INFORMATION
SYSTEMS IN
MANAGEMENT

Information Systems in Management (2018) Vol. 7 (3) 155−170
DOI: 10.22630/ISIM.2018.7.3.1

THE APPLICATION OF THE ARIMA MODEL IN FORECASTING THE PASSENGER TRAFFIC ON THE EXAMPLE OF BORDER CROSSINGS BETWEEN THE SUBCARPATHIAN PROVINCE AND UKRAINE

DOROTA DEJNIAK

Institute of Technical Engineering,
State Higher School of Technology and Economics in Jaroslaw, Poland

Accession of Poland to the European Union meant that its eastern border became the external frontier of the Community. The next step in the European integration was joining the Schengen Zone by Poland. Polish citizens may freely travel throughout the Schengen Zone and the state was obliged to tighten its eastern border. Under these circumstances conducting research on passenger traffic has become a vital issue, with particular focus on the eastern frontier. In the article an attempt is made at examining the possibility of forecasting passenger traffic on the example of border crossing points between the Subcarpathian Province and Ukraine using the ARIMA models. Confirmation of these possibilities seems to be crucial as the number of people crossing the border is characterized by high variability and sensitivity to the political situation. The study is based on the information provided by the Polish Border Guard. The conducted time series analysis is of a multi-purpose character. It may be used to support decision making processes of investment, organizational, as well as socio-political nature.

Key words: passenger border traffic, ARIMA models, time series analysis

1. Introduction

Accession of Poland to the European Union meant that its eastern border became the external frontier of the Community. The next step in the European integration was joining the Schengen Zone by Poland. Polish citizens may freely travel throughout the Schengen Zone and the state was obliged to tighten its eastern border. Under these circumstances conducting research on passenger traffic has become a vital issue, with particular focus on the eastern frontier. The data and reports prepared by Central Statistical Office (GUS) provide information dealing with the volume of passenger traffic, the purpose of their visit, and the structure of expenses paid by foreigners. Due to a high sensitivity of the data, modelling of the phenomenon in question poses some difficulties. Acquiring the right database is problematic, as the examples of border traffic modelling available refer mainly to air passenger traffic [8, 10, 11, 14].

The aim of the analysis presented in the article is to examine the possibility of forecasting passenger traffic on the example of border crossing points between the Subcarpathian Province and Ukraine using the ARIMA models. Confirmation of these possibilities seems to be crucial as the number of people crossing the border is characterized by high variability and sensitivity to the political situation. The Subcarpathian Province is the area of Poland where the majority of incoming foreigners, especially Ukrainian citizens, make their migration decisions. The study is based on the information provided by the Polish Border Guard. For the purposes of the analysis the monthly data on passenger traffic from January 2012 to December 2016 were used pertaining to the number of clearances performed by the Border Guard. The data refer to the number of crossings by Ukrainian citizens, not the total number of individuals crossing the frontier. The conducted time series analysis is of a multi-purpose character. It may be used to support decision making processes of investment, organizational, as well as socio-political nature.

2. Passenger border traffic at border crossings between the Subcarpathian Province and Ukraine

The state border between Poland and Ukraine with a total length of 535 kilometers (15.2% of the total length of the Polish border) is at the same time the external frontier of the European Union. Its length across the Subcarpathian Province and the Lublin Province amounts to 239 kilometers and 296 kilometers respectively.

In the Subcarpathian Province, the Bieszczady Border Guard Department covers the crossing points with Ukraine in Medyka, Korczowa, Krościenko, Budomierz, Werchrata (a railway crossing point, freight traffic only) and Jasionka (an air crossing point).

Map 1.. Border crossing points in the Subcarpathian Province [17]
Source: own preparation on the basis of [17]

Polish accession to the Schengen Zone resulted in the opening of borders for Poles inside the European Union, but also made it more difficult for citizens of Ukraine to cross the border (the need for having a visa). In order to facilitate the border traffic an agreement between the Government of the Republic of Poland and the Cabinet of Ministers of Ukraine on local border traffic was signed in March 2008, and entered into force on 1st July 2009. It was then modified by the Act of 25th June 2015 on the ratification of the Second Protocol between the Government of the Republic of Poland and the Cabinet of Ministers of Ukraine on local border traffic regime, signed on 17th December 2014 in Warsaw amending the first Agreement.

The provisions of the agreement cover citizens of both countries living in the border area stretching not further than 30 kilometers from the common border. However, if some part of an administrative district lies between 30 and 50 kilometers from the border line, it is considered part of the border area. According to the concluded agreement, the Polish border area in the Subcarpathian Province comprises 588 locations in 43 municipalities and 8 districts. The border area includes, inter alia, Przemyśl, Jarosław, Sanok, Zagórz, Lesko, Lubaczów. The municipalities on the Polish side cover an area of 13,400 square kilometers, inhabited by 800,000 people. Ukrainian border area comprises 1,111 locations in 23 regions covering a total area of 24,000 square kilometers, inhabited by 1,200, 000 people. Local Border Traffic is the regular crossing of an external land border by residents in order to stay in border area for example for social, cultural or substantiated economic reasons, or for family reasons, for a period not exceeding the time limit laid down in the Regulation (EU) No 1342/2011of the European Parliament.

In the last several years at the Polish-Ukrainian border, there has been a smaller increase of Poles going through clearance compared to the number of cleared foreigners. Price relations between both countries and recently an unsettled international situation in Ukraine are undoubtedly important factors influencing the situation. The analyses of the data from the first half-year of 2017, and presented by the Border Guard show that it is going to be a record-breaking year in terms of the number of passengers cleared at the border crossing points within the authority of the Bieszczady Border Guard Department (BBGD). Citizens of Ukraine constitute 85% of people crossing the border. The number of Ukrainians crossing the border on the basis of the Local Border Traffic (LBT) regime has been decreasing significantly in favor of visa-free regime. This is a consequence of the abolition of visa requirements for Ukrainians who use biometric passports since 11th June 2017. Currently, approximately 30% of Ukrainians cross the border without a visa. In the first half-year of 2017, the Subcarpathian border was crossed by more than 860,000 Poles, which constitutes an increase of approximately 1%. On the other hand, in the first quarter of 2017 there was a 2.4% increase (compared to the analogous data from 2016) of foreigners crossing the border into Poland from Ukraine and a 4.9% decrease of Polish citizens crossing the border into Ukraine [13].

High dynamics, variability and spatial distribution of cross-border traffic is conditioned by a number of factors, the most important of them being geopolitical situation, development of road infrastructure and economic situation for the cross-border trade.

The formal agreements on the local border traffic and recent events in Ukraine have determined the size and directions of passenger traffic. As a result an increase in border traffic, particularly the LBT, has been observed. It may be linked to deteriorating economic situation in Ukraine resulting in the search for additional sources of income, mostly in small-scale trade taking advantage of price differences, and the possibility of employment in Poland. Changes in the level of traffic concentration can be observed over a longer span of time. Initially, a tendency towards its de-concentration dominated, which was associated with opening of new border crossing points and their low capacity (choosing remote border points). In the recent years the situation has changed considerably. Concentration of passenger streams, mostly along the main routes can be observed. In this case, an important factor is the development of Polish network of motorways and express ways [2]. Poles travelling to Ukraine in order to purchase excise goods are mainly residents of the immediate border area, choosing the closest check points, taking advantage of the LBT. Undoubtedly, queues and long border wait times have significantly contributed to the decline in the number of Poles crossing the Ukrainian border [16]. Therefore, it is not surprising that there is an interest in the train connection between Przemyśl and Kiev and that plans are being made to launch a Kraków – Lviv connection.

In order to learn more about the structure of expenses paid by foreigners crossing the border between Poland and Ukraine, the Polish Central Statistical Office (GUS) conducted a survey, whose results are presented in the publication entitled *Ruch graniczny oraz wydatki cudzoziemców w Polsce i Polaków za granicą w 2015 roku* [13]. According to the survey, 10,700,000 crossings of the Polish – Ukrainian border within the local border traffic (27.6% more than in 2014) were reported. 56.6% of crossings at this section of the border were made by foreigners (in 2014 – 53.6%). Within the LBT most foreigners (75.3%) crossed the border several times a week; those who crossed the border several times a month accounted for 18.3%, every day – 5.3%, several times quarterly – 0.9%, a few times a year or less frequently – 0.1%, respectively. The estimated value of expenses paid in Poland by foreigners crossing the Polish – Ukrainian border within the LBT in 2015 amounted to PLN 2.9 billion (23.5% more than in 2014), which constituted 44.6% of their total spending (41.3% in 2014). For all foreigners crossing the border the majority of expenses (87.8%) was spent on non-food goods, similarly to the total number of foreigners crossing the Polish – Ukrainian border (86.5%). There were only minor differences in the structure of expenses of both groups of foreigners. In 2015 the average spending in Poland by a foreigner crossing the Polish – Ukrainian border within the LBT amounted to PLN 540 (3% less than in 2014) and was 22.1% lower when compared to the average spending of all foreigners crossing this border (including the LBT). The analysis shows that the implementation of regulations facilitating border crossing has significantly influenced the revival of traffic in the border region. It has also resulted in the increasing number of commercial partnerships involving foreign capital. By the end of 2015, in the area close to the border with Ukraine there has been a 124.3% increase in the number of those partnerships since 2009.

On the above described factual conditions, the problem of modelling passenger border traffic will be presented using statistical methods of analysis of time series by means of ARIMA models.

3. Model construction

ARIMA (Autoregressive Integrated Moving Average) models are a very general class of time series models. Their construction is based on the phenomenon of autocorrelation. These models form the basis for the forecasting method known as Box–Jenkins [1] method. They can be used to model stationary time series and those non-stationary time series which can be transformed into stationary ones. There are three basic models of this class: autoregression models (AR), moving average models (MA) and mixed autoregression and moving average models (ARMA). The symbol I used in the model name indicates that a time series was subject to differencing. In order to formulate an ARIMA model (p, d, q) a notation stipulating the row of individual model components is used: autocorrelation – p, differencing – d, moving average – q. The process of model construction consists of parts relating to: identification, estimation, diagnostic checking [1, 3].

In the first stage, the initial identification of a time series in terms of stationarity is made through checking the function of autocorrelation (ACF) and partial autocorrelation (PACF). The fact that in stationary processes the autocorrelation function decreases (as a rule, quite rapidly) is used to determine whether the process is stationary or non-stationary. If the decrease of the autocorrelation function is very slow, it means that the time series is non-stationary and should be reduced to the stationary form using differencing. During the second phase the parameters of given models are estimated. The identification phase provides information on the possible variants of the process. The final choice is based on an analysis of several criteria: relevance/validity of (the) model parameters, mean squared error, information criterion. Then the model is subject to the analysis of properties of the model residuals. If the model residuals are a white noise process and there are no significant values of functions ACF or PACF of the model residuals, the model can be used in forecasting. Otherwise, another model should be chosen or the model should be identified again. Having estimated the model parameters and their statistical significance checked, an assessment of model fit should be made [3, 9]. From the goodness of fit tests based on the analysis of correlation of residuals, the Q test should be used:

$$Q = n(n+2)\sum_{\tau=1}^{M}(n-\tau)^{-1} r_{\tau}^{2}(a) \tag{1}$$

where $r_\tau(a)$ is the function of autocorrelation of residuals, and M is the maximum delay of this function. The Q statistics has the distribution of M-p-q degrees of freedom. For testing the significance of parameters as well as for building confidence intervals for the forecasts, it is important that a white noise has normal distribution.

Finally, the model should be used to make a forecast. The basic difficulty in the use of ARIMA models is the fact that there is no way to automate the procedure for their construction.

The article provides an analysis of the time series describing the volume of passenger border traffic on the example of the Subcarpathian crossing points with Ukraine. As the citizens of Ukraine constitute more than 80% of all foreigners crossing the eastern border, the variable showing monthly data from January 2012 to December 2016 on crossing the border by the citizens of Ukraine into Poland was selected to construct the model. The data was obtained from the Bieszczady Border Guard Department. Modeling was preceded by an analysis of the data pertaining to the number of border crossings by Polish citizens from and into Poland, and Ukrainian citizens from Poland and into Poland. Statistica program was used in order to construct the model.

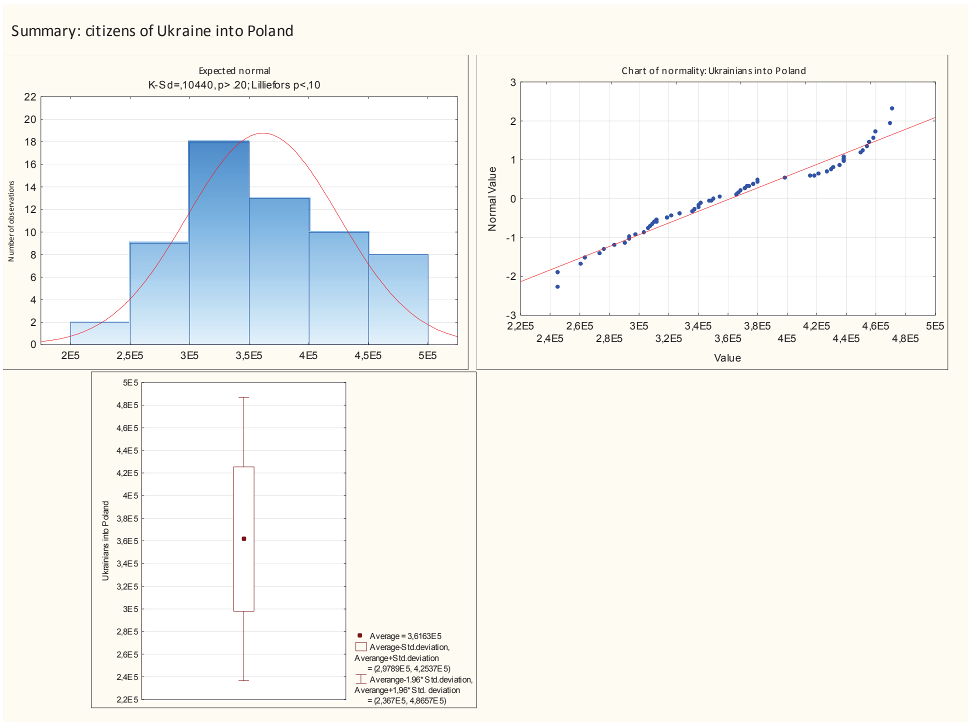

Figure 1. Basic statistics for the variable describing monthly number of border crossings by the citizens of Ukraine entering Poland (own work)

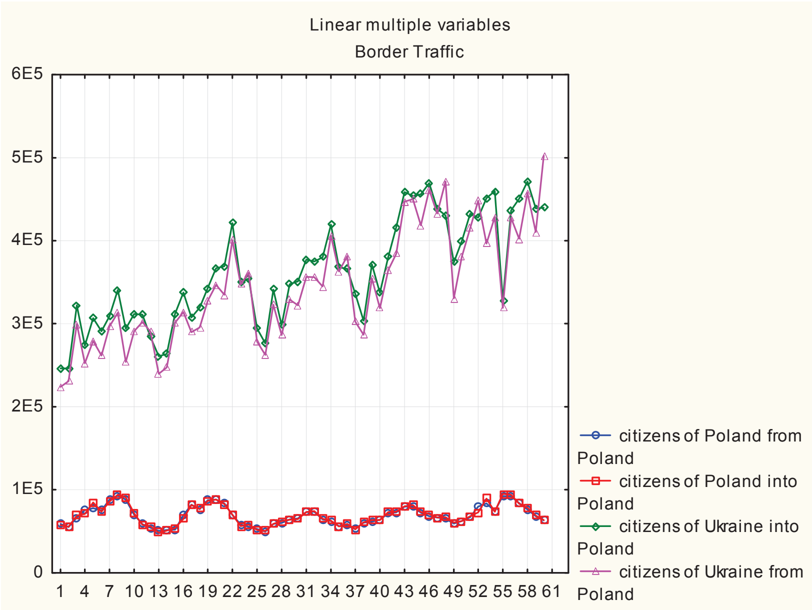

Figure 2. Line chart of variables describing the passenger traffic at the Subcarpathian part of the border between Poland and Ukraine on a monthly basis (own elaboration)

After the initial analysis of relationship between the variables, a conclusion can be drawn that there is no relationship between the amount of Polish and Ukrainian citizens entering Poland and those leaving Poland. It confirms the opinion about the influence of political situation and prices of goods, including excise goods, on the volume of passenger border traffic.

Table 1. Correlations between variables relating to passenger traffic at the Polish – Ukrainian border (own work)

Variable	Correlations (Border Traffic Copy) Marked correlation coefficients are significant with p < .05000 N = 60			
	Poles from Poland	Poles into Poland	Ukrainians into Poland	Ukrainians from Poland
Poles from Poland	1.000000	0.984250	0.324401	0.239433
Poles into Poland	0.984250	1.000000	0.344845	0.255672
Ukrainians into Poland	0.324401	0.344845	1.000000	0.961961
Ukrainians from Poland	0.239433	0.255672	0.961961	1.000000

To construct the ARIMA model we have used the variable describing passenger border traffic at the Subcarpathian part of the border between Poland and Ukraine referring to the number of crossings by Ukrainian citizens entering Poland. The chart created suggests that it is a non-stationary time series which should be submitted to differencing. The non-stationarity of the series is confirmed by the autocorrelation function. For the assumed significance level α = 0.05, the autocorrelation coefficients for delays of 1 to 14 months appeared to be statistically different from zero. This confirms the earlier assumption that the time series of the variable in question is non-stationary.

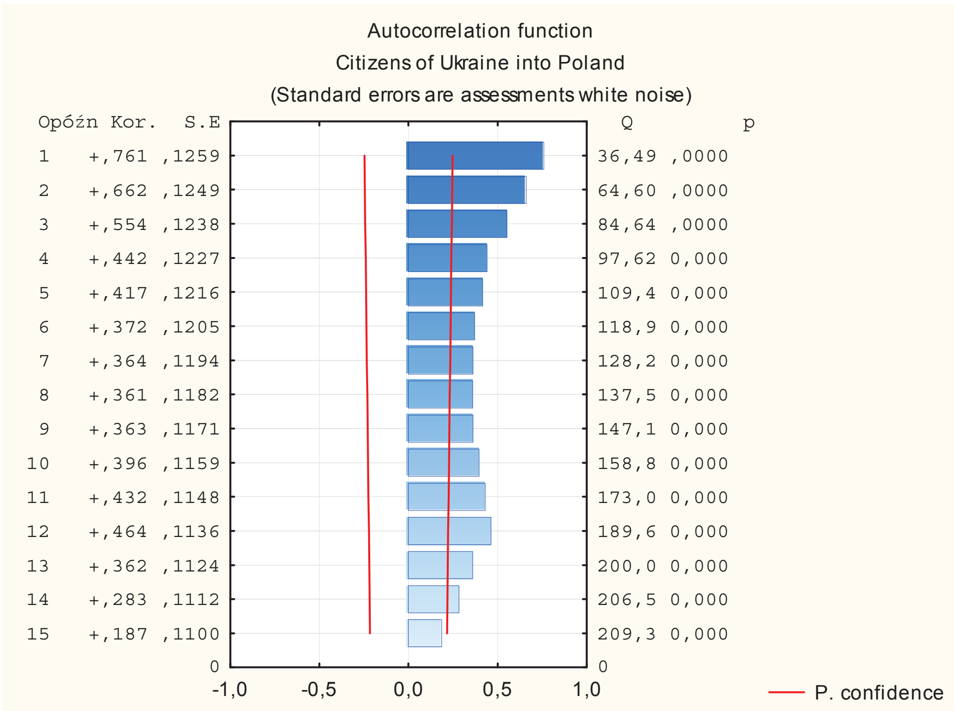

Figure 3. Coefficients of autocorrelation of the time series of border crossings by citizens of Ukraine into Poland

The image of PACF function is not entirely clear. At the assumed significance level α = 0.05, the values of the PACF function for the delays 1 and 13 are significant.

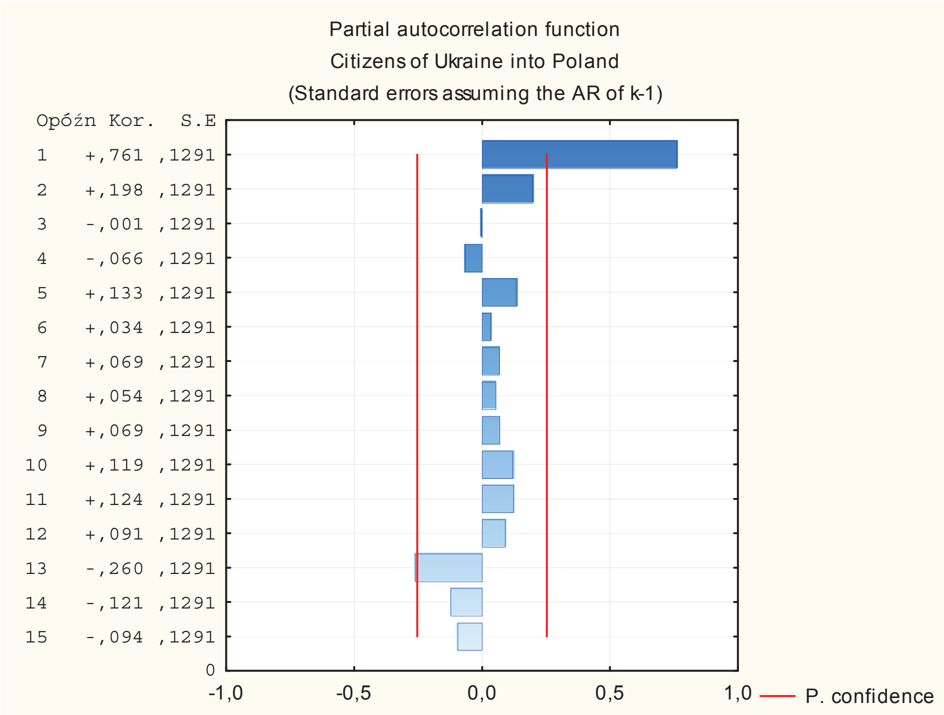

Figure 4. Coefficients of partial autocorrelation of the time series of border crossings by citizens of Ukraine into Poland

The given time series has been described using different ARIMA models. The following criteria were taken into account when choosing the best model: significance of model parameters, lowest mean squared error value and information criterion. On the basis of the conducted analysis the ARIMA (1,1,2) model was chosen.

Table 2. Summary, parameter assessment of the ARIMA(1,1,2) model

Param eter	Data: Ukrainian citizens into Poland (Border traffic 26-07-17) Transformations: D(1) Model:(1,1,2) Residual MS = 1474E6					
	Parameter	Asymptote Std. error	Asymptote t (56)	p	Lower limit of 95% p.confidence	Upper limit of 95% p.confidence
p(1)	–0.980966	0.065486	–14.9799	0.000000	–1.11215	–0.849783
q(1)	–0.547048	0.133989	–4.0828	0.000143	–0.81546	–0.278637
q(2)	0.452772	0.138181	3.2766	0.001807	0.17596	0.729582

Having established the significance of assessments, we proceeded to the analysis of residuals. The chart of residual autocorrelation function shows that they are insignificantly differentiating from zero. Combined Q criterion for the maximum delay is 23.51 with the test probability p = 0.0740 greater from the standard significance level α = 0.05, which means that the selected model is correct.

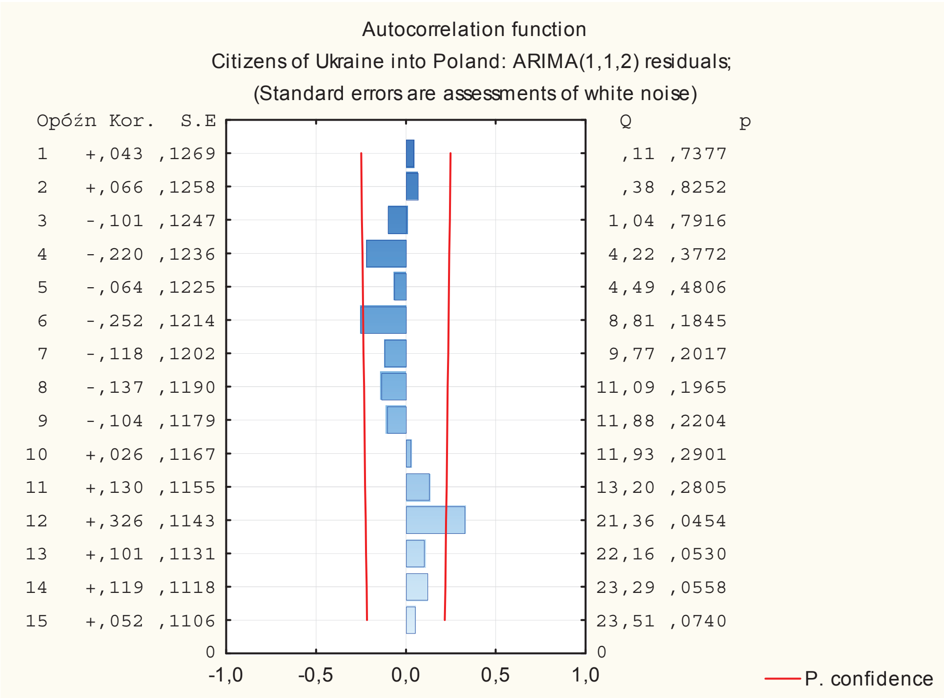

Figure 5. Chart of ACF of residuals of the ARIMA(1,1,2) model

Partial autocorrelation chart confirms that residuals constitute the white noise due to the lack of significant correlations of autoregressive type.

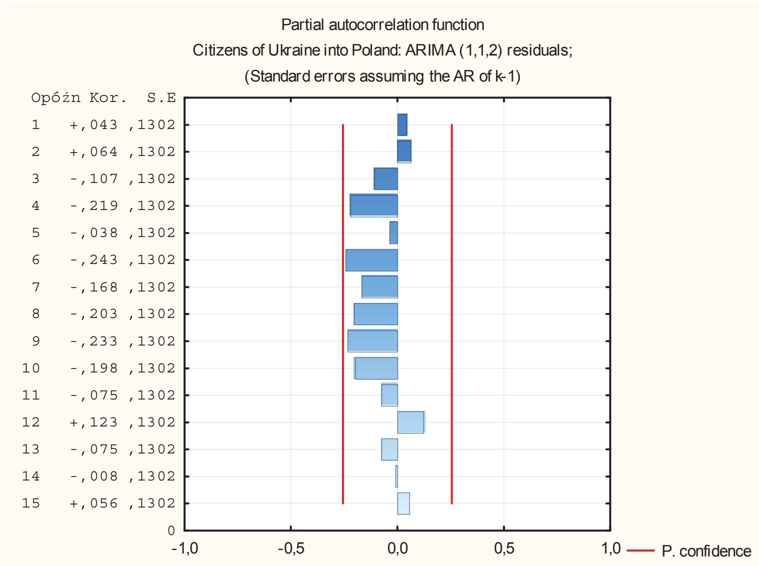

Figure 6. Chart of PACF of residuals of the ARIMA(1,1,2) model

Examination of the distribution of residuals confirmed approximately the shape of normal distribution. Normal chart of residuals shows shaping of quantiles against a straight line which corresponds to the quantiles of normal distribution. If the points distribute on the line or close to it, it can be considered that the residuals have normal distribution.

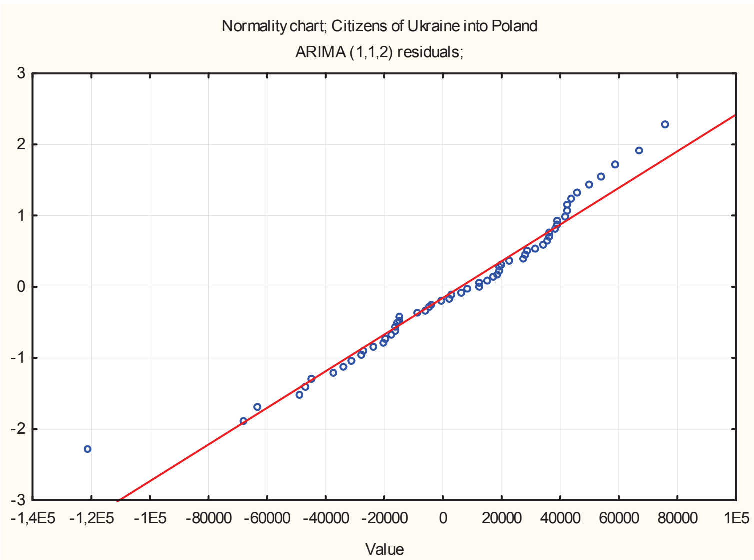

Figure 7. Chart of normality of residuals of the ARIMA(1,1,2) model

Having obtained positive results of the diagnostic check, the designated model was recognized to be useful for developing the forecast. The model was used to forecast the volume of passenger border traffic for Ukrainian citizens entering Poland in subsequent three months. They were compared with the actual data received from the BBGD. For one-step ahead forecast the error does not exceed 9%. The results of the calculations are presented in tables 3 and 4. The graphical chart of the forecast is presented in Figure 8.

Table 3. Forecast determined for the ARIMA(1,1,2) model

Obs. no	Forecast; Model: (1,1,2) Seasonal delay: 12 (Border traffic 26-07-17) Data: Ukraine citizens into Poland Database beginning: 1 Database end: 60			
	Forecast	Lower 90.0000%	Upper 90.0000%	Std. error
61	444662.2	380447.4	508877.0	38393.98
62	441327.1	367537.4	515116.9	44118.83
63	444598.7	363095.6	526101.9	48730.67

Table 4. Forecast determined for the ARIMA(1,1,2) model compared with the observed data

No of the next month	Forecast; Model: (1,1,2) Seasonal delay: 12 (Border traffic 26-07-17) Data: Ukraine citizens into Poland Database beginning:1 Database end: 60		
	Forecast	Observed data	Relative error
61	444662.2	435727	–2.05%
62	441327.1	406345	–8.60%
63	444598.7	435028	–2.20%

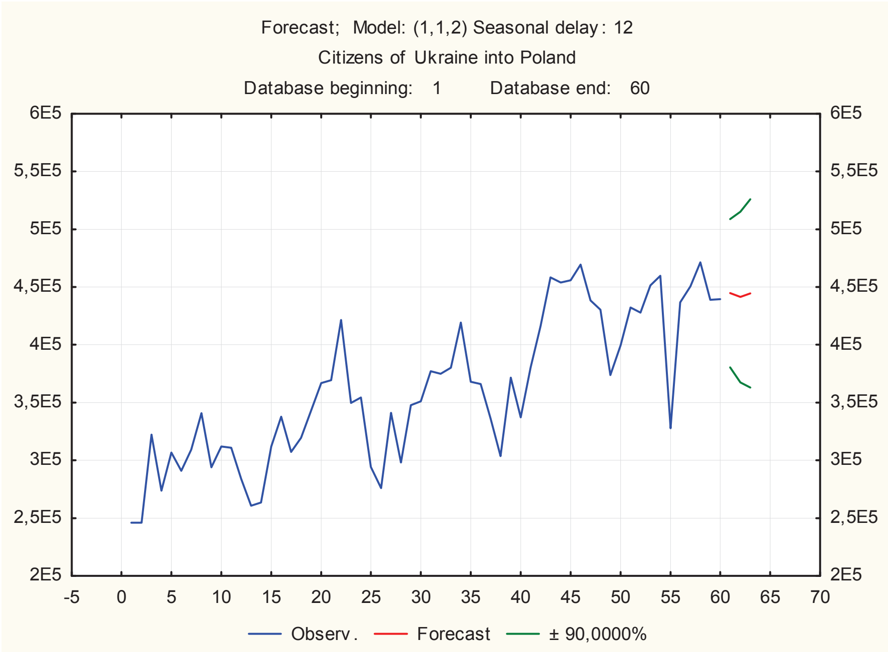

Figure 8. Forecast for three months of Q1 of 2017, using the ARIMA(1,1,2) model

4. Conclusion

An important decision-making tool can be obtained by making use of ARIMA models. The advantage of using this type of models is that, despite many problems connected with their construction and testing, we obtain information about the time series structure and mechanism of its creation. Creating models for passenger border traffic is very interesting due to its complex structure and huge sensitivity to social and economic conditions. Construction of an ARIMA model requires using over 50 observations, therefore a time step longer than a month could generate incorrect forecasts and conclusions. There is a direct, positive correlation between the volume of passenger border traffic and the expenses paid by people crossing the border, therefore the obtained forecasts may be of business nature. They could be used to support making decisions connected with communication infrastructure and commercial infrastructure. The universal character of the study allows for the use of ARIMA models in relation to research on border traffic at any section or border crossing point. The forecasts obtained may be used to analyze regional migration.

REFERENCES

[1] Box E.P., Jenkins G.M. (1983) *Szeregi czasowe – analiza i prognozowanie*, Wydawnictwo Naukowe PWN, Warszawa, Poland.

[2] Komornicki T. (2016) *Zmiany w ruchu przez polską granicę wschodnią w roku 2014 na tle sytuacji geopolitycznej* Prace i Studia Geograficzne T.61, 103 – 119, Warszawa.

[3] Kot T., Jakubowski J., Sokołowski A. (2007) *Statystyka*, Wydawnictwo Difin, Warszawa, Poland.

[4] Luszniewicz A., Słaby T. (2001) *Statystyka z pakietem komputerowym Statistica PL* Teoria i zastosowania, Wydawnictwo C. H. Beck, Warszawa, Poland.

[5] Maddala G.S (2006) *Ekonometri*a, Wydawnictwo Naukowe PWN, Warszawa.

[6] Maksimczuk A., Sidorowicz L.(2007) Graniczny ruch osobowy i towarowy w Unii Europejskiej, Almamer, Warszawa, Poland.

[7] Maksimczuk A., Sidorowicz L. (2007) *Ochrona granic i obsługa ruchu granicznego*, Wydawnictwo Prawnicze LexisNexis, Warszawa, Poland.

[8] Rudnicki M. (2012) *Zastosowanie modelu ARIMA do krótkookresowego prognozowania wielkości ruchu pasażerskiego na przykładzie Międzynarodowego Portu Lotniczego w Krakowie – Balicach*, Zeszyty Naukowe Uniwersytetu Ekonomicznego w Krakowie T.885, 175 -191, Kraków, Poland.

[9] Sokołowski A. (2010) *Analiza szeregów czasowych i prognozowanie*, Statsoft, Kraków.

[10] Tang X., Deng G. (2016) Prediction of Civil Aviation Passenger Transportation *Based on ARIMA Model*, Open Journal of Statistic, 6, 824 – 834 http://dx.doi.org/10.4236/ojs.2016.65068.

[11] Tsui Wai Hong Kan, Hatice Ozer Balli, Hamish Gow (2011), *Forecasting airport passenger traffic: the case Hong Kong International Airport*, Aviation Education and Research Proceedings, vol 2011, p 54-62.

[12] Zeliaś A., Pawełek B., Wanat S. (2004) *Prognozowanie ekonomiczne. Teoria, przykłady, zadania,* Wydawnictwo Naukowe PWN, Warszawa, Poland.

[13] Główny Urząd Statystyczny (2016) Ruch graniczny oraz wydatki cudzoziemców w Polsce i Polaków za granicą w 2015 r., Warszawa – Rzeszów, Poland.

[14] Zhao X., Zhao H. (2015) Empirical Analysis and Predictions of Civil Aviation Passenger Traffic Based on ARIMA model, 3rd Conference on Machinery, Materials and Information technology Applications (ICMMITA 2015)

https://www.atlantis-press.com/proceedings/icmmita-15/25841854

[15] https://stat.gov.pl/obszary-tematyczne/kultura-turystyka-sport/turystyka/osobowy-ruch-graniczny-z-ukraina-bialorusia-i-rosja-wrzesien-2017-r-,7,37.html [3.11.2017]

[16] http://www.bieszczadzki.strazgraniczna.pl/bie/aktualnosci/25368,12-milionow-podroznych-odprawionych-w-przejsciach-granicznych-Bieszczadzkiego-OS.html [30.11.2017]

[17] http://www.bieszczadzki.strazgraniczna.pl/bie/dla-podroznych/14923,Gdzie-mozna-przekroczyc-granice-na-Podkarpaciu.html [3.03.2018]