Bartosz Kozicki
Wojskowa Akademia Techniczna im. Jarosława Dąbrowskiego
e-mail: Bartosz.kozicki@wat.edu.pl
Jacek Wozniak
Wojskowa Akademia Techniczna im. Jarosława Dąbrowskiego
e-mail: Jacek.wozniak@wst.edu.pl
Jarosław Tomaszewski
Ministerstwo Obrony Narodowej
e-mail: jarek7@wp.pl

FORECASTING RESIDENTS KILLED IN ROAD ACCIDENTS IN EUROPE BETWEEN 2018-2020

ABSTRACT

Background: The article discusses the problem of forecasting residents killed in road accidents. The observation of data contained in the literature allows for the statement that in the last dozen years or so in Europe there is a decreasing trend of the number of people killed in road accidents. The prediction for the future is important from the point of view of the article. The purpose of this article is to attempt to predict data on 100,000 inhabitants killed in road accidents in thirty-two European countries between 2018-2020.

Methods: The article uses research methods: literature analysis of the subject, which deals with issues related to road transport, forecasting, analysis of source documents, computer simulation methods and comparisons in groups. The research began with the analysis and evaluation of primary data. The original data for the study was obtained from the Eurostat website. Then, through the detected dependencies in the time series, dynamic forecasting methods were selected, which were analyzed and evaluated. In addition, a research technique in the form of the Statistica computer program and the following research tools were used: quantile graph, autocorrelation, partial autocorrelation, multiple regression and histogram.

Results: The results of the primary time series forecast for 2018-2020 were obtained by three methods: Holt exponential smoothing [Dittmann P. et al., 2016, Forecasting in sales and financial management of an enterprise, Wydawnictwo Nieoczywiste, imprint GAB Media, p.

P. Cyplik et al. (eds.), *Contemporary challenges in supply chains, Vol. 2*,
https://doi.org/10.17270/B.M.978-83-66017-89-4.5

69-70], Klein model and multi-model forecasting (as the best one). They confirm the strong trend of the declining trend.

Conclusions: The use of the multi-model method for forecasting, with the observed dependencies in the considered primary time series, may improve planning in terms of predicting the amount of deaths in road accidents, so as to minimize their occurrence, by implementing appropriate investment expenditure to improve safety in this matter.

Key words: methods, forecasting, dead, road transport.

INTRODUCTION

One of the elements of logistics is transport, which in literature is considered one of the most important aspects of the logistics system. Transport is a production process aimed at overcoming space [Jacyna M. and Lewczuk K., 20016, p. 88]. It is considered the foundation of the world economy and society [Gołembska 1998, p. 7]. One of the types of transport is road.

In the European Union countries, economic growth is observed, which is fueled by the development of road transport - freight and passenger. The road infrastructure of Europe in which vehicle users travel varies, and factors affecting traffic safety are also different in European countries. The basis for addressing this issue was the Internet research conducted as part of the ESRA1 (European Survey of Road users' safety Attitudes) program by the Belgian VIAS Institute with a population of 35,000. users (993 in Poland) from 32 countries in 2018 [www.esranet.eu]

What are the causes of road accidents? The answers were obtained on the basis of the questions contained in the surveys. The list of issues included in the survey concerned:

- driving after consuming alcohol, drugs and some medicines;
- speeding;
- use of seat belts;
- distraction;
- cycling, motorcycle without a helmet;
- pedestrian behavior on the road,

which were the cause of road accidents, including fatalities.

The purpose of the article is to attempt to forecast data concerning people killed per 100,000 inhabitants in road accidents in thirty-two European countries between 2018-2020. The observation of data contained in the literature allows to state that in the last dozen years or so

P. Cyplik et al. (eds.), *Contemporary challenges in supply chains, Vol. 2*,
https://doi.org/10.17270/B.M.978-83-66017-89-4.5

in Europe there is a decreasing tendency of the number killed in road accidents. From the point of view of the article, their prediction for the future becomes important. Primary data for the study was obtained from the Eurostat website. The article uses research methods in the form of literature analysis, which deals with issues related to road transport, forecasting, analysis of source documents, computer simulation methods and comparison. In addition, a research technique was used in the form of the Statistica computer program and the following research tools: quantile graph, autocorrelation, partial autocorrelation, multiple regression, histogram. The article includes introduction, four substantive points, summary and conclusions.

ANALYSIS OF STATISTICAL DATA

Research included the collection of data concerning road fatalities in 32 European countries per 100,000 inhabitants between 2010-2017. The main research effort focused on:

- analysis of statistical data on those killed, taking into account the trend and time distribution;
- building a multiple regression model;
- forecasting methods;
- assessment of the phenomenon.

As part of the research procedure, the above methods were used with varying degrees to demonstrate the best method in the forecasting process. Figure 1 presents a categorized frame-mustache graph with raw data on those killed in road accidents in thirty-two selected European countries per 100,000 inhabitants in 2010-2017.

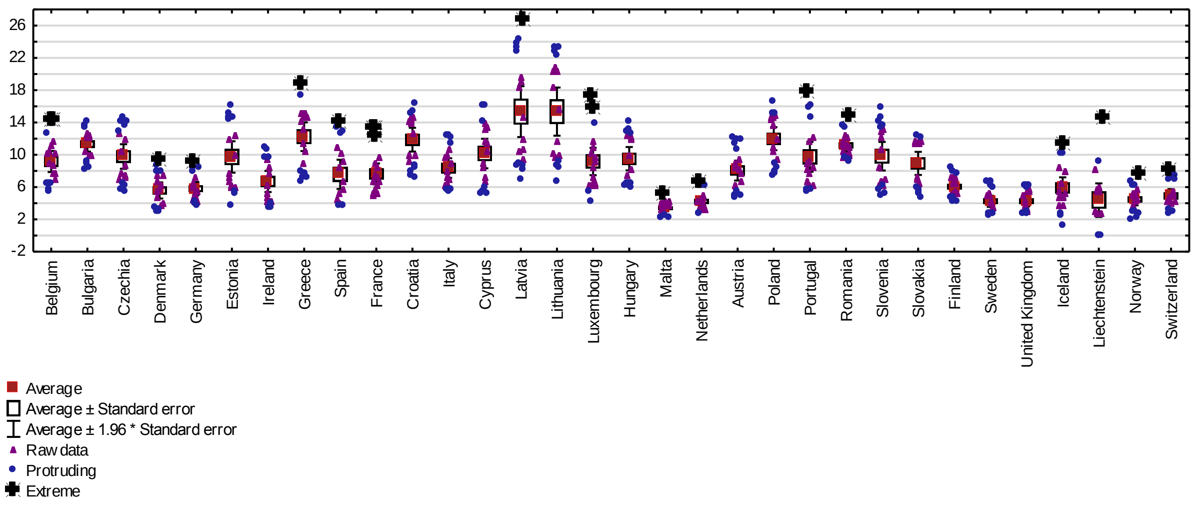


Fig.1. Categorized frame-mustache chart with raw data of killed in car accidents in selected European countries in 2000-2017 per 100,000 inhabitants.
Source: own study based on data obtained from the website:
https://ec.europa.eu/eurostat/tgm/download.do?tab=table&plugin=1&language=en&pcode=sdg_11_40

P. Cyplik et al. (eds.), *Contemporary challenges in supply chains, Vol. 2*,
https://doi.org/10.17270/B.M.978-83-66017-89-4.5

The interpretation of Figure 1 shows that the highest accident rate per 100,000 inhabitants in the analyzed European countries was in Lithuania and Latvia. These countries showed the highest inter-quantile range. In each of the analyzed countries there are visible values deviating from the arithmetic mean. However, in half of the analyzed countries extreme values are observed. The lowest level of the dead is visible in Malta. For research purposes, data on the dead were aggregated in individual countries and grouped in 2000-2017. The results are shown in Figure 2.

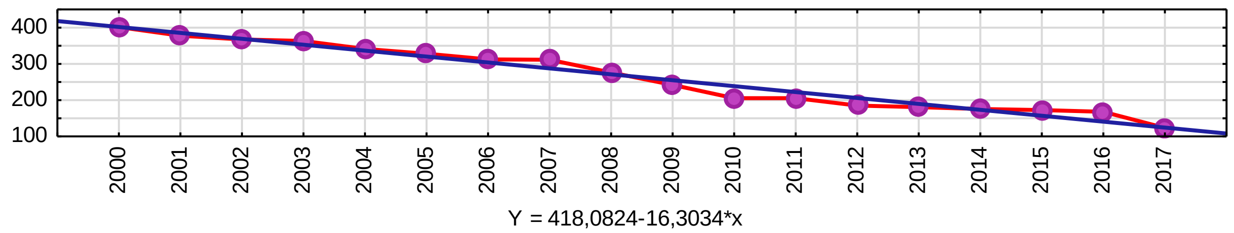


Fig.2. Line graph of the sum of people killed in car accidents in selected European countries in 2000-2017 per 100,000 inhabitants, along with a trend line.
Source: own study based on data obtained from the website:
https://ec.europa.eu/eurostat/tgm/download.do?tab=table&plugin=1&language=en&pcode=sdg_11_40

The analysis of data on road fatalities per 100,000 inhabitants (named by the authors with the primary series) summarized in Figure 2 allowed for the finding of a downward trend. The observed trend was outlined with a blue line and described by the function Y = 418.0824-16.3034 * x. The next stage of the analysis was to examine the existence of outliers and extremes in the primary series. For this purpose, a frame-mustache chart was drawn together with raw data Figure.

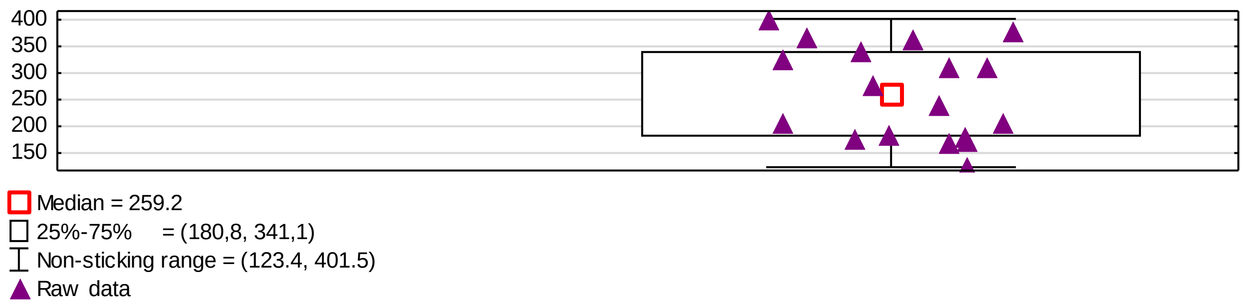


Fig. 3. Frame-mustache chart together with outlined random data on the sum of people killed in car accidents in selected European countries in 2000-2017 per 100,000 inhabitants.
Source: own study based on data obtained from the website:
https://ec.europa.eu/eurostat/tgm/download.do?tab=table&plugin=1&language=en&pcode=sdg_11_40

The analysis of Figure 3 shows that there are no outliers and extreme values.

P. Cyplik et al. (eds.), *Contemporary challenges in supply chains, Vol. 2*,
https://doi.org/10.17270/B.M.978-83-66017-89-4.5

The next stage of the analysis was to examine the distribution. For this purpose, two research tools were used in the form of: a histogram Figure 4 and a normality chart together with the Shapiro-Wilk (SW-W) test Figure 5.

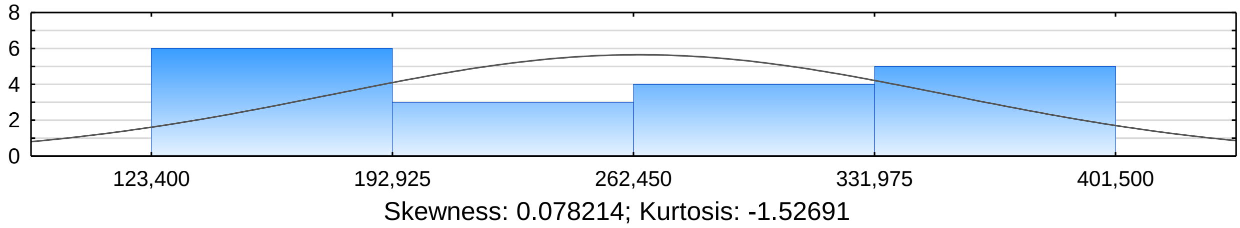


Fig. 4. Histogram of the sum of people killed in car accidents in selected European countries in 2000-2017 per 100,000 inhabitants.
Source: own study based on data obtained from the website:
https://ec.europa.eu/eurostat/tgm/download.do?tab=table&plugin=1&language=en&pcode=sdg_11_40

As shown in Figure 4, the distribution of the analyzed series looks like bimodal. What is more, it can be assumed that it is the right side which is oblique and more flattened than normal.

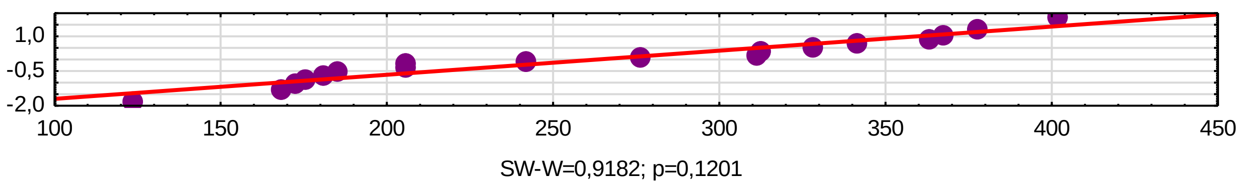


Fig.5. Graph of normality together with the Shapiro-Wilk test of the sum of people killed in car accidents in selected European countries in 2000-2017 per 100,000 inhabitants.
Source: own study based on data obtained from the website:
https://ec.europa.eu/eurostat/tgm/download.do?tab=table&plugin=1&language=en&pcode=sdg_11_40.

The quantiles are close to the straight line outlined (Figure 5). The Shapiro-Wilk test clearly indicates that there are no grounds to reject the null hypothesis about the normality of the distribution of the analyzed time series. Thus, it was assumed that the distribution of the original time series is normal.

For research purposes, it was decided to investigate the existence of a trend in the original time series. For this purpose, a multiple regression model was built consisting of three predictors in the form of variables: t, t2 and lnt. Significant predictors are summarized in Table 1.

Table 1. Multiple regression model of the sum of people killed in car accidents in selected European countries in 2000-2017 per 100,000 inhabitants

N=18	R= ,98 R^2= ,97 adjusted R^2= ,96 Standard estimation error: 15,76					
	b*	Standard error	b	Standard error	t(16)	p
Word free			418,08	7,75	53,93	0,00
t	-0,98	0,04	-16,30	0,72	-22,77	0,00

P. Cyplik et al. (eds.), *Contemporary challenges in supply chains, Vol. 2*,
https://doi.org/10.17270/B.M.978-83-66017-89-4.5

Source: own study based on data obtained from the website:
https://ec.europa.eu/eurostat/tgm/download.do?tab=table&plugin=1&language=en&pcode=sdg_11_40

The built multiple regression model in Table 1 is very well suited. Multiple R^2 was 0.96 and the standard error of estimation was 15.76. An important predictor was the variable t, whose regression coefficient was -16.30, which clearly confirms the decreasing linear trend. For research purposes, the rest of the analyzed model was examined. For this purpose, the following research tools were used: plot of predicted and observed values in Figure 6, line graph of model residuals in Figure 7, autocorrelation in Figure 8, partial autocorrelation in Figure 9, histogram in Figure 10 and normality chart with Shapiro test -Wilk in Figure 11.

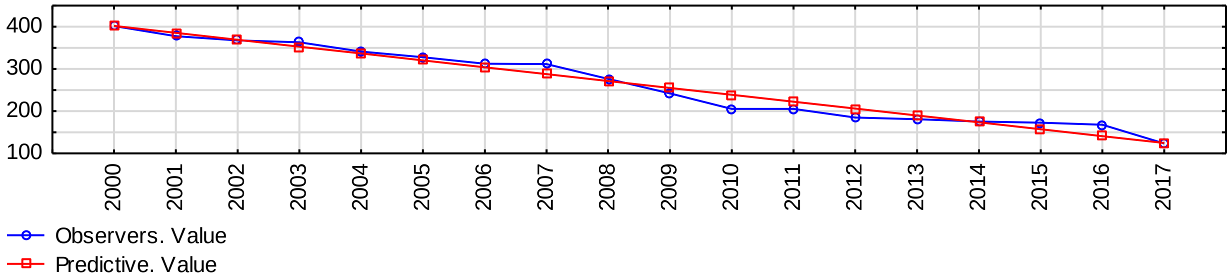


Fig. 6. Chart of predicted and observed sums of people killed in car accidents in selected European countries in 2000-2017 per 100,000 inhabitants.
Source: own study based on data obtained from the website:
https://ec.europa.eu/eurostat/tgm/download.do?tab=table&plugin=1&language=en&pcode=sdg_11_40

Figure 6 shows that there is a good match between predicted and observed values.

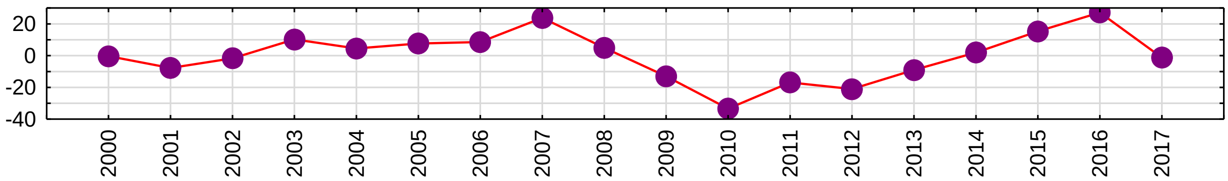


Fig. 7. Residual graph of multiple regression model.
Source: own study based on data obtained from the website:
https://ec.europa.eu/eurostat/tgm/download.do?tab=table&plugin=1&language=en&pcode=sdg_11_40

The rest of the multiple regression model (Figure 7) is both positive and negative with a long mustache, which may directly indicate the existence of dependence.

P. Cyplik et al. (eds.), *Contemporary challenges in supply chains, Vol. 2*,
https://doi.org/10.17270/B.M.978-83-66017-89-4.5

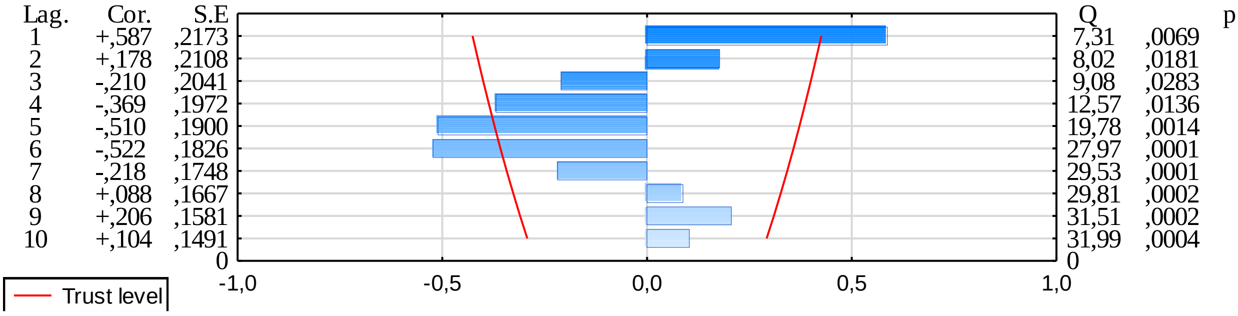


Fig. 8. Autocorrelation of residuals of the multiple regression model.
Source: own study based on data obtained from the website:
https://ec.europa.eu/eurostat/tgm/download.do?tab=table&plugin=1&language=en&pcode=sdg_11_40

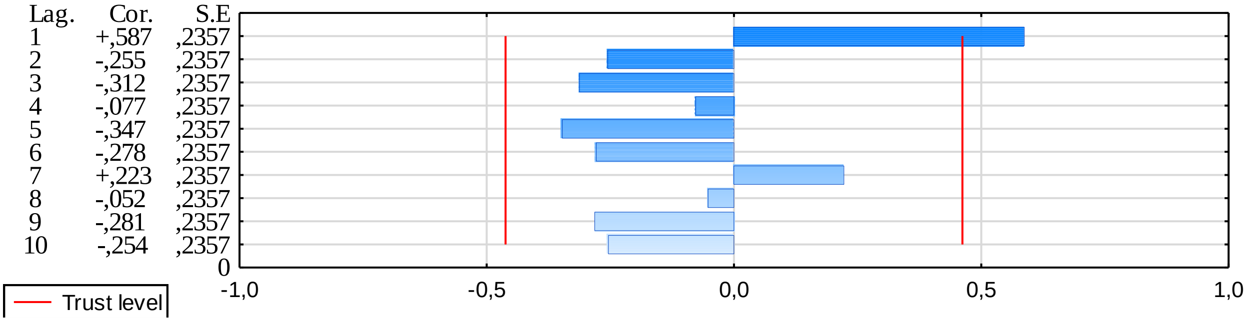


Fig 9. Partial autocorrelation of multiple regression model residues.
Source: own study based on data obtained from the website:
https://ec.europa.eu/eurostat/tgm/download.do?tab=table&plugin=1&language=en&pcode=sdg_11_40

Autocorrelation of residues in Figure 8 and partial autocorrelation in Figure 9 confirm the existence of dependencies in the residues of the built model, as do Statistics P and Q.

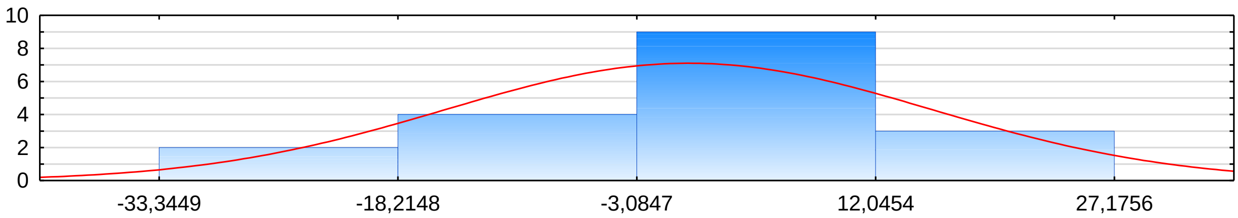


Fig.10. Histogram of multiple regression model residuals.
Source: own study based on data obtained from the website:
https://ec.europa.eu/eurostat/tgm/download.do?tab=table&plugin=1&language=en&pcode=sdg_11_40

The assessment of Figure 10 is that the distribution of the analyzed data is more slender than normal.

P. Cyplik et al. (eds.), *Contemporary challenges in supply chains, Vol. 2*,
https://doi.org/10.17270/B.M.978-83-66017-89-4.5

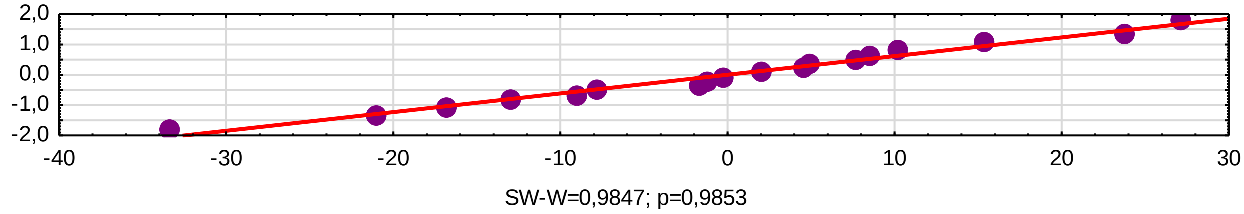


Fig. 11. Normality chart with the Shapiro-Wilk test of the multiple regression model residues.
Source: own study based on data obtained from the website:
https://ec.europa.eu/eurostat/tgm/download.do?tab=table&plugin=1&language=en&pcode=sdg_11_40

The distribution of residuals of the multiple regression model is normal in Figures 10 and 11. The next stage of the analysis was to examine the existence of dependencies in individual delays of the original time series. The following research tools were used for this purpose: autocorrelation in Figure 12 and partial autocorrelation in Figure 13.

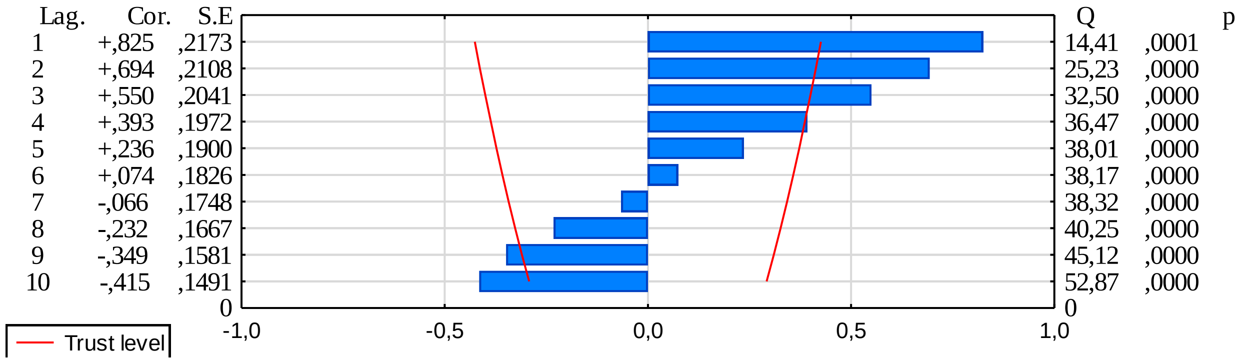


Fig.12. Auto-correlation of the sum of people killed in car accidents in selected European countries in 2000-2017 per 100,000 inhabitants.
Source: own study based on data obtained from the website:
https://ec.europa.eu/eurostat/tgm/download.do?tab=table&plugin=1&language=en&pcode=sdg_11_40

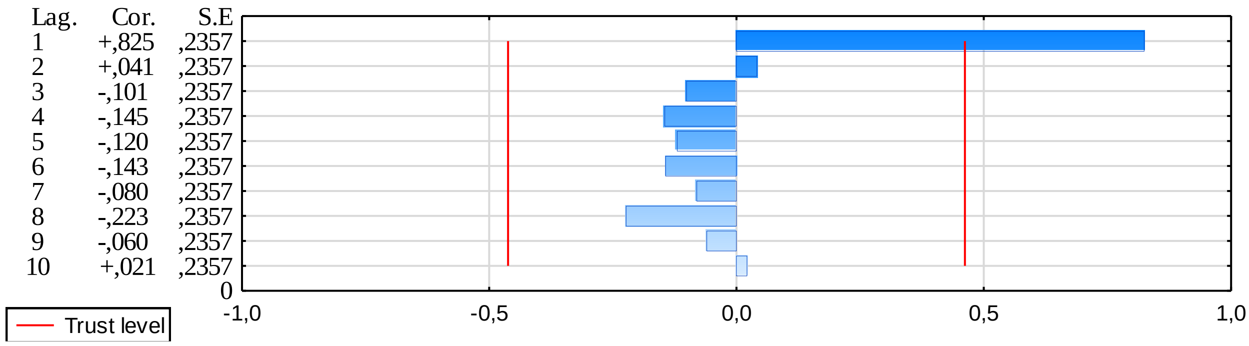


Fig. 13. Auto-correlation of the sum of people killed in car accidents in selected European countries in 2000-2017 per 100,000 inhabitants.
Source: own study based on data obtained from the website:
https://ec.europa.eu/eurostat/tgm/download.do?tab=table&plugin=1&language=en&pcode=sdg_11_40

On the other hand, it can be seen in Figures 12 and 13 that the analyzed primary time series is non-stationary. There are dependencies in it. This was confirmed by observing and assessing the autocorrelation coefficients that fall off in an oscillating manner taking the cosine shape.

P. Cyplik et al. (eds.), *Contemporary challenges in supply chains, Vol. 2*,
https://doi.org/10.17270/B.M.978-83-66017-89-4.5

The existence of dependence is also confirmed by partial autocorrelation, whose first coefficient is very strong.

The article attempts to reduce the analyzed primary time series to stationarity by performing the first degree differentiation. The results are shown in Figure 14.

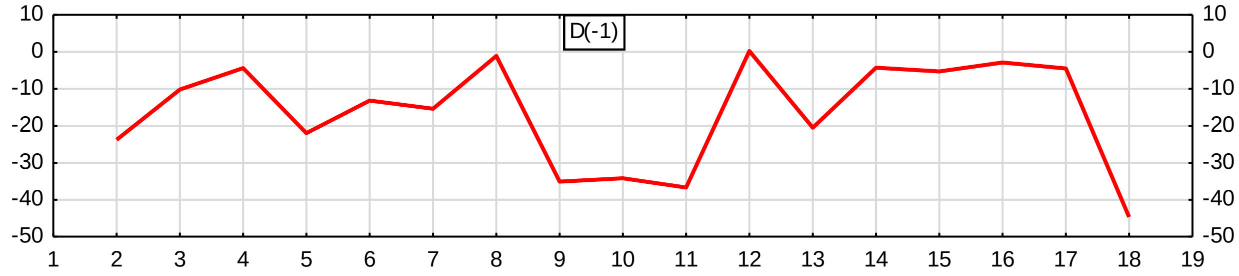


Fig.14. Differentiation of the first degree of the sum of people killed in car accidents in selected European countries in 2000-2017 per 100,000 inhabitants.
Source: own study based on data obtained from the website:
https://ec.europa.eu/eurostat/tgm/download.do?tab=table&plugin=1&language=en&pcode=sdg_11_40

Visual observation of the time series presented in Figure 14 confirms that the analyzed primary series after differentiation is stationary. For research purposes, through the use of autocorrelation in Figure 15 and partial autocorrelation in Figure 16, an attempt was made to confirm the existence of stationarity of the original time series after differentiation.

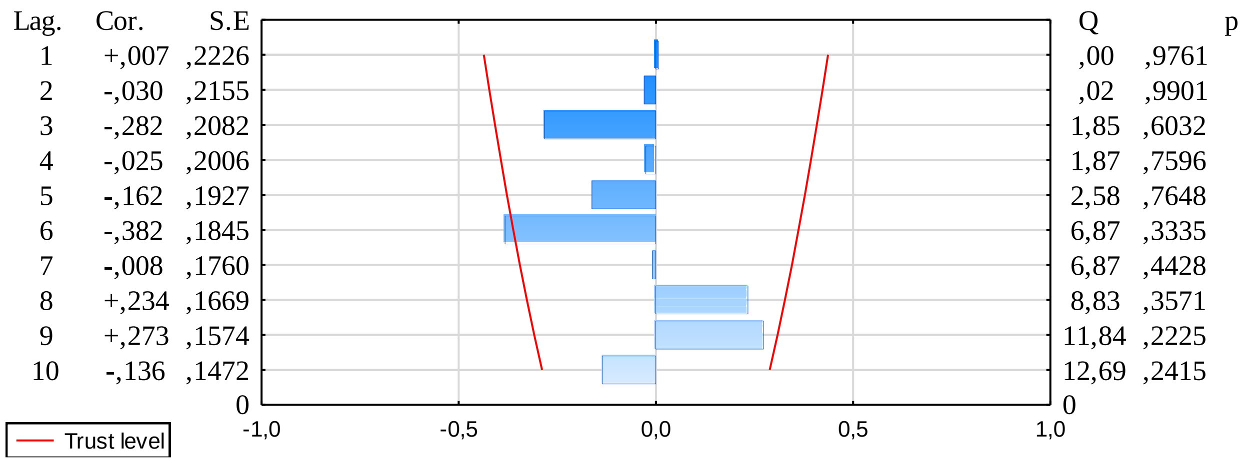


Fig.15. Time series autocorrelation after differentiation.
Source: own study based on data obtained from the website:
https://ec.europa.eu/eurostat/tgm/download.do?tab=table&plugin=1&language=en&pcode=sdg_11_40

P. Cyplik et al. (eds.), *Contemporary challenges in supply chains, Vol. 2*,
https://doi.org/10.17270/B.M.978-83-66017-89-4.5

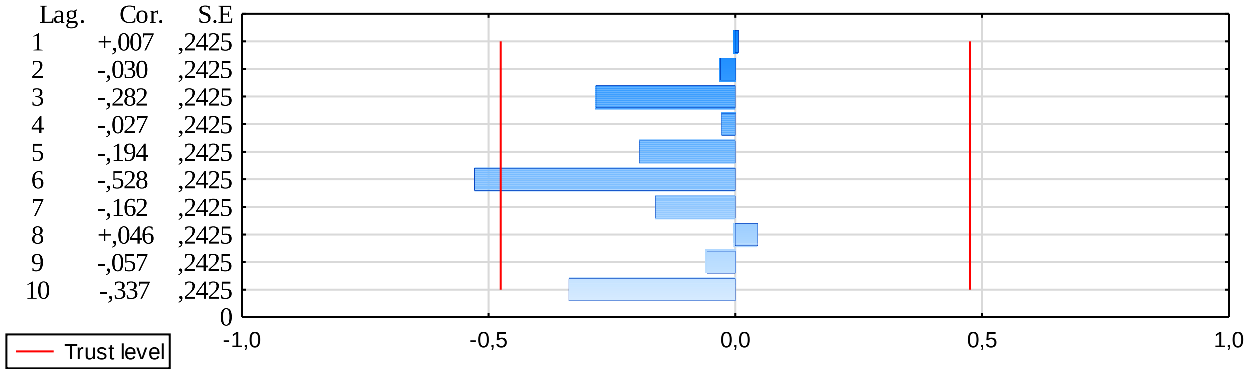


Fig.16. Partial autocorrelation of time series after differentiation.
Source: own study based on data obtained from the website:
https://ec.europa.eu/eurostat/tgm/download.do?tab=table&plugin=1&language=en&pcode=sdg_11_40

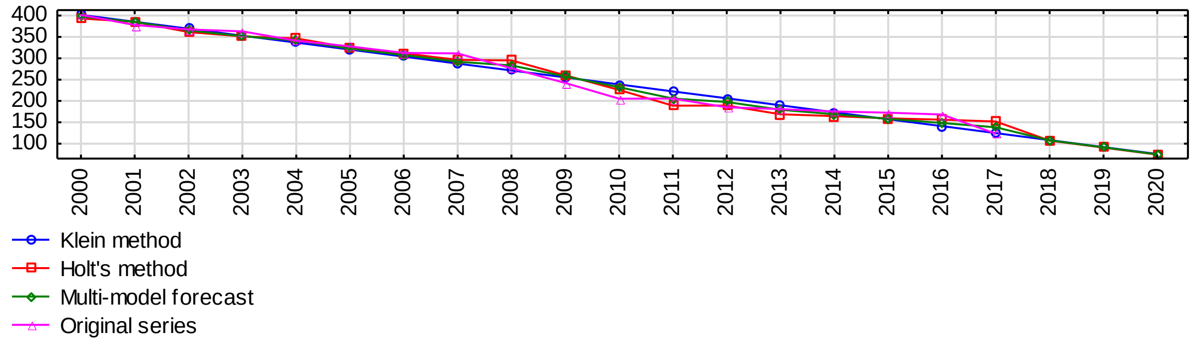
The assessment of the analysis carried out in Figures 15 and 16 is that the primary series analyzed after differentiation is stationary. This is indicated by the P and Q statistics, and the lack of significant dependencies in autocorrelation delays and partial autocorrelation.

EVALUATION OF RESEARCH

The evaluation of the analyzes related to primary time series made in the article using the various research tools is to observe a decreasing linear trend. This became a direct premise for choosing forecasting methods. The critical analysis of the literature made it possible to state that with the decreasing linear trend and the small amount of data contained in the original time series (eighteen), the best forecasting methods will be: the Holt exponential smoothing method, the Klein model and multi-model forecasting.

PREDICTION

The results of the forecast of the original time series for the years 2018-2020 by three methods: Holt exponential smoothing, Klein model and multi-model are presented in Figure 17 and Table 2.



P. Cyplik et al. (eds.), *Contemporary challenges in supply chains, Vol. 2*,
https://doi.org/10.17270/B.M.978-83-66017-89-4.5

Fig. 17. Line graph of the primary series forecast by Holt, Klein and multi-model methods together with the primary series.
Source: own study based on data obtained from the website:
https://ec.europa.eu/eurostat/tgm/download.do?tab=table&plugin=1&language=en&pcode=sdg_11_40

Table 2. Analysis of the primary series forecast for 2018-2020

Years	Original series	Klein method	Holt method	Multiple model forecast
2000	401,500	401,779	393,3206	397,55
2001	377,700	385,476	385,1412	385,308
2002	367,500	369,172	361,3412	365,257
2003	363,100	352,869	351,1412	352,005
2004	341,100	336,565	346,7412	341,653
2005	327,900	320,262	324,7412	322,502
2006	312,500	303,959	311,5412	307,75
2007	311,400	287,655	296,1412	291,898
2008	276,300	271,352	295,0412	283,196
2009	242,100	255,048	259,9412	257,495
2010	205,400	238,745	225,7412	232,243
2011	205,600	222,442	189,0412	205,741
2012	185,100	206,138	189,2412	197,69
2013	180,800	189,835	168,7412	179,288
2014	175,500	173,531	164,4412	168,986
2015	172,600	157,228	159,1412	158,185
2016	168,100	140,925	156,2412	148,583
2017	123,400	124,621	151,7412	138,181
2018		108,318	107,0412	107,679
2019		92,0144	90,6824	91,3484
2020		75,711	74,3235	75,0173

Source: own study based on data obtained from the website:
https://ec.europa.eu/eurostat/tgm/download.do?tab=table&plugin=1&language=en&pcode=sdg_11_40

Forecasts made using three different methods (Holt, Klein and Multimodel) keep the downward trend. In order to select the best one, the absolute percentage error in the forecast was examined. The results are shown in Table 3.

Table 3. Analysis of errors in forecasts for 2018-2020

	Kleine method	Holt method	Multiple model forecast
Forecast error	5,521	5,663	4,530

Source: own study based on data obtained from the website:
https://ec.europa.eu/eurostat/tgm/download.do?tab=table&plugin=1&language=en&pcode=sdg_11_40

P. Cyplik et al. (eds.), *Contemporary challenges in supply chains, Vol. 2*,
https://doi.org/10.17270/B.M.978-83-66017-89-4.5

The assessment of the analysis performed in Table 3 is that the best method for forecasting is the multi-model method.

For research purposes, the rest of the multi-model forecasting was examined. The following research tools were used for this purpose: histogram in Figure 18, normality chart with the Shapiro-Wilk test in Figure 19, autocorrelation in Figure 20 and partial autocorrelation in Figure 21.

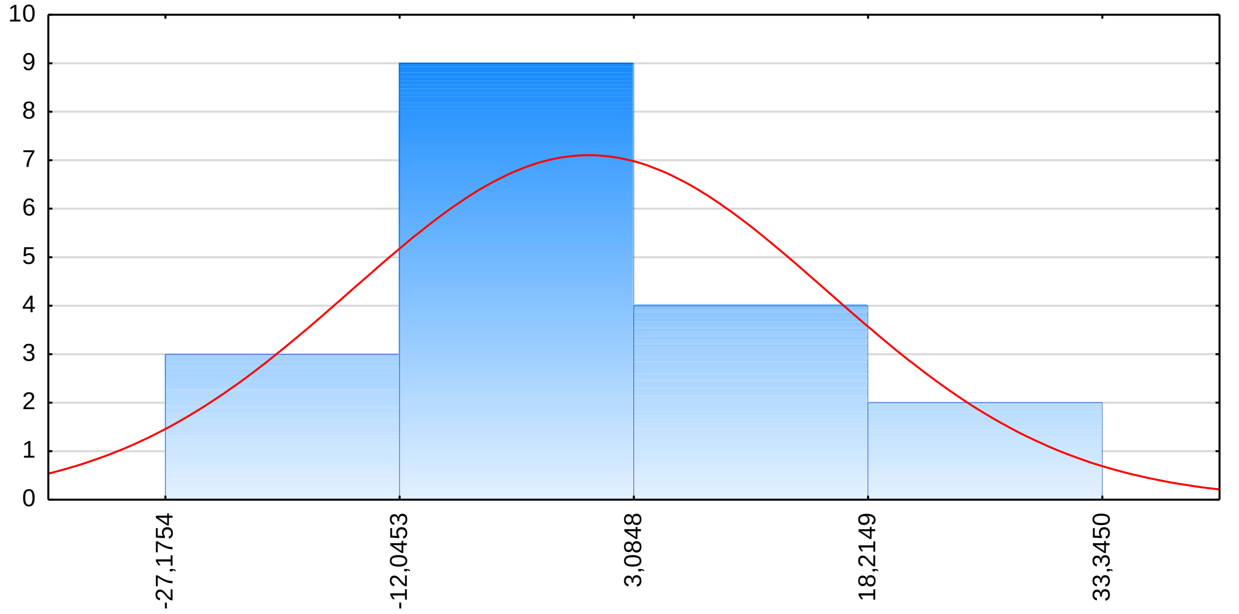


Fig. 18. Histogram of residuals in a multi-model forecast.
Source: own study based on data obtained from the website:
https://ec.europa.eu/eurostat/tgm/download.do?tab=table&plugin=1&language=en&pcode=sdg_11_40

The assessment of Figure 18 is that the distribution of the analyzed data is more slender than normal.

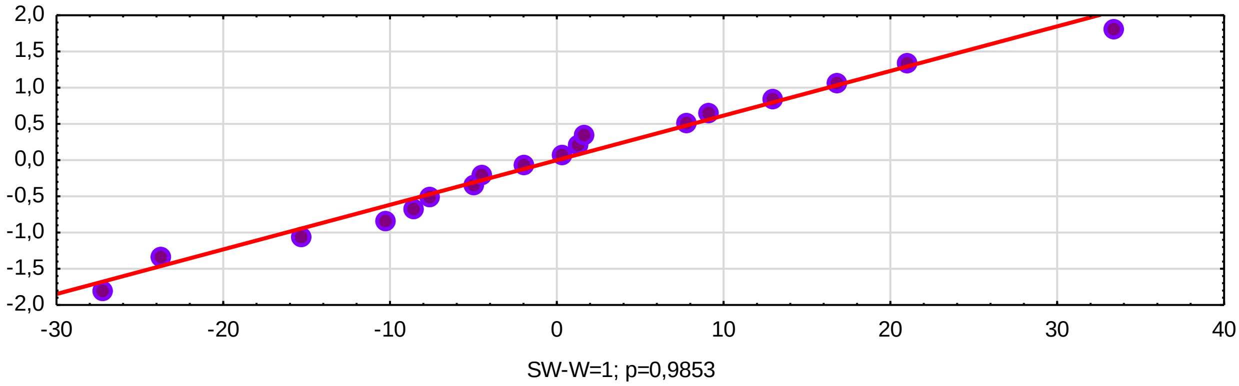


Figure 19. The normality chart together with the Shapiro-Wilk test of the multimodal forecast residuals.
Source: own study based on data obtained from the website:
https://ec.europa.eu/eurostat/tgm/download.do?tab=table&plugin=1&language=en&pcode=sdg_11_40

P. Cyplik et al. (eds.), *Contemporary challenges in supply chains, Vol. 2*,
https://doi.org/10.17270/B.M.978-83-66017-89-4.5

The analyzed multi-model prediction residuals indicate that their distribution is typically normal (histogram - Figure 18 and normality chart - Figure 19). This is also confirmed by the Shapiro-Wilk test.

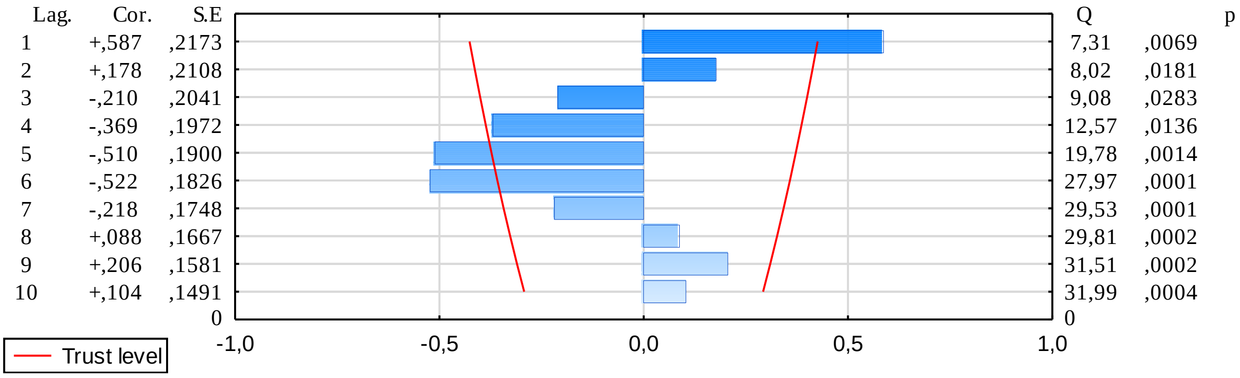


Fig. 20. Autocorrelation of multi-model residual forecasts.
Source: own study based on data obtained from the website:
https://ec.europa.eu/eurostat/tgm/download.do?tab=table&plugin=1&language=en&pcode=sdg_11_40

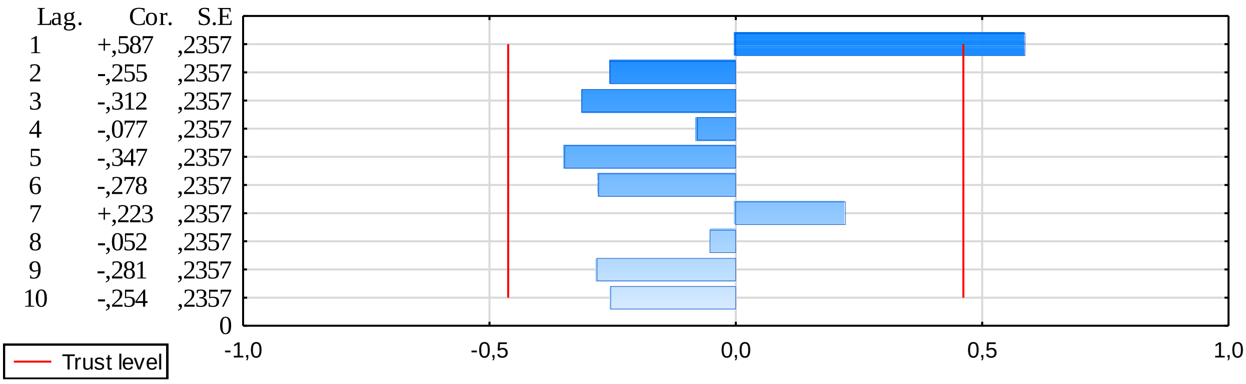


Fig. 21. Partial autocorrelation of residues of a multi-model forecast.
Source: own study based on data obtained from the website:
https://ec.europa.eu/eurostat/tgm/download.do?tab=table&plugin=1&language=en&pcode=sdg_11_40

In the remainder of the multi-model forecast, there is a relationship in Figures 20 and 21.

CONCLUSIVE REMARKS

In Europe, there is a visible decline in deaths per 100,000 inhabitants in road accidents in dynamic terms. This is due to increased expenditure on improving road infrastructure and other aspects that directly overlap this phenomenon.

The main goal of the article has been achieved. The obtained results of the assessment of conducted analyzes became a premise for the selection of forecasting methods. The original time series were forecast for 2018-2020. The obtained results of the forecast were analyzed and evaluated. The best method for forecasting was multi-model. The results of the multi-model

P. Cyplik et al. (eds.), *Contemporary challenges in supply chains, Vol. 2*,
https://doi.org/10.17270/B.M.978-83-66017-89-4.5

forecast are as follows: 2018 - 107.679, 2019 - 91.3484, 2020 - 75.0173. They confirm the further strong downward trend of the linear trend.

The use of the multi-model method for forecasting, with the observed dependencies in the considered primary time series, can improve planning in terms of predicting the number of people killed in road accidents, so as to minimize their occurrence by implementing appropriate investment outlays to improve safety in this matter.

REFERENCES

Dittmann P. et al., 2016, Forecasting in sales and financial management of an enterprise, Non-obvious Publisher, imprint GAB Media.

Gołembska E., 2008, Air transport in the development of Wielkopolska in the 21st century, Scientific Notebooks of the University of Szczecin, No. 491, Economical Problems of Services, No. 15, p. 304–312.

Jacyna M., Lewczuk K., 2016, Logistic systems design, PWN Scientific Publishing House, Warsaw.

WEBSITES

https://ec.europa.eu/eurostat/tgm/download.do?tab=table&plugin=1&language=en&pcode=sdg_11_4, (as of May 22, 2019).

http://:www.esranet.eu, (as of May 22, 2019).

P. Cyplik et al. (eds.), *Contemporary challenges in supply chains, Vol. 2*,
https://doi.org/10.17270/B.M.978-83-66017-89-4.5