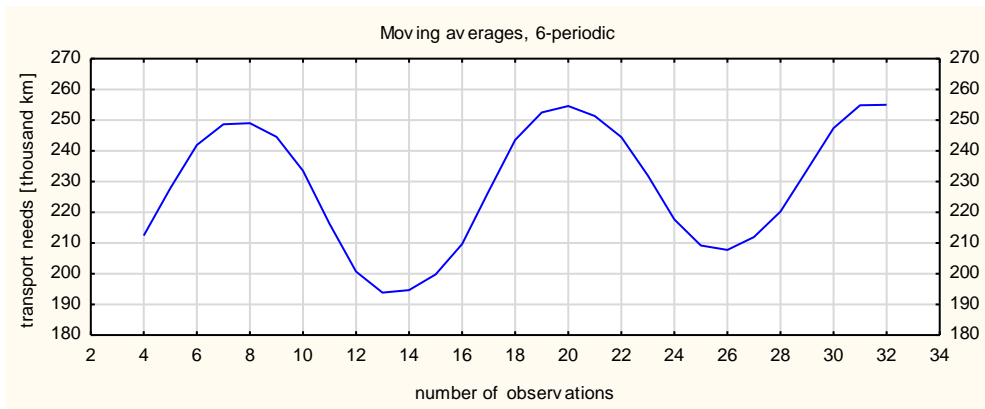Fig. 3. Moving averages, 3-periodic
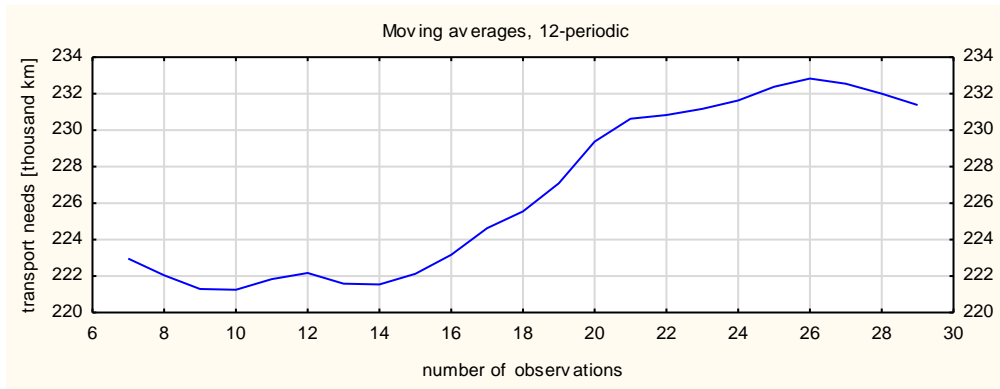

Fig. 4. Moving averages, 6-periodic

Fig. 5. Moving averages, 12-periodic

Already in the figure with moving 6-period averages, the existence of the trend and its increasing character can be seen, which is finally confirmed by the graph with 12-period averages. The above analyses indicate, therefore, that the model should describe the long-term trend and short-term seasonality, for particular months of the year. As mentioned earlier, the qualitative nature of the explanatory variable makes it necessary to estimate a model in which the number of variables will be lowered by one. Such a model consists of the free word, trend, and the sum of eleven products of structural parameters and binary variables D_k for $k\varepsilon$ {1,11} representing seasonality, as in the following formula (4)

$$y = a_0 + a_1\, t + b_1\, D_1 + \cdots + b_k D_k. \qquad (4)$$

In most publications, it is proposed to estimate a model lowered by the variable with the lowest indication [10, 15]. In this case, it is January. This means that estimates of seasonal parameters will refer to the 'January' level. Therefore, all parameter values for particular months will be positive, as each of them will have a higher than January level. The results of the estimation of multiple regression parameters and estimation errors are presented in tab. 2.

The adjusted coefficient of determination, which indicates what percentage of variability of the dependent variable (Y — explained) is explained by the independent variable (X — explanatory) is satisfactory and amounts to 96%, the errors of parameter estimates are small, at about 6%. In addition, the quality of the model is evidenced by the fact that all the estimated parameters are statistically significant, which is also satisfactory due to the substantive interpretation of the model and no need to remove irrelevant explanatory variables. The graph of empirical and forecasted values is presented in fig. 6.

Tab. 2. Estimate results of multiple regression model

| $N = 35$	$R^2 = 0.97$, adjusted $R2 = 0.96$ F(12.22) = 66.750 $p <0.00000$ Standard error: 7.3084			
	b	SE for b	t(22)	p
a_0	167.93	4.54	37.01	0.000000
trend	0.34	0.13	2.64	0.014975
January	12.66	5.97	2.12	0.045404
February	39.66	5.97	6.64	0.000001
March	43.98	5.98	7.36	0.000000
April	46.98	5.99	7.84	0.000000
May	64.31	6.00	10.71	0.000000
June	73.97	6.02	12.29	0.000000
July	118.96	6.03	19.71	0.000000
August	107.29	6.06	17.72	0.000000
September	64.62	6.08	10.63	0.000000
October	32.61	6.10	5.34	0.000023
November	26.47	6.70	3.95	0.000682
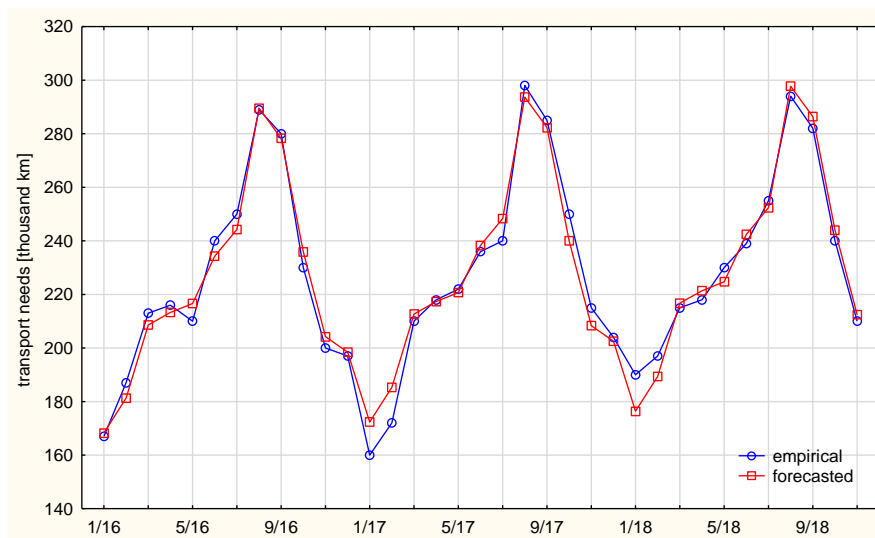
Fig. 6. Graph of empirical and forecasted values

The forecast was the worst for January, it is the month in which the forecast errors received are the largest. It is influenced by both a small number of observations and the highest of all coefficients of variation. Further research and estimation of the model from a longer period would make the model more credible.

DIAGNOSIS OF THE MODEL

The basis for assessing the accuracy of the fit of the theoretical function to the empirical data is to analyze the differences between empirical and theoretical values, called model residues. Confirmation of the correctness of the regression model requires, therefore, reviewing basic assumptions regarding residues, which include examining the normality of their distribution and the existence of significant dependencies of the autocorrelation function.

Fig. 7 presents a histogram of the distribution of residuals which shows that the distribution is close to normal, which is confirmed by the calculated statistics of the Shapiro-Wilk test, whose design value is 0.9861 and test probability $p = 0.9288$, which means that there are no grounds for rejecting the $h0$ hypothesis at the significance level $\alpha = 0.05$, which tells us that the distribution of the variable is close to normal. Similar results were obtained for the Lilliefors test for which the calculation value is 0.08424, with the value of $p > 0.2$ confirming that the residual distribution is normal.
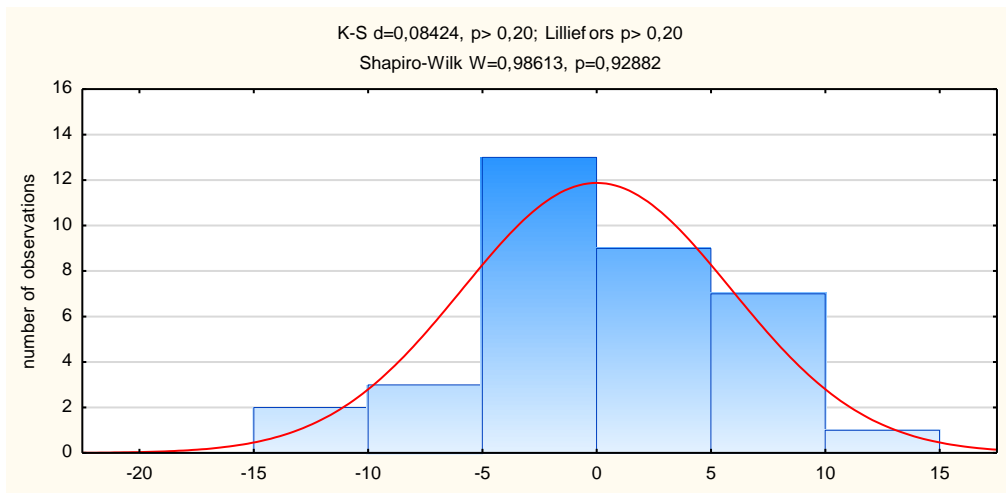

Fig. 7. Histogram of model residues

Another factor that decides about the correctness of the model is the lack of statistically significant correlations between the residual values, or confirmation that there was no unexplained dependency model. For this purpose, the graphs of the autocorrelation and partial autocorrelation functions presented in fig. 8 and fig. 9 were examined. Since all the values of the functions are statistically insignificant, the correct model construction was confirmed in this way.
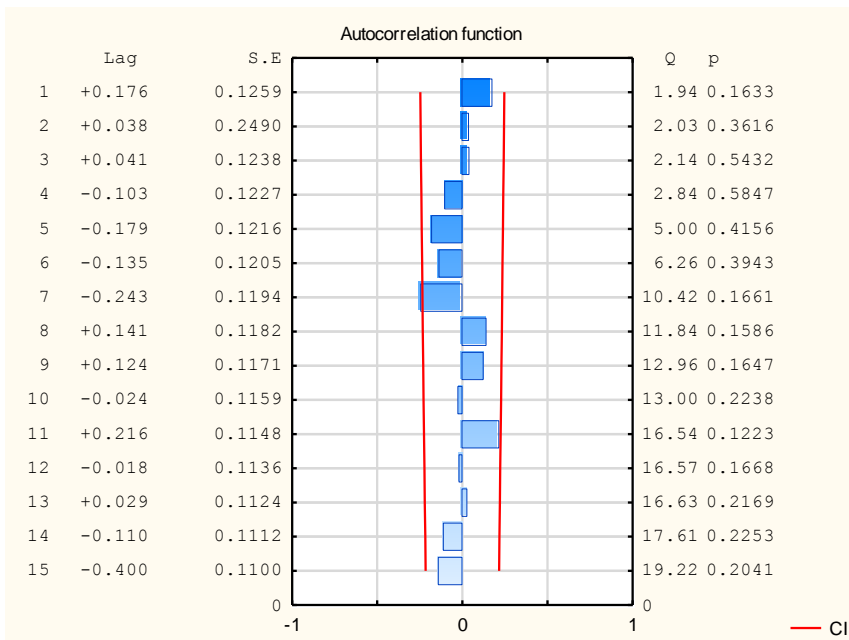
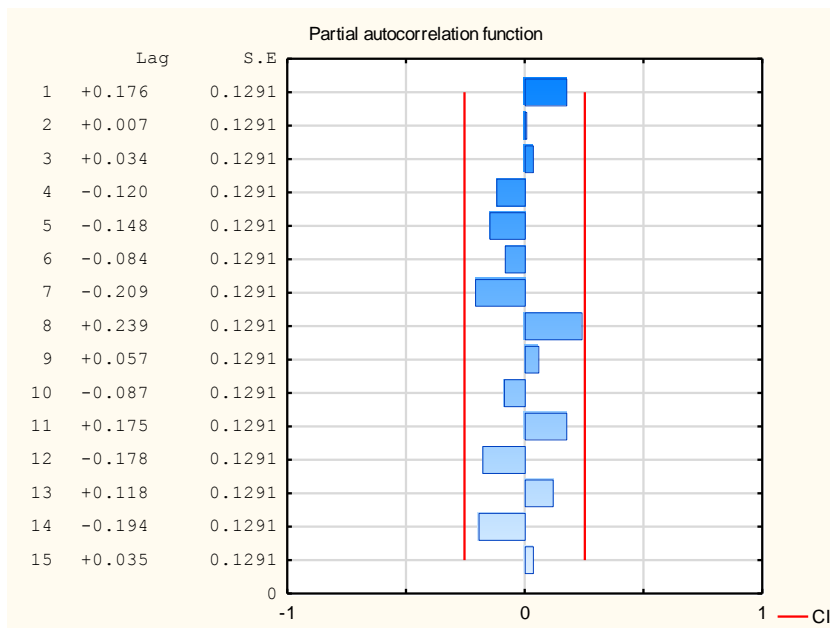Fig. 8. Autocorrelation function of regression model residues

Fig. 9. Partial autocorrelation function of regression model residues

On the basis of the results obtained, a forecast of transport needs for 2019 has also been introduced, presented in fig. 10.
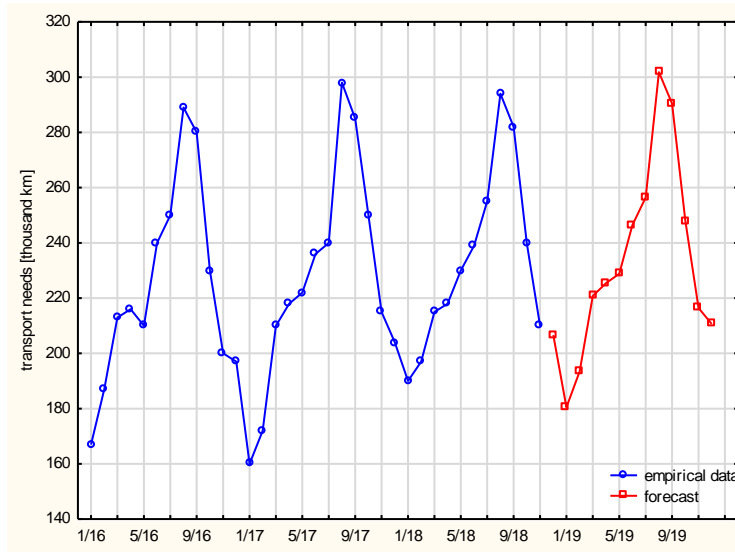

Fig. 10. Forecast of transport needs for 2019

The analysis of the chart in fig. 10 again indicates the underestimation of the forecast for January. For many industries this is a difficult moment, which is reflected in the demand for transport services. It should also be emphasized that each forecast must be monitored, verified and adjusted to the dynamically changing market. In addition, the forecast function is only advisory, allows you to indicate the main directions of change and provides excellent support for management processes. However, it does not provide decision readiness.

CONCLUSIONS

The study showed the possibility of forecasting transport needs in a transport company using multiple regression. The proposed model revealed not only the clear seasonality of the process, but also the existing upward trend. For an enterprise, especially one that operates its own transport (as tested), this information is very important. First of all, it shows that there are months when transport potential may not be fully utilized, and on the other hand, that constantly growing demand may in the near

future cause that your fleet does not fully meet your needs, especially in the months in which they are the greatest. This means that, it is necessary to verify the company's transport strategy and propose directions that minimize the level of unused potential of vehicles on one hand and securing the interests of the company in case of demand greater than the transport possibilities on the other.

REFERENCES

[1]   Bielińska E., *Prognozowanie ciągów czasowych*, Publ. Politechnika Śląska, Gliwice 2007 [*Time series forecasting* — available in Polish].

[2]   Bitner A., *Konstrukcja modelu regresji wielorakiej przy wycenie nieruchomości*, 'Acta Scientiarum Polonorum. Administratio Locorum', 2007, Vol. 6, No. 4, pp. 59–66 [*Construction of a multiple regression model for property valuation* — available in Polish].

[3]   Borucka A., *Forecasting of fire risk with regard to readiness of rescue and fire-fighting vehicles*, 'Interdisciplinary Management Research XIV', Croatia, 2018, pp. 397–395.

[4]   Borucka A., *Funkcjonowanie wojskowych oddziałów gospodarczych w nowym systemie logistycznym sił zbrojnych*, 'Logistyka', 2013, No. 6, pp. 39–48 [*Functioning of military economic units in the new armed forces logistics system* — available in Polish].

[5]   Borucka A., *Markov models in the analysis of the operation process of transport means*, Proceedings of the ICTTE International Journal for Traffic and Transport Engineering Conference, Belgrad, 2018, pp. 1073–1082.

[6]   Borucka A., *Model of the operation process of aircraft in the transport system*, Proceedings of the ICTTE International Journal for Traffic and Transport Engineering Conference, Serbia, 2018, pp. 22–30.

[7]   Borucka A., *Przewóz osób w Polsce w świetle zmian legislacyjnych*, 'Logistyka', 2014, No. 6, pp. 468–475 [*Transport of people in Poland in the light of legislative changes* —available in Polish].

[8]   Borucka A., *Risk Analysis of Accidents in Poland Based on ARIMA Model*, 'Transport Means', 2018, Proceedings of the 22nd International Scientific Conference, Part I, Lithuania, 2018, pp. 162–166.

[9]   Borucka A., *Three-state Markov model of using transport means*, Proceedings of the 18th International Scientific Conference 'Business Logistics In Modern Management', Croatia, 2018, pp. 3–19.

[10] Dittmann P., *Metody prognozowania sprzedaży w przedsiębiorstwie*, Publ. Akademia Ekonomiczna, Wrocław 200.0 [*Methods of sales forecasting in the enterprise* — available in Polish].

[11] Dittmann P., Szabela-Pasierbińska E., Dittman I., Szpulak A., *Prognozowanie w zarządzaniu sprzedażą i finansami przedsiębiorstwa*, Wolters Kluwer, Warszawa 2011 [*Forecasting in the management of sales and finances of the company* — avaliable in Polish].

[12] Maciąg A., Pietroń R., Kukla S., *Prognozowanie i symulacja w przedsiębiorstwie*, PWE, Warszawa 2013 [*Forecasting and simulation in the enterprise* — avaliable in Polish].

[13] Mikosz B., Borucka A., *Organizacja gospodarki odpadami w siłach zbrojnych na tle zmian militarnych i nowych wyzwań stawianych polskiej armii*, 'Archiwum Gospodarki Odpadami i Ochrony Środowiska', 2008, No. 8, pp. 1–12 [*Organization of waste management against the background of military changes posed to the Polish Army* — available in Polish].

[14] Mitkow S., Borucka A., *Mathematical model of travel times related to a transport congestion: an example of the capital city of Poland — Warsaw*, Proceedings of the 18th International Scientific Conference 'Business Logistics In Modern Management', Croatia, 2018, pp. 501–526.

[15] Sokołowski A., *Prognozowanie i analiza szeregów czasowych*, teaching materials StatSoft, Kraków 2016 [*Time series forecasting* — available in Polish].

[16] StatSoft, *StatSoft Electronic Statistics Textbook*, Kraków 2006, [online], http://www.statsoft.pl/textbook/stathome.html [access 08.01.2019].

[17] Świderski A., Borucka A., Jacyna-Gołda I., Szczepański E., *Wear of brake system components in various operating conditions of vehicle in the transport company*, 'Eksploatacja i Niezawodność — Maintenance and Reliability', 2019, Vol. 21, No. 1, pp. 1–9.

[18] Świderski A., Borucka A., *Mathematical Analysis of Factors Affecting the Road Safety in Selected Polish Region*, 'Transport Means', Proceedings of the 22nd International Scientific Conference, Part II, Lithuania, 2018, pp. 651–654.

[19] Świderski A., Borucka A., Skoczyński P., *Characteristics and Assessment of the Road Safety Level in Poland with Multiple Regression Model*, 'Transport Means', Proceedings of the 22nd International Scientific Conference, Part I, Lithuania, 2018, pp. 92–97.

[20] Wielgosik M., Borucka A., *Istota i znaczenie służby przygotowawczej i szkolenia rezerw*, 'Systemy Logistyczne Wojsk', 2016, No. 45, pp. 51–66 [*The essence and importance of the preparatory and reserve training service* — available in Polish].

MODELOWANIE PROCESÓW TRANSPORTOWYCH Z UWZGLĘDNIENIEM SEZONOWOŚCI PRZEWOZÓW

STRESZCZENIE

Sukces przedsiębiorstwa zależy od wielu czynników. Jednym z nich jest zdolność do spełnienia oczekiwań klientów i dopasowanie do potrzeb zgłaszanych przez rynek. Pomocne w ocenie popytu są metody i narzędzia matematyczne. Formułowane prognozy powinny brać pod uwagę wszystkie czynniki kształtujące zapotrzebowanie na dobra i usługi, jednak często są one trudne do zdefiniowania, nie tylko ze względu na znaczną ich liczbę, ale również na trudny do określenia wpływ poszczególnych zmiennych. W wielu przypadkach liczba złożonych zamówień silnie zależy od czasu, w którym są one składane. Potrzeby mogą być zróżnicowane w zależności od pory dnia, tygodnia czy roku. Wówczas uaktywnia się tzw. sezonowość, której uwzględnienie jest bardzo istotne w przedsiębiorstwie i pozwala lepiej dostosować jego działalność do wymagań klientów.

Andrzej Świderski

W artykule dokonano charakterystyki sezonowości popytu w firmie świadczącej usługi transportu krajowego drogowego realizowanego pojazdami wysokotonażowymi. Wskazano zasadność prowadzenia takich analiz i potencjalne korzyści.

Słowa kluczowe:

prognozowanie popytu, sezonowość, transport drogowy, model regresji wielorakiej.