Network Intelligence Studies
Volume IV, Issue 1 (7) / 2016

Kamer Ainur M. AIVAZ
Ion Danut I. JUGANARU
Mariana C. JUGANARU
Ovidius University of Constanta

THE ANTICIPATION OF THE NUMBER OF TOURISTS ARRIVED IN MAMAIA USING THE TYPE OF MODELS ARIMA

Empirical study

Keywords
Tourists arrivals,
Auto regressive models,
Prevision

JEL Classification
C10, C21, C53, M21, J63, Y32, Z33

Abstract

The Mamaia station is, at the moment, the biggest and the most looked for touristic station from the Romanian seaside of the Black Sea. From the analysis of the evolution of the main indicators of touristic circulation from the last 10 years (2006-2015), we can notice a significant increase, but we are also interested in knowing the tendency of their modification in the near future. For this reason, in the present study, we wanted to test the contribution of the models ARIMA to the elaboration of an anticipation regarding the indicators: arrivals of tourists, totally and structurally: Romanians and foreigners, for Mamaia station. We consider that the results obtained in this study may contribute to the defining of the strategy of development of the station and ensuring the necessary conditions for hosting a significant greater number of tourists, in the following years.

INTRODUCTION

From the previous study, made by us, that had as main aim the identification of resemblances and differences between 17 localities and touristic stations of the Romanian seaside (Năvodari, Mamaia Sat, Mamaia, Constanţa, Eforie Nord, Techirghiol, Eforie Sud, Costineşti, Neptun, Olimp, Cap Aurora, Jupiter, Venus, Saturn, Mangalia, 2 Mai şi Vama Veche), through the use of the ACP method (analysis of the main components) and a data base of the main indicators of touristic circulation in 2015, we noticed that 3 cities, respectively touristic stations Constanta, Mamaia and Eforie Nord emphasize themselves, obviously, through the biggest values recorded by the indicator "tourists' arrival (total, Romanian and foreigners)" and "tourists that stay overnight (total, Romanians and foreigners). The interpretations of these results leads to the conclusion that the 3 cities/touristic stations have been the most preferred by the Romanian and foreign tourists, in 2015.

We selected, for the present study, Mamaia station, considered the "Pearl of the Black Sea", the most well-known station of the Romanian seaside. The database used for the mathematic modeling, for the elaboration of anticipation, is taken from the National Institute of Statistics – INS and includes indicators of touristic circulation for 2006-2015 for the touristic station Mamaia (see table 2.1). We mention that "tourists arrivals total/ Romanians/ foreigners", tourists that stayed overnight total/ Romanians/ foreigners and the medium stay of tourists total/ Romanians/ foreigners represent some of the most used indicators, specific to the touristic circulation, for the internal and international tourism (Minciu, 2005; Jugănaru, 2007). These are physical indicators, that express only the quantitative dimension of the touristic demand. The weak point of these indicators is that they don't reflect the total touristic demand, the real one, because they refer only to the tourists accommodated in the structures of touristic reception having classified functions of accommodation. Moreover, these structures of touristic reception are re-established in the physical indicators " accommodation capacity", that express the quantitative dimension of the official touristic offer. In our study we used only official data, recorded by the National Institute of Statistics , but, in reality, there are pretty much more other tourists (especially Romanians, but also foreigners) that, during vacation, are either accommodated at their families or friends, either own a vacation house in his station It is clear that, in the last 10 years, in Mamaia, a lot of vacation residences have been built, especially apartments, in residential buildings, but there isn't a clear evidence of those. We estimate that they have a total capacity between 3.000 and 5.000 of places.

In these conditions, we can talk about 2 dimensions of the touristic market of Mamaia station: a touristic market recognized officially (where the touristic demand is expressed through the indicators of touristic circulation used in our study, and the touristic offer can be found in the "accommodation capacity" indicator) and a real touristic market, unofficial, bigger than the official, but unrecognized market, due to a lack of modern recording system, of the tourists hosted in other spaces than the classified ones, but also of the buildings in Mamaia stating that they are used as vacation houses or secondary houses.

HISTORIC LANDMARKS IN THE EVOLUTION OF MAMAIA STATION

Mamaia station celebrates 110 years in August 2016. The history of the apparition of this station is connected with the beginning of the extension works of the port in Constanta, in 1889, by the team of the engineer Anghel Saligny. Due to this project, the beach of Constanta, known as "Baile de la Vii" was to be disbanded. At the same time, the concern of finding a new place for the beach emerged, so that Constanta develops as a "summer station of sea baths (http://constanta-imagini-vechi.blogspot.ro; www.i-tour.ro/).

The attention of the local authorities directed towards the Northern area of the city where, at a distance of 5 km of Constanta, the area between Siutghiol Lake and the Black Sea presented itself as a stripe of desolated seaside, with sand dunes, whose height reached 5-7 km, and in the end of this stripe a modest settlement of fishermen could be found, with distanced houses and a famous water mill, at the end of the Siutghiol Lake. (www.academia.edu/).Even from the end of the 19th century, this settlement has represented a place especially chosen for spending the summer holiday. The first constructions were made in 1906 (www.academia.edu/). The establishment, built from wood, resembled a mountain house and had, at the same time, a role as a train station. We can also mention that, on the Northern and Southern swills of the establishment, a long pavilion was built, with a watchtower at the end, and each watchtower had 56 cabins (http://constanta-imagini-vechi.blogspot.ro; www.i-tour.ro/). This has the endowment in 1906, considered the moment when the activity of the new station "Mamaia baths" began.

Mamaia was recognized as a station on the 22nd of August 1906, and the official recognition took place on the 28th of August 1906, when 2 trains that went from Constanta to Mamaia baths, in one hour and a half (the access in the station was made

through trains, but also through carriages). It is mentioned the fact that, in that year (1906), there were over 45.000 tourists registered in Mamaia. (http://constanta-imagini-vechi.blogspot.ro)
Even from the beginning of its activity, everyone predicted a bright future for the station. That is why in the newspaper "La Roumanie", the journalists wrote:"Constanta has, beginning with today, a beach that situates itself among the great stations in the East" (http://constanta-imagini-vechi.blogspot.ro).

Between 1906 and the First World War, the development of the station was little significant, which meant only the apparition of several small villas, of some restaurants for fishermen, and the beginning of building a park in 1913, the works being interrupted by the beginning of the war. (http://constanta-imagini-vechi.blogspot.ro) Even in this situation, of modest development, the interest of spending the vacation at the sea and especially at the Mamaia station was growing. We can make this assertion based on two sources of information from that time. The first information refers to the fact that a commercial appeared in the newspaper "The morning" in 30th of August 1912. The content of the message referred to the Mamaia station, but especially to a new way of spending the leisure time during the summer vacation (www.historia.ro/; adevarulfinanciar.ro/).

The second source regards the writer Cella Serghi, who, referring to the atmosphere of the summer vacations specific to that period, mentioned that: "You could get to Mamaia with the train: very crowded, a very colorful world, speaking French and German, a lot of children, governesses, comfortable wagons, very hot, numerous halts, in order to pick people up from the way" (http://constanta-imagini-vechi.blogspot.ro). The writer also mentions that, in those times, only the rich people used to go for a vacation at the sea side and her mention that most of them used to leave for Mamaia at the end of the week is very interesting. A marketing interpretation of these aspects leads to the following conclusions: Mamaia was an attractive destination for the summer vacations ("very crowded"…" numerous halts, in order to pick people up from the way"), for segments of tourists with high incomes ("only the rich people used to go to the sea for the summer vacation"), from superior social classes and a high level of education ("a world..speaking French and German), belonging to different age groups (" a very colorful world"…"many children"). Even from the beginning of its activity, the image of expensive station accessible only to a certain segment of tourists was created for Mamaia, preferred especially for spending time at the end of the week. The station begins to develop at the end of the First World War, beginning with the creation of the summer residence of the royal family. After the royal family received in 1920, from the town hall, a land of 4 ha, between the Siutghiol Lake and the sea, a small palace in the Romanian style was beginning to be built, as a summer residence (vladimirrosulescu-istorie.blogspot.com/).

During the years between the wars, through the building of new accommodation spaces, the station records, in 1938, a capacity of hosting 1067 people, being visited, in 1939, by 110 506 tourists. *(*www.academia.edu/)**.** With the aid of these two indicators, 110 506 tourists and 1067 accommodation places, assuming that it could have been a maximum degree of occupation of 100%, we can measure that the duration of the season was in 1939, of at least 103 days, which means almost 3 months and a half of a year. In reality, the season was longer because the degree of occupation was not the maximum possible for the entire duration of the season. Paradoxically, now, after almost 80 years, the data shows that the duration of the season has decreased, in comparison with that moment.

After the Second World War, for a period of almost 10 years, there weren't any new constructions built in Mamaia station. A few intervals of time follow when the station had a remarkable development. So, between 1959-1962, several hotels were built, in the whole station, which meant almost 10.000 new accommodation places. A new stage, of great constructions, in the whole station, with an alternation of small and big hotels, was in the years 1966-1970, and the interval 1975-1985 is remarkable through the apparition, in the middle of the station of another 8 new hotels, distinguished through the quality and diversity of the services. This was also the period when the station recorded a maximum of notoriety and/ attractiveness among the foreign tourists. The accommodation units with a high level of comfort were preferred by them, and the charges for the touristic services recorded significant growths. In 1987, over 352000 tourists spent their vacation here, and 90.000 were foreigners *(*www.academia.edu/).

From the data we analyzed, this level of the number of tourists, was exceeded in 2008 (when a number of 355124 tourists was recorded), then in 2009 (when the highest level of tourists so far was reached – 358707) and in 2012 (354519 tourists). Regarding the number of foreign tourists, registered in 1987, this number has never been reached so far. If we refer to the analyzed period, (2006-2015), we can notice that the highest level was registered in 2006 and it was of 31839 foreign tourists (which means almost 35.4% comparing with the level in 1987). We must take into account that, in comparison with the moment 1987, the accommodation capacity of the Mamaia station recorded a growth, in the period we analyzed, reaching 20030 accommodation places, in 2015.

After 1990, the station recorded evolutions sometimes contradictory from the point of view of the accommodation capacity. On one side, certain accommodation units (villas and hotels), from various reasons, degraded themselves, and on the other side, improvement and modernization works of the already existing units have been made and new units with a high comfort level have been built. *(*www.academia.edu/)**.** So, if in 2004, in Mamaia there were 80 accommodation units, with 19326 places, in 2015 there were recorded 86 accommodation units, with a capacity of 20030 places (www.constanta.insse.ro).

THE METHODOLOGY OF INVESTIGATION

It has been noticed in the economic practice that the economic variables (like the number of tourists' arrivals), together with the important influences fromfactorial variables (like, for instance, the level of prices, the level of income, the rate of inflation, exchange course), they also suffer a modification due to their auto-regressive character , based on the memorization of their previous behavior. From an econometric point of view, the term autoregressive defines the extent to which an economic variable presents the characteristic of autocorrelation, which means that its current level is determined, to a significant extent, by its previous levels, delayed with one or more periods of time beyond (expressed in calendar years, quarters or months). In this situation, we can assume that the effect on resultative variable (the number of tourists' arrivals in the case of our study) is not caused only by the direct influence of some factorial variables, but also of its past values, meaning the informational charge of the studied variable.

Mainly, the autoregressive effect emphasizes itself more or less in the way that the actual level of the analyzed variable it influenced by its previous levels, the difference in time of the influences having different values. For instance, the actual level of the analyzed variable can be influenced by the level registered by the same variable a year ago, two years ago etc. The bigger this difference is, which means that the movement behind is more emphasized, the weaker the influences over the current level of the variable. The problem that appears in this situation is establishing the moment when the delayed influences become insignificant, in order to be eliminated from the model. This autoregressive characteristic is related to the capacity of the environment of retaining the previous behaviors of the studied variable and reacting according to these. The economic subjects interested in formulating some forecasts highly accurate, which mean that those interested in the quantitative aspect of the modification emerged in the studied variable, follow in a systematic way this evolution, being oriented for the estimation of the anticipated level of the variable of its level from the previous periods. This sort of relationship can be studied with the aid of the autoregressive patterns of different orders. An autoregressive model requires a connection between the current level of the studied variable and its previous behaviors.

Depending on the number of the amounts of delayed time, there are, in the specialized literature, various types of econometric models: autoregressive models of the first order, where the effect in time is analyzed for a period of time (one year) behind (autoregressive model AR (1)) and models of a superior order, two years (the autoregressive model AR (2)), three years (the autoregressive model AR (3)) etc, where the effect in time is studied for a k periods of time behind. Depending on the existing conditions in economy, the forecasts of the operators can be based not only on the immediately previous level of the studied variable (level that can be emphasized by the autoregressive model of the first order), but also by its older behavior, delayed with two or more periods of time behind. This way, autoregressive models of superior orders can be built.

In the modeling of the series of time, a very important aspect is represented by the concept of stationarity. In general, the stochastic process is considered stationary if the media and the variation are constant in time, and the covariance from two periods of time depends only on the distance from the two periods. In other words, a series is considered stationary if its evolution is constant in time.

The conditions that must be fulfilled in order for a stochastic process to be considered stationary can be synthesized this way:

$M(Y_t) = \mu$, the average is constant and independent in time

$Var(Y_t) = M(Y_t - \mu)^2 = \sigma^2$, the variance is constant in time

$M(Y_t - \mu)(Y_{t-k} - \mu_t) = \gamma_k$, covariance depends only on the temporal difference (Asteriou, D., Hall, S., 2011).

The stationary characteristic of the series of time is important because, based on the analysis of the behavior of the stationary variables, different forecasts can be made of the studied variables. When the series of time are un-stationary, the results obtained from the modeling cannot be generalized, and, consequently, cannot be used for anticipating moments in the future. Also, when making research based on the un-stationary series of time (this being predominant in economy), the phenomenon of false regression can emerge (spurious regressions). In this situation, the value of the report of determination is high, which indicates, according to interpretations, a strong connection between the variables, even though there is no correlation between the analyzed variables

(Mahadeva şi Robinson, 2004). Because of this, the macroeconomic series of time must be subdued to a logarithm and stationaried, because the results obtained after the regression analysis must be conclusive.

In economy, the effects of the impact of the independent variables over a dependent variable are realized after a certain period of time, not immediately. That is why, the “lag operator” (delay) is used, which realized the connection between the values of the series of time from a *t* moment and the values registered at a previous moment. When an un-stationary series of time is differentiated until the order *d*, when a stationary series is obtained, we can conclude that the initial series of time is integrated in the *d* order.

In the case of a stationary series(which doesn’t need the differentiation), this is considered integrated in the order 0. In general, in economy, we can notice that the un-stationary series are integrated in the first order (which means they only need one differentiation).

In econometry, for the modeling of the series of time, we can use several types of stochastic models: autoregressive models (AR), models of mobile medium type(MA) and models built based on these(ARMA). The variables have a strong inertial character, the actual values of the variables being influenced by the past evolution of the phenomenon (the autoregressive component), and the shocks produced on them are quantified by the medium mobile type component. (Gujarati, D.,2003).

The autoregressive models (AR) are characterized by the fact that the values of the variable Y at the moment t depend on the previous values of the same variable. The auto determination, this dependence, according to its previous values, emphasized through the autoregressive models, can generate a series of compensatory series of actions, with the aim of counteracting some interference (mix) from outside the system.

For instance, the absence of a required product on the market produces an excessive growth of the sales, in the following period. Or, a speculative action at the bursary can trigger a significant deviation of cotations, apart from the level of normal fluctuations. This modification will be, very probably, followed by an opposed deviation, which must be understood as a correction movement, with the aim of situating themselves around the equilibrium level

In the formal representation, specific to forecasting levels, this sort of fixing manifestations are rendered through the mobile medium level MA (movie average). O mobile medium model is a lineary combination of residual terms. The general form of a medium mobile model of order q MA(Q), is defined through the relationship:

$Y_t = \alpha_0 + \alpha_1\varepsilon_t + \alpha_2\varepsilon_{t-1} + \ldots + \alpha_q\varepsilon_{t-q}$,

where $\alpha_0, \alpha_1, .., \alpha_p$ – the parameters of the model

ε_t – the rezidual term(Gujarati, D.,2003).

Based on these elements, an ARMA model is defined as an aggregate which includes the recepted impulses with delays on p intervals of time, as the reactions expressed in average at the accidental deviations from the linear evolution, manifested with q periods of time behind. (Brooks, C.,2008). Consequently, an autoregressive – medium mobile model ARMA 9(p, q) is made of an autoregressive component and a medium mobile component $Y_t = \beta_0 + \beta_1 Y_{t-1} + \ldots + \beta_p Y_{t-p} + \varepsilon_t + \alpha_1\varepsilon_{t-1} + \ldots + \alpha_q\varepsilon_{t-q}$,

where p is the order of the autoregressive part, q is the order of the medium mobile one and ε_t is a white noise type process.

The models AR(p) and MA(q) can be described as particular forms of the ARMA type model 9(p,q):

AR(p)=ARMA(p,0)

MA(q)=ARMA(0,q)

When the initial series of time is un-stationary, the operator difference of order d is applied and we obtain a stationary station that can be modeled through the ARIMA types processes. (p,d,q). The ARIMA models (p,d,q) are based on the ARMA models that do not respect the stationarity property.

Box and Jenkins (1976) elaborated a general methodology to create forecasts for the series of time univariated. These specialists saw that the majority of the series of time are un-stationary, but these can be modeled using integrated models and medium mobile types (ARIMA). By applying the Box-Jenkins methodology, in our study, we will estimate one model for each series of time (for each variable) analyzed and we will realize forecasts for two consecutive years. (2016, 2017). The steps necessary for the estimation of an ARIMA (p,d,q) type process, using the Box-Jenkins type, are the following: testing the stationarity of the series; the identification of the type of model; the estimation of the parameters of the model; the validation of the model; realizing forecasts based on the chosen model.

DATA, RESULTS AND DISCUSSIONS

For the econometric modeling of the three indicators showing the total number of tourists, the number of Romanian tourists and number of foreign tourists in the Mamaia station, we used an annual series of data from 2006-2015. The remaking of data was made with the E-views program. Because the used series were unstationary, we used more statistical tests, in order to transform them in a stationary series. After applying these tests, series became stationary and integrated in the first order; therefore, in the determination of ARMA, we used the Box-Jenkins procedure, after which, the performances of the chosen ARIMA model were verified based on the classic statistical model. The last step of the study also had in view the realization of some forecasts for 2016 and 2017.

4.1. The modeling of the variable "total number of tourists"

a. Testing the stationarity of the series of time
The results obtained and presented in the tables 4.1. and 4.2 indicated the fact that the series of time is un-stationary (the probability of the test t-student is bigger that the significance threshold of 5%), but, applying the difference operator, the null hypothesis is validated, the series of time becoming stationary after the first differentiation (d=1), for a level of relevance of 10% (Dickey, D., Fuller, W.,1979).

b. The identification of the type of modelWith the aid of the data from the graphic of the autocorrelation and partial correlation functions of the stationary series (table 4.3), we could determine p and q parameters, also keeping in mind that:
° The function of correlation decreases suddenly towards 0, after the second term, anticipating a MA process.
° The value of the partial correlation function decreases towards 0, beginning with the third term, so it is recommended a AR(2) type process. (Engle, R.,1982).
Consequently, the order of the autoregressive component is p=2 and the order of the medium mobile component is q=2.

c. Criteria of choosing the best model
Based on the results obtained in the previous step, we proposed the analysis of various autoregressive integrated and medium mobile models, depending on the p and q indicators. The parameters of the models are estimated through the method of the smallest squares. According to value of informational criteria that are emphasized by the values of the determination report (column $\mathbf{R^2}$ from table 4.4 and keeping in mind the differentiation from the previous step, for the modeling of the initial series of time, the specification ARIMA is chosen (2,1,2).
The equation of the ARIMA(2,1,2) model is the following:
$$\Delta \mathbf{Y_t} = \mathbf{12.67+0.718}\ \Delta \mathbf{Y_{t-1}}$$
where Y_t = the logarithm of the analyzed variable

d. Testing the validity of the ARIMA model (2,1,2)
With the aim of appreciating the quality of the econometric process, which is the validity of the model, we analyzed a series of aspects related to: testing the parameters of the model (table 4.5), testing the hypothesis according to which the average of errors is null (table 4.6), testing the hypothesis of homoscedasticity (table 4.7), testing the hypothesis of normality of the errors (table 4.8), testing the hypothesis of the un-correlation of the errors (table 4.9).
The results obtained after the tests indicate the following:

- Because the value of the probability of the test t is bigger that the chosen threshold of significance of 0.05, we accept the hypothesis according to which the average of errors is null
- Because the value of the associated probability associated with the test Chi-square is bigger than 0.05, we accept the hypothesis of homoscedasticity, with a probability of 95%
- Because the probability associated to the Jarque Bera test is bigger than 0.05 , we accept the null hypothesis, according to which the series is normally distributed,
- Acording to the correlogram of the errors, we notice that there isn't a serial autocorrelation of the errors until lag 8, because the probability associated to the test for every delay is bigger than the value of the significance threshold.

e. Forecasting the analyzed series of time
In our study, in order to test the capacity of the estimated model of forecasting the evolution of the variable "total number of tourists", based on the real registered data, we used the dynamic version of forecasting (out-of-sample). We made a forecast that includes the annual values of the variable for 2016-2017.
In the table 4.10 the forecasted values of the variable "total number of tourists" in Mamaia are foreseen, for 2016-2017.

Table 4.10 The nominal values foreseen for the total number of tourists

Year	**The total number of tourists foreseen by the model ARIMA (2,1,2) in Mamaia**
2016	317094
2017	321180

The foreseen growth of the total number of tourisst in Mamaia station in 2016, in comparison with 2015, should be relatively modest, of only 102%. We have reasons to think that, in reality, the growth of the total number of tourists, in total, will be much bigger, because, on one side, of the declarations of the representative figures of the Patronat of tourism of Mamaia station, which foresaw big growths of the number of early booking reservations, in comparison to the previous years, but also the contracts with the foreign tour-operators. An important reason of this grown touristic demand is also due to the fact that the first step rehabilitation and extension of the beaches is finished, which included also the Southern part of Mamaia station, the tourists being able to enjoy here, in 2016, generously extended beach areas.

4.2. The modelation of the number of Romanian tourists in Mamaia

a. Testing the stationarity of the series of time

The results from tables 4.11 and 4.12 indicate the fact that the series of time is un-stationary (the probability of the t-student test is bigger than the significance threshold of 5%), but, by applying the difference operator, the null hypothesis is validated, the series of time being stationary, after the first differentiation (d=1), for a relevance level of 5%.

b. Identifying the type of model

With the aid of the data from the graphic of autocorrelation and partial autocorrelation functions of the stationary series (table 4.13), we could determine the p and q parameters, keeping in mind that:

- the autocorrelation station decreases suddenly towards 0, after the first term, a process of the MA(1) type being anticipated.
- the value of the partial autocorrelation function decreases towards 0, beginning with the second term, so a process of the AR(1) type is recommended.

In conclusion, the order of the autoregressive component is p=1 and the order of the medium mobile component is q=1.

c. Criteria of choosing the best model

According to the value of informational criteria that are emphasized by the values of determination report (column $\mathbf{R^2}$ from table 4.4) and keeping in mind the differentiation of the series from the previous step, for the modeling of the initial series of time, the ARIMA (1,1,1) specification is chosen.

The values of the informational criteria (Akaike, Schwarz, Hannan-Quinn) for the analyzed models are indicated in table 4.14.

The equation of the ARIMA model (1,1,1) is the following:

$\mathbf{\Delta Y_t = 12.66 + 4.242\, \varepsilon_{t-1}}$

where Y_t = logarithm of the analyzed variable

d. Testing the validity of the ARIMA model (1,1,1)

The results obtained after testing this model (tables 4.15, 4.16., 4.17, 4.18, 4.19) indicate the following:

- because the value of the probability of the t test is bigger than the significance threshold of 0.05, we accept the hypothesis according to which the average of errors is null.
- Because the value of the probability associated to the test Chi-square is bigger than 0.05, we accept the homoscedasticity hypothesis, with a probability of 95%
- Because the probability associated to the Jarque Bera test is bigger than 0.05, we accept the null hypothesis, according to which the series is normally distributed.
- According to the correlogram of errors, we notice that there isn't a serial autocorrelation of the errors, until lag 8, because the probability associated to the test for every delay is bigger than the value of the significance threshold.

e. The forecast of the analyzed series of time

In our study, in order to test the capacity of the estimated model in order to predict the evolution of the variable "number of Romanian tourists", based on the real registered data, we used the dynamic forecasting version (our-of-sample). This way, we made a forecast that contains the annual values of the variable for 2016-2017.

In the table 4.20 the forecasted values of the variable "number of Romanian tourists" in Mamaia are foreseen, for 2016-2017.

Table 4.20 The nominal values foreseen for Romanian tourists:

Year	**Number of Romanian Tourists foreseen by the ARIMA model (1,1,1) in Mamaia**
2016	313875
2017	313881

The foreseen growth of the number of Romanian tourists in Mamaia station in 2016 , will be only of 5,6% but in 2015, the registered growth was of only 0.5% in comparison with the previous year.

4.3. The modeling of the number of foreign tourists in Mamaia

a. Testing the stationarity of the series of time

The results from the table 4.21 and 4.22 indicate the fact that the series of time is un-stationary (the probability of the t-student test is bigger than the significance threshold of 5%), but, by applying the difference operator, the null hypothesis is validated, the series of time being stationary, after the first differentiation (d=1), for a relevance level of 5%.

b. Identifying the type of model

With the aid of the data from the graphic of autocorrelation and partial autocorrelation functions of the stationary series (table 4.13), we could determine the p and q parameters, keeping in mind that:

- -he autocorrelation station decreases suddenly towards 0, after the first term, a process of the MA(1) type being anticipated.
- the value of the partial autocorrelation function decreases towards 0, beginning with the second term, so a process of the AR(1) type is recommended.

In conclusion, the order of the autoregressive component is p=1 and the order of the medium mobile component is q=1.

c. Criteria of choosing the best model

The values of the informational criteria (Akaike, Schwarz, Hannan-Quinn) for the analyzed models are indicated in the 4.24 table. According to the value of the informational criteria that are emphasized by the values of the determination report (column $\mathbf{R^2}$ from table 4.24 **)** and keeping in mind the differentiation from the previous step, for the modeling of the initial series of time , the ARIMA (1,1,1) specification is chosen.

The equation of the ARIMA model (1,1,1) is the following:

$\mathbf{\Delta Y_t = 9.82+0.412\ \Delta Y_{t-1}}$

where Y_t = logarithm of the analyzed variable

d. Testing the validity of the ARIMA model (1,1,1)

The results obtained after the testing this model (tables 4.25, 4.26., 4.27, 4.28, 4.29) indicates the following:

- because the value of the probability of the t test is bigger than the significance threshold of 0.05, we accept the hypothesis according to which the average of errors is null.
- Because the value of the probability associated to the test Chi-square is bigger than 0.05, we accept the homoscedasticity hypothesis, with a probability of 95%
- Because the probability associated to the Jarque Bera test is bigger than 0.05, we accept the null hypothesis, according to which the series is normally distributed.
- According to the correlogram of errors, we notice that there isn't a serial autocorrelation of the errors, until lag 8, because the probability associated to the test for every delay is bigger than the value of the significance threshold.

e. The forecast of the analyzed series of time

In our study, in order to test the capacity of the estimated model in order to predict the evolution of the variable "number of foreign tourists", based on the real registered data, we used the dynamic forecasting version (out-of-sample). This way, we made a forecast that contains the annual values of the variable for 2016-2017.

In table 4.30 the forecasted values of the variable number of foreign students in Mamaia are emphasized, for 2016-2017.

Table 4.30 The nominal values foreseen for foreign tourists

Year	**Number of foreign toutists foreseen by ARIMA(1,1,1) model,in Mamaia**
2016	18300
2017	19378

The foreseen growth of the number of foreign tourists in 2016, in Mamaia, should be of 14.48%, and in 2015, a significant decrease of 16% was registered, in comparison with 2014. We think that the efforts of touristic publicity of Mamaia, from the last two years, will begin to be productive, from 2016, the actual number of foreign tourists arrived in this station being under the possible one and under the one registered in the previous periods. (21861, in 2012, respectively 29273, in 2007 or 31839, in 2006).

CONCLUSIONS

The present study is part of a large series of analysis, from the area of touristic activity, where our aim is to realize, by applying certain econometric model, various quantitative research of the touristic phenomenon/touristic market, that could offer, at the same time, a database for qualitative approaches/interpretations also. As the content of the work also shows, after going through the specific work steps, the ARIMA model probed itself appropriate for processing our series of database (from the interval 2006-2015), and it allowed us to provide the evolution of the indicators of touristic circulation followed in our study, respectively to calculate the number of tourists (total, Romanian and foreigners) that should arrive in 2016 and 2017 in Mamaia station.

The results obtained show that growths of the three indicators are expected in 2016 and 2017. A limit of our study is that we provided only one component of the touristic market, respectively the touristic demand. The other component, the touristic offer, we did not provide it (for technical, spatial reasons), but we will include it in our work agenda, for future studies.

Our concern is to interpret the eventual opportunities and threats, that can emerge as an effect of this evolution, to identify who it is and how will it experience these opportunities/threats, who and how should it interfere in order to capitalize in a better way the opportunities, in order to diminish even more the threats.

We mentioned that in the content of the study the processed data have official value of the indicators of touristic circulation. But these official values do not express the dimension of the effective market, reached in certain moments (every year in the interval 2006-2015). The lack of data of the real demand and offer, un-official, does not allow us to evaluate of some aspects as: if there ever was/is/will be un-satisfied touristic demand of offer specific to every moment or if there remained/remains/will remain touristic offer that is not covered by demand. (Jugănaru,1998)

The tendency of growth of the number of tourists, in the following two years, in Mamaia station, should be of interest for establishing the strategy and the business plan of several entities, among which:

-actual economic agents that develop their activity in this station

- all the economic organizations that appear as providers of goods and services and have a connection with the touristic activity in Mamaia

-local administration

-potential investors etc.

The foreseen growth of the number of tourists may be equal to a growth of the charges of the economic operators that activate in the touristic domain or the ones that have a connection with the touristic activity, but also growth of the charges at the local budget. These income growths may be a base for future investments, that may lead to the development of the activity from a quantitative perspective (the growth and diversification of the material base), but also from a qualitative perspective (the growth of the comfort level, the growth of the quality of the services provided).All the aspects of the quantitative and qualitative growth must be reflected in the growth of value offered to the tourist, in the degree of satisfaction experienced by the tourist, but also in the growth of the image and prestige of the station ("Mamaia" brand). It is known that, in the case of a satisfied tourist, there is the possibility for him to come back and recommend the Mamaia station to other people, causing new growths of the number of tourists. (Kotler, Kartajaya şi Setiawan, 2010, Kotler şi Armstrong,2008; Juganaru,2000). We should also pay attention to the other alternative as well: the foreseen growth of the number of tourists can lead to the emergence of eventual unpleasant situations as: over-crowded (in station, in the accommodation units, in restaurants, parking spots, transportation etc) or the incapacity of some economic agents and authoritities of the state/local administration to manage and face, through their activity, these growths of the number of tourists. Moreover, as an example, we must say that the lack of insufficient number of parking spots, for the tourists' cars, in this station, is still an unsolved issue, despite the fact that at the beginning of the touristic season 2016, a new bunk parking lot was opened, with a capacity of 422 parking spot. The eventual appearance of some unpleasant situations may lead to the decrease of the quality of the offered services, the decrease of the degree of the tourist's satisfaction, reason for which he will refuse to come back in this station and, in this way, it is possible to lower the number of tourists at a certain moment.

In order to avoid these situations, we think that it is necessary to elaborate a strategy of developing the area, carefully analyzed and able to provide a correlation of all the facts that influence the demand, but also the touristic offer.

REFERENCES

[1] Asteriou, D., Hall, S.,2011, Applied econometrics. Palgrave Macmillan

[2] Brooks, C., 2008, Introductory econometrics for finance. Cambridge University Press, New York

[3] Dickey, D., Fuller, W.,1979, Distribution of the estimators for autoregressive time series with a unit root. Journal of the American Statistical Association, Vol.74, Nr.366, pp.427-431

[4] Engle, R.,1982, Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. Econometrica, Vol50, Nr.4, pp.987-1007

[5] Gujarati, D.,2003, Basic econometrics (fourth edition). McGraw-Hill Higher Education, New York

[6] Jugănaru, I. D., 2007. Politici şi strategii în turismul mondial. Bucureşti: Editura Expert.

[7] Juganaru,M.1998.Teorie şi practica in cercetarea de marketing.Bucureşti: Editura Expert

[8] Juganaru,M.2000.Marketing.Bucureşti: Editura Expert

[9] Kotler, Ph., Kartajaya, H. şi Setiawan, I., 2010. Marketing 3.0: de la produs la consumator şi la spiritul uman. Bucureşti: Editura Publica.

[10] Kotler,Ph., Armstrong,G.,2008.Principiile marketingului.Bucureşti: Editura Teora

[11] Mahadeva, L., Robinson, P., 2004, Unit root testing to help model buiding. Centre for Central Banking Studies, Nr.22, Bank of England

[12] Minciu,R.,2005.Economia turismului. Bucureşti :Editura URANUS

[13]www.academia.edu/7300594/Istoricul_stațiunii_Mamaia

[14] adevarulfinanciar.ro/.../109-ani-de-mamaia-o-istorie-a-fitelor

[15] http://constanta-imagini-vechi.blogspot.ro/2012/07/mamaia-numele-unei-legende-partea-i.html

[16] www.constanta.insse.ro

[17] www.historia.ro/exclusiv_web/general/articol/concediu-1912-b-ile-mare-mamaia

[18] www.i-tour.ro/istoria-statiunii-mamaia)

[19] (www.turismistoric.ro).

[20] vladimirrosulescu-istorie.blogspot.com/2012/09/mamaia.html

Tables

Table 2.1. *Number of tourists arrived in Mamaia2006-2015*

Nr..	Year	Total number of tourists	Number of Romanian tourists	Number of foreign tourists
1	2006	278707	246868	31839
2	2007	351496	322223	29273
3	2008	355124	335679	19445
4	2009	358707	340233	18474
5	2010	336106	315215	20891
6	2011	331823	314462	17361
7	2012	354519	332658	21861
8	2013	327899	310340	17559
9	2014	314610	295633	18977
10	2015	313232	297248	15984

Table 4.1. *The values of the Augmented Dickey-Fuller test for the total arrivals of tourists*

Logarithmic variable	**t-statistic**	**Prob.**	**t-statistic (differentiated series)**	**Prob.**
Total number of tourists	-0.404	0.777	-5.418	0.0002

* the critical value of the ADF test for a significance level of 5% is -3.25
* the critica value of the ADF test for a significance level of 10% is -2.77

Source: E-Views

Table 4.2. *Values of Phillips-Perron test for total arrivals of tourists*

Logarithmic variable	**t- statistic**	**Prob.**	**t-statistic (differentiated series)**	**Prob.**
Total number of tourists	-0.404	0.777	-6.174	0.0001

* the critical value of ADF test for a significance level of 5% is -3.25
* the critical value of ADF test for a significance level of 10% is -2.77

Source: E-Views

Table 4.3. *The autocorrelation and partial correlation functions for the total number of the tourists' arrivals*

Sample: 2006 2017
Included observations: 8

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.217	-0.217	0.5397	0.463
		2	-0.045	-0.097	0.5666	0.753
		3	0.066	0.037	0.6369	0.888
		4	-0.271	-0.267	2.1103	0.715
		5	0.261	0.169	3.9247	0.560
		6	-0.134	-0.105	4.6454	0.590
		7	-0.159	-0.177	6.6738	0.464

Table 4.4. *The values of informational criteria for the total number of tourists arrivals*

	R^2	**AIC**	**Schwarz**	**Hannan-Quinn**
ARIMA(2,1,2)	0.65	**-3.46**	**-3.43**	**-3.66**
ARIMA(1,1,2)	0.06	-2.51	-2.44	-2.65
ARIMA(1,1,1)	0.38	-2.98	-2.92	-3.12
ARIMA(2,1,1)	0.46	-3.01	-2.98	-3.22

Table 4.5. *Testing the coefficients of the model for the toal number of tourists arrivals*

Variable	**Coeficient**	**Prob.**
C	12.675	0.0000
AR(1)	0.718	0.0350
AR(2)	-0.073	0.8802
MA(1)	-0.260	0.8173
MA(2)	-0.739	0.5136

Table 4.6. *Testing the hypothesis according to which the average of the errors is null for the total number of tourists' arrivals*

Test of Hypothesis: Mean = 0.000000

Sample Mean = 0.007630
Sample Std. Dev. = 0.030419

Method	Value	Probability
t-statistic	0.709440	0.5010

Table 4.7. *ARCH test for the total number of tourists' arrivals*
Heteroskedasticity Test: ARCH

F-statistic	0.176950	Prob. F(1,5)	0.6915
Obs*R-squared	0.239262	Prob. Chi-Square(1)	0.6247

Figure 4.8 *The histogram of the distribution for the total number of tourists' arrivals*

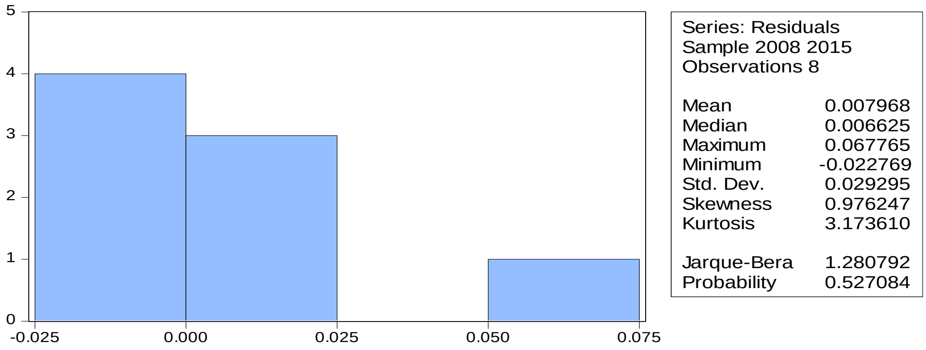


Table 4.9. *Corelogram of errors for the total number of tourists' arrivals*

Sample: 2008 2015
Included observations: 8
Q-statistic probabilities adjusted for 4 ARMA term(s)

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.352	-0.352	1.4147	
		2	-0.069	-0.220	1.4780	
		3	-0.006	-0.136	1.4786	
		4	-0.235	-0.373	2.5848	
		5	0.175	-0.141	3.4036	0.065
		6	-0.054	-0.200	3.5194	0.172
		7	0.041	-0.141	3.6532	0.301

Table 4.11. *The values of Augmented Dickey-Fuller test for the number of Romanian tourists arrivals*

Logarithmic variable	t-statistic	Prob.	t-statistic (differentiated series)	Prob.
Total number of Romanian tourists	-1.519 0.0067		0.473	-4.926

* critical value of the ADF test for a significance level of 5% is -3.25
* critical value of the ADF test for a significance level of 10% is -2.77

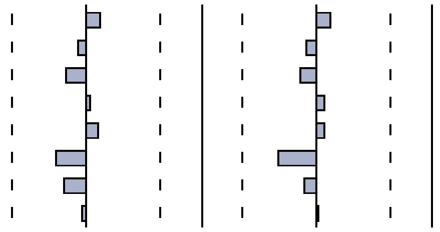
Table 4.12. *Values of Phillips-Perron test for the number of arrivals of Romanian toursists*

Logarithmic variable	t- statistic	Prob.	t-statistic (differentiated series)	Prob.
Total number of Romanian tourists	0.581 0.0001		0.821	-6.792

* critical value of the ADF test for a significance level of 5% is -3.25
* critical value of the ADF test for a significance level of 10% is -2.77

Table 4.13. *The autocorrelation and partial correlation functions for the number of Romanian tourists arrivals*

Sample: 2006 2017
Included observations: 9



Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.129	0.129	0.2071	0.649
		2	-0.071	-0.089	0.2781	0.870
		3	-0.173	-0.156	0.7741	0.856
		4	0.033	0.073	0.7950	0.939
		5	0.104	0.072	1.0648	0.957
		6	-0.275	-0.340	3.5609	0.736
		7	-0.203	-0.108	5.6042	0.587
		8	-0.044	0.010	5.7942	0.670

Table 4.14. *The values of informational criteria for the number of Romanian tourists arrivals*

	R^2	AIC	Schwarz	Hannan-Quinn
ARIMA(1,1,1)	0.95	**-5.66**	**-5.59**	**-5.80**
ARIMA(1,1,0)	0.01	-2.82	-2.78	-2.92
ARIMA(0,1,1)	0.78	-3.13	-3.07	-3.20

Table 4.15. *Testing the coefficients of the model for the number of Romanian tourists arrivals*

Variable	Coeficient	Prob.
C	12.669	0.0000
AR(1)	-0.017	0.8099
MA(1)	4.242	0.05

Table 4.16. *Testing the hypothesis according the which the average of the errors is null for number of Romanian tourists arrivals*

Test of Hypothesis: Mean = 0.000000

Sample Mean = -0.000368
Sample Std. Dev. = 0.010822

Method	Value	Probability
t-statistic	-0.101947	0.9213

Tabelul 4.17. *Testul ARCH pentru sosiri turişti români*

Heteroskedasticity Test: ARCH

F-statistic	0.708793	Prob. F(1,6)	0.4321
Obs*R-squared	0.845210	Prob. Chi-Square(1)	0.3579

Table 4.18*.Testing the hypothesis of normality of the errors for number of Romanian tourists arrivals*

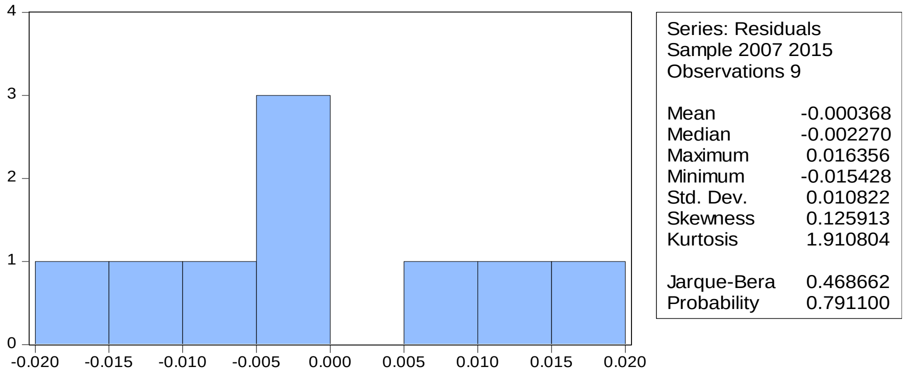


Table 4.19. *Correlogram of errors for the number of Romanian tourists arrivals*

Sample: 2007 2015
Included observations: 9
Q-statistic probabilities adjusted for 2 ARMA term(s)

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.364	0.364	1.6373	
		2	-0.175	-0.354	2.0707	
		3	0.046	0.344	2.1050	0.147
		4	0.130	-0.172	2.4386	0.295
		5	-0.240	-0.243	3.8608	0.277
		6	-0.401	-0.181	9.1702	0.057
		7	-0.197	-0.137	11.098	0.049
		8	-0.026	0.008	11.165	0.083

Table 4.21. *The values of the Augmented Dickey-Fuller test for the number of Romanian tourists arrivals*

Logarithmic variable	t-statistic	Prob.	t-statistic (differentiated series)	Prob.
Total number of foreign tourists	-2.229	0.209	-3.917	0.023

* critical value of the ADF test for a significance level of 5% is -3.25
* critical value of the ADF test for a significance level of 10% is -2.77

Table 4.22. *Valorile testului Phillips-Perron pentru sosiri turişti străini*

Logarithmic variable	t- statistic	Prob.	t-statistic (differentiated series)	Prob.
Total number of foreign tourists	-2.521	0.1415	-3.991	0.023

* critical value of the ADF test for a significance level of 5% is -3.25
* critical value of the ADF test for a significance level of 10% is -2.77
Source: *E-Views*

Table 4.23. *The autocorrelation and partial correlation for the number of Romanian tourists arrivals*

Sample: 2006 2017
Included observations: 9

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.439	-0.439	2.3878	0.122
		2	0.220	0.034	3.0743	0.215
		3	-0.103	0.007	3.2498	0.355
		4	-0.206	-0.316	4.0872	0.394
		5	0.097	-0.132	4.3194	0.504
		6	-0.170	-0.150	5.2689	0.510
		7	0.098	-0.100	5.7460	0.570
		8	0.002	-0.056	5.7465	0.676

Table 4.24. *The values of the informational criteria for the number of Romanian tourists arrivals*

	R^2	AIC	Schwarz	Hannan-Quinn
ARIMA(1,1,1)	0.71	**-1.32**	**-1.25**	**-1.46**
ARIMA(1,1,0)	0.27	-0.62	-0.57	-0.71
ARIMA(0,1,1)	0.18	-0.04	0.01	-0.11

Table 4.25. *Testing the coefficients of the model for the number of Romanian tourists arrivals*

Variable	Coeficient	Prob.
C	9.821	0.0000
AR(1)	0.412	0.0047

Table 4.26. *Testing the hypothesis according to which the average of errors is null for the number of Romanian tourists arrivals*

Test of Hypothesis: Mean = 0.000000

Sample Mean = 0.034646
Sample Std. Dev. = 0.087615

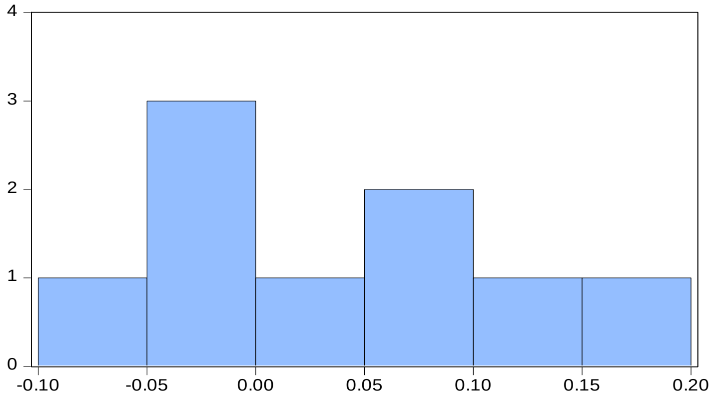
Method	Value	Probability
t-statistic	1.186312	0.2695

Table 4.27. *ARCH test for the number of Romanian tourists arrivals*

Heteroskedasticity Test: ARCH

F-statistic	1.606712	Prob. F(1,6)	0.2519
Obs*R-squared	1.689784	Prob. Chi-Square(1)	0.1936

Table 4.28. *Testing the hypothesis of normality for the number of Romanian tourists arrivals*



Series: Residuals
Sample 2007 2015
Observations 9

Mean	0.034646
Median	0.026097
Maximum	0.167750
Minimum	-0.092620
Std. Dev.	0.087615
Skewness	0.146115
Kurtosis	1.784777
Jarque-Bera	0.585812
Probability	0.746092

Table 4.29. *Corelogram of errors for the number of foreign tourists'arrivals*

Sample: 2007 2015
ncluded observations: 9
Q-statistic probabilities adjusted for 2 ARMA term(s)

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.400	-0.400	1.9799	
		2	0.140	-0.023	2.2588	
		3	-0.326	-0.330	4.0081	0.045
		4	-0.045	-0.389	4.0487	0.132
		5	0.064	-0.187	4.1498	0.246
		6	0.102	-0.086	4.4909	0.344
		7	0.174	0.108	5.9906	0.307
		8	-0.209	-0.124	10.325	0.112