Analysis of the Impact of the Evolution of the Car Rental Sector in the Automotive Logistics: Application to the Entreposto Logística Case Study

Francisco Spohr Ribeiro Coelho[1]

[1] Master in Industrial Engineering and Management, Instituto Superior Técnico, Universidade de Lisboa - Portugal

Abstract. The current socio-economic context forces companies to focus on optimizing operations and processes in order to strengthen their market value and maximize their competitiveness. It is in this context that an effective prediction becomes a key factor, since it results in a larger and more accurate decision making both in the long and short term. Therefore, there are various levels of planning under which the prediction proves to be valuable. The more accurate is the prediction, the greater is the responsiveness of the company as well as more optimized are the processes and the information and material flows. It is in this context that the present work arises, whose aim is to forecast Entreposto Logística´s (EL) activity volume, described by the quarterly and annual volume of vehicles exits from the enterprise´s facilities.

The record growth of the tourism sector in Portugal has boosted the success of the national car rental industry, which is characterized by the need for high traffic in order to fulfil its mission of bringing mobility to all customers, and by the brief use of rental cars (whose life spans are thus decreasing), so that rent-a-car companies can often provide new vehicles to their customers by renewing their inventory on a regular basis.

Thus, it is expected that the fact that this sector, which possesses particular characteristics, is gaining more prominence in the automotive market will have repercussions on automotive logistics and related operations and activities.

On the other hand, since the majority of Entreposto Logística customers are car importer brands/firms whose main customers are rent-a-car companies, this industry has a high impact on the company's revenues. This is another reason why the EL Board showed interest in an estimate that also took into account the impact of the car rental market evolution on the demand for the company's logistics activities.

Therefore, this master's dissertation will focus on the study and development of a resolution approach that will allow to support EL´s operational strategy in the near future, namely in terms of available capacity planning and operational costs minimization.

Based on the case studies analysed on the literature review section, the Holt-Winters (additive and multiplicative) method, the time series decomposition method and the linear regression analysis (single and multiple) procedure were selected as the approach to be applied to the problem in hands.

At the end of this study, it was possible to conclude that in a quarterly perspective the Holt-Winters´s multiplicative model was the one that proved to be the most adequate and, in turn, in an annual perspective the model named Model 2 was the one that proved to be the most accurate.

Keywords: Forecasting Models, Logistics, Rent-a-Car.

1. Introduction

The car rental business has been undergoing a considerable growth over the years and has gained a position of increasing importance in the market (Pachon, 2006). This growth has been boosted by the development of the automotive sector, which currently allows the acquisition and maintenance of a heterogeneous fleet at low costs. Also, the increasing adhesion to air transport, strongly driven by the increase in low-cost flights and the wide choice of schedules and destinations, boosted the car rental market, breaking the paradigm of public transport or vehicles for own use (Ribeiro, 2016).

Today, consumers are characterized by the eagerness to want to see their needs satisfied in the right place at a pre-defined time and with a high level of service. (Mahajan, 2015). In order to respond positively to these requirements, companies are faced with the growing need to optimize their logistics activity, so as to improve their service level, reduce the incurred costs and, consequently, increase their profits.

In addition to adapting to changing consumption patterns, companies must also adjust and react to the economic environment in which they find themselves.

According to Associação dos Industriais de Aluguer de Automóveis sem Condutor (ARAC), tourism is currently the largest market for the Portuguese car rental companies, accounting for around 60% of their revenues (Jornal Económico, 2018).

According to Banco de Portugal (2017), tourism revenues in Portugal increased steadily from 2012 to 2016, with an average annual growth of 10.2%. In 2016, revenues reached close to 12.7 billion euros, a significant increase of 10.7% over the previous year.

In 2017, tourism revenues were € 15 billion, an increase of € 3.4 billion over 2016, according to the World Tourism Organization (Observador, 2018). Moreover, in 2018, Portuguese tourism revenues increased again, reaching a new national record of 16.6 billion euros, thus increasing 45% since 2015 (Observador, 2019).

Portuguese revenues derived from tourism from 2012 to 2018 can be seen in Figure

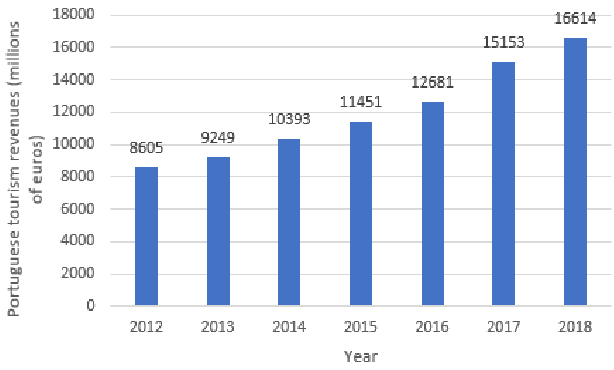

Figure 1: Revenues of Portugal´s tourism sector (Millions of Euros).

In 2017, revenues from tourism were €3.39 billion, more than 16.6% compared to 2016, and was registered a total of 20.6 million guests and 57.5 million of nights, according to Instituto Nacional de Estatística (INE) (Observador, 2018).

The Portuguese rent-a-car market had a sales value of approximately 542 million euros in 2015 and grew approximately 6.5% compared to 2014 (Economic, 2016), while in 2014 there was a billing increase in the order of 13% over the previous year (Fleet Magazine, 2015).

The car rental companies had revenues of around 700 million euros in 2017 and a growth of 45% compared to 2016 (Diário de Notícias, 2018).

Thus, the evolution of this market in terms of revenues from 2014 to 2017 can be seen in Figure 2.

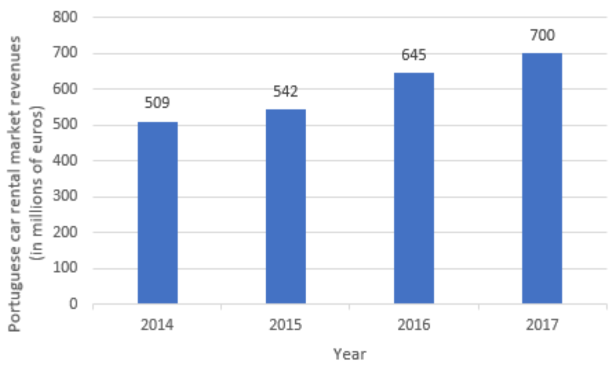

Figure 2: Evolution of revenues registered by the Portuguese rent-a-car market between 2014 and 2017.

It is in this sense that the company Entreposto Logística, which provides automotive logistics services in its various components and has as a major source of revenue car importer brands, whose main costumers are car rental companies, shows interest in an estimate that also takes into account the impact of the car rental market evolution on the demand for the company's logistics activities.

In fact, by taking into account the described context, the company wants to ensure that it manages to maximize, in terms of revenue and profit generated, the strong growth of the Portuguese car rental industry. In this way, it is crucial to define an operational strategy that will allow EL to present a high level of service and increase its competitiveness, without incurring in an excessive increase of costs and, therefore, to maintain or increase its clientele of rent-a-car companies in a way that enables the company to benefit indirectly from the increase in their revenue.

Despite the high growth of this industry, there are few studies conducted in its scope as well as scientific literature associated with its operation and limitations.

2. Related Literature

According to Archer (1994), demand forecast is "the art of predicting the level of demand that may occur in the future or in a given period of time" and is considered to be an essential element of the decision-making process by managers of a particular business or area of activity.

Viegas (2000) describes econometric models that can be used in complex transportation estimation problems when travellers have multiple alternatives to move from one point to another.

The framework of the models to be applied is described in three steps:

1 - Detailed observation and description of reality in order to enable the representation of the phenomena under study.

2 - Development and validation of mathematical models that allow understanding of people behaviour (and other social agents, namely companies) through their adequate capacity to represent past evolutions and scenarios.

3 - Towards alternative scenarios of intervention on the transport system, estimate the changes in people's behaviour regarding their mobility in this new context.

The most frequently used models in this domain are the linear, exponential or logistic growth models (regression models).

Carmona-Benitez et al. (2017) proposes and demonstrates how to use an Econometric Dynamic Model (EDM) to estimate the demand for air passengers in a state context (pax demand) based on economic indicators. Economic indicators determine the pax demand per route at a particular airport, region or city. The most common economic indicators are Gross Domestic Product (GDP), GDP per capita, population (Pop), income, and income per capita (Carmona-Benitez, 2012).

The pax demand forecast is very important, and it can be done in the short, medium and long-term. A short-term forecast is usually carried out for operational planning, while the medium and long-term are used to evaluate large capital investments.

In this paper, the econometric method of the panel data is used to determine the economic variables that significantly explain the pax demand. This method is chosen since there is data over time and by State. A panel data is a combination of time series with cross-sectional units (countries, states, cities or regions, business, etc.) (Cobacho, 2000).

The EDM used was as follows:

$$Pax_{it} = \beta_0 + \beta_1 Pax_{it-1} + \beta_2 ITAEE_{it} + \beta_3 PEA_{it} + \beta_4 INPC_{it} + \beta_5 freq_{it} + \beta_6 TUR_{it} + \beta_7 ME_{it} + u_i + \varepsilon_{it}$$

For i = 1,…, 30 and t = 2006,…, 2014.

Where:

Pax	= Air transport passenger demand	[thousand people]
u	= individual effects	[-]
ITAEE	= Indicator of Economic Activity	[thousand pesos]
PEA	= Indicator of Economically Active Population	[People]
INPC	= National Consumer Price Index	[Current prices]
FREQ	= Total number of flights (national and international)	[flights]
TUR	= Hotel occupancy index of domestic and foreign visitors	[rooms occupied]
ME	= Foreign Exchange earnings from international arrivals	[thousand pesos]

Wang et al. (2016) reports that regarding regional logistics demand forecasting, various quantitative forecasting methods have been developed by scholars in recent years. Firstly, regional logistics demand is the total volume of logistics flow within the activities including transportation, storage, packing, loading and unloading, handling, and distribution processing, and is one of the major expenditures for businesses, thereby affecting and being affected by other economic activities. Secondly, those forecasting methods can be roughly divided into regression analysis method, grey forecasting method, artificial neural network (ANN) and support vector machine (SVM).

Grey system method predicts the change of logistics demand by the generation and development of original data sequence and excavation of inherent rules of the original data. Regression method obtains the change rules through the analysis between explained variable and explanatory variables.

Research showed that the regional economic scale, industrial structure and economic space layout are the significant aspects of regional logistics demand. Therefore, this paper selects the total freight volume as the logistics demand target and constructs the forecast index system as shown in Figure 3.

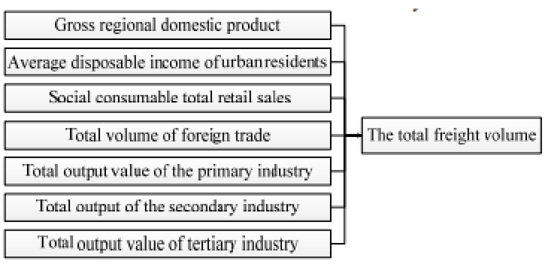

Figure 3: Forecast Index System. Source: Wang et al. (2016)

Tsekeris and Tsekeris (2011) consider that the demand for passenger and freight transport and logistics services is influenced by a multitude of factors. The economy, as reflects the GDP or total added value of a country, and the gross output or added value of a region, affects the general derived demand. The structure of the economy, in terms of the resources, goods, and services, e.g., specialization on particular products, and cultural, trade and tourism services, affects the level of transport demand and modal shares. The supply or upgrading of transport infrastructure and services of better quality (higher frequencies, larger seat capacities, higher speed, increased safety and comfort), as well as the supporting infrastructure (e.g., those of Intelligent Transport Systems or ITS) may result in induced demand.

Transport demand forecasting models can be generally categorized according to the steps involved in the traditional four-stage transport planning process (see Figure 1). These steps include: (1) trip generation, (2) trip distribution, (3) modal split (or mode choice), and (4) traffic assignment. Figure 4 illustrates this process and model.

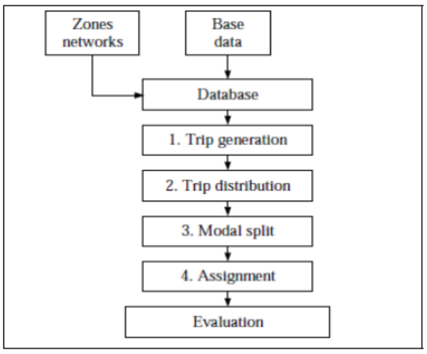

Figure 4: The traditional four-stage transport planning model. Source: Tsekeris and Tsekeris (2011)

Only the first stage (trip generation) is useful for the scope of this work, since it consists in predicting the total logistics demand from one point of origin to another without considering any alternatives and new (variables) modes of transport that may influence the first. Meanwhile, the remaining stages address scenarios that have no association with the objectives of the work and they "run away" from the generic forecasting process as they seek only to distribute the already estimated total logistics demand by different "circuits". Therefore, this first stage will be the only step to be taken in this work.

Trip generation from and attraction to specific origin-destination (or production-attraction) traffic zones in which the study area is partitioned are based on the socio-economic, demographic and land use characteristics of each zone. Most of the trip generation studies employ econometric and, at a lesser extent, time series analysis techniques. The econometric models use linear or log-linear regression analysis to estimate the relationship between transport demand and its determinants.

The time series techniques involve extrapolating historic trends in transport demand into the future without considering the underlining causes of the trends. The most frequently used methods in this approach include (seasonal) exponential smoothing and Box-Jenkins procedure, which can predict the medium- or long-term demand for shipment of goods and movement of passengers.

Rojas et al. (2015) present in their work a selection mechanism of forecast models to contribute to demand estimation in a supply chain.

Due to the existence of unknown and uncontrollable future sceneries, several quantitative and qualitative forecasting techniques appear in order to approximately visualize future events, techniques that become relevant in decision making for companies and organizations.

The selection mechanism of forecast models developed a five steps sequence to make the forecast. Two of them correspond to a basic sequence to forecast as they define the product or service and gather data, however, this work suggests a different third step (analyse data pattern).

This third step consists of a previous data analysis to visualize the time series characteristics; to accurately and precisely select the best forecast techniques.

Although it is recommended, as a prior in intuitive analysis, a graphic representation of the time series evolution in order to analyse its behaviour. However, it is not always possible to correctly conclude about patterns, so, a more sophisticated tool is needed in this case. The tool that this article proposes is to analyse the data pattern with an autocorrelation analysis, which aim is to determine the time series relationship, when it has a time lag of one or more periods, according to the following formula:

$$r_k = \frac{\sum_{t=1}^{n-k}(x_t-\bar{x})(x_{t+k}-\bar{x})}{\sum_{t=1}^{n-k}(x_t-\bar{x})^2} \quad (1)$$

Moreover, in this article it´s considered that t-statistics, associated to the correlation coefficient calculation, allows also judge about the significance of it.

For r_k (lag of k time units), t_{rk} is estimated like r_k/S_{rk}. Being S_{rk} the significance interval for the correlation coefficient that can be built with the standard error calculation and calculated with the following mathematical expression:

$$S_{rk} = \sqrt{\frac{1+2\sum_{i=1}^{n-1} r_i^2}{n}} \quad (2)$$

To significantly consider r_k different from zero, so there is autocorrelation, the value of the statistical t_{rk} must be higher than 1,96 in absolute value.

On the fourth step, which consists in the selection of the forecasting model, a decision tree of forecast methods was developed to illustrate the selection of the best forecasting method, according to the series data pattern. This decision tree helps to pre-select alternative methods to forecast future demand.

Finally, to select the most adequate alternative proposed by the decision tree, a retrospective analysis of data with each pre-selected method must be done. It is necessary to determine a common error measure for all models, and to choose the one that best fits the data, that is, the one that presents the lowest forecast error.

To perform an error analysis the following measures can be used: Medium Squared Error (MSE), Root Medium Squared Error (RMSE) and Medium Absolute Percentage Error (MAPE). In addition, if time and resources are available, it is recommended to use combinations of forecasts with different models according to the data pattern and, analysing the best method to get the lowest error. With this we can get, on average, more precise forecasts than by using individual methods solely.

Finally, in the fifth step, an evaluation of the method(s) used should be carried out by measuring their respective error(s) and monitoring its (their) performance based on the errors obtained in future forecasts.

3. Description of the Problem

The main objective of this work is to forecast (with an adequate statistical foundation) the annual and quarterly volume of vehicle exits from EL´s facilities, while also considering the impact of the performance of the national car rental segment in the company´s activity. Therefore, in order to comply with this objective is necessary to have access to past data regarding the volume of exits from EL (annual and quarterly) and to statistics relative to the Portuguese car rental market, both in an annual and quarterly perspective, and to Portugal´s economy and tourism sector.

3.1. Entreposto Logistica´s Data

This data set, which was made available by the own company, refers to EL´s activity volume from 2010 to 2017 and is expressed in terms of the entry date of vehicles into the company´s facilities and the exit date of vehicles from those same facilities.

Although both the entry and exit volumes of vehicles were made available, EL is only interested in having a forecast of the volume of exits.

This work does not take into account the type of service (Pre-Delivery Inspection (PDI), reconditioning of new vehicles or assembly of extras and preparation of special series) to which the vehicles were subjected, being only relevant the fact that a certain vehicle was subjected to the company´s services and that it consequently left the former. Therefore, we end up considering the total exits of vehicles from EL occurred in a certain period of time.

It is possible to visualize in the Excel file made available by EL, for each vehicle that enters the company: the data related to its trademark, its model, VIN code, date of entry and exit, name of the customer that owns it and transportation request and start of the distribution process dates.

3.1.1. EL´s Annual Volume of Vehicles Exits

The time series of the annual number of vehicles exits from the company from 2010 to 2017 can be observed in Figure 5.

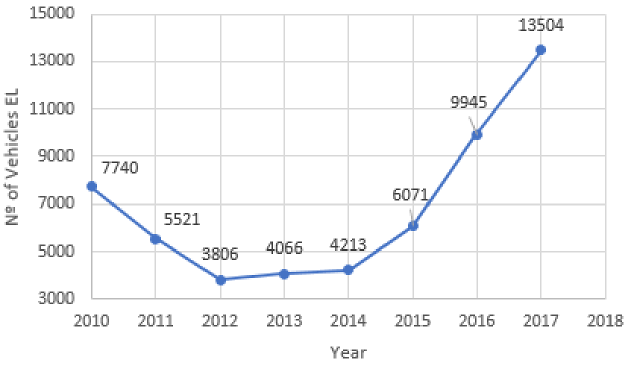

Figure 5 – Annual evolution of EL´s volume of vehicle exits from 2010 to 2017

After analysing Figure 5 it´s possible to conclude that in 2011 and 2012, EL saw a significant decrease in its volume of operations compared to 2010 and 2011, respectively. This situation can supposedly be explained by the Portuguese financial crisis that occurred from 2010 to 2014, which must have had a negative impact on the company´s activity. However, from 2013 onwards it is possible to observe an increasing tendency of the company´s operational volume which, from 2015 onwards is remarkable. It is also important to mention that in 2016 the volume of vehicles exits registered in 2010 (pre-crisis year) was exceeded.

3.1.2. EL´s Monthly Volume of Vehicles Exits

The time series of the monthly number of vehicles exits from EL from 2010 to 2017 (96 months) can be observed in Figure 6.

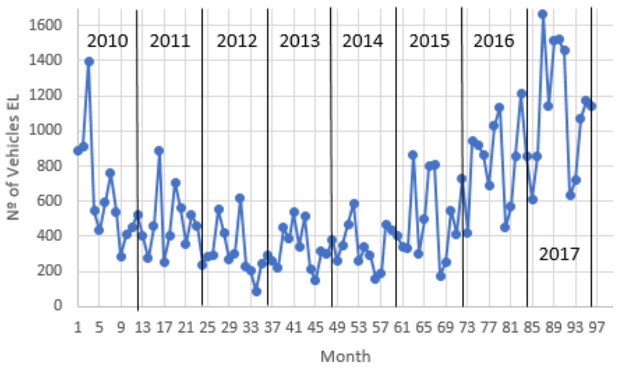

Figure 6 – Monthly evolution of EL´s volume of vehicle exits from 2010 to 2017

From the observation of Figure 6 it is possible to draw a set of conclusions:

• In March 2010, EL registered a very high volume of vehicles exits compared to the observations registered up to 2016, and this value is only surpassed again in 2017 (in March as well).

• In every year from 2010 to 2017, except in 2016, the company assisted in March to an increase in the number of vehicles exits when compared to the months of February and January and recorded one of the highest values in each year.

• In every year from 2010 to 2017, the company recorded a sharp fall in the volume of vehicles exits in August. This scenario is most likely a consequence of the fact that this is a month in which many company employees go on vacation, thus reducing the volume and operational activity of the company.

• In October 2012 EL registered the minimum number of monthly vehicles exits (88) in the sample space under analysis.

• In March 2017 the company registered the maximum monthly value of vehicles exits (1665) in the sample space considered.

• From the first observation (January 2010) to the 34th observation (October 2012), the company seems to register a downward trend in the monthly volume of vehicles exits.

• From the 34th observation onward the company registered an increasing trend in the monthly volume of vehicles exits.

3.1.3. EL´s Quarterly Volume of Vehicles Exits

The time series of the quarterly number of vehicles exits from EL from 2010 to 2017 (32 quarters) can be observed in Figure 7.

Figure 7 – Quarterly evolution of EL´s volume of vehicle exits from 2010 to 2017

After analysing the Figure 7 it´s possible to arrive at the following conclusions:

• Since 2014, EL recorded a quarterly maximum annual volume in the 4th quarter (with the exception of the year 2017 when this occurred in the 2nd quarter).

• Since 2013, this variable always shows the same evolutionary pattern (in the sample space): the second quarter registers a higher number of vehicles exits than the 1st quarter (sometimes quite significantly, as in the year 2017, for example), while also occurring in the 3rd quarter a very sharp decrease of the values observed by the variable in the 2nd quarter and, finally, in the 4th quarter there is a positive and significant variation compared to the 3rd quarter.

• From 2010 to 2012 the variable had a very significant downward trend.

• It is possible to note that from 2013 onwards, the variable presents a very significant upward trend, especially since 2015. It is also possible to visualize that since 2013 the value registered in each quarter increases every year, except for the year 2014, whose 2nd and 3rd quarters presented lower results than the 2nd and 3rd quarters of 2013, respectively.

3.2. Portuguese Rent-a-Car Market Data

ARAC is the only association of employers that covers the entire Portuguese rent-a-car industry, currently having about 150 members (companies) distributed throughout the Portuguese territory.

The data provided by ARAC consists of a set of annual statistics related to the Portuguese car rental market.

These statistics refer to the investment made by the industry, its billing, average fleet, contracts made, days of rental, average daily billing, average billing per vehicle, average rental period and usage rate from 2010 to 2017, inclusive, and are expressed by the type of vehicle in question (light passenger vehicles, light commercial vehicles and heavy commercial vehicles).

3.3. Portuguese Socio-Economic Data

This data set was collected from INE's public documentation and refers to various annual and quarterly statistics that fall within Portugal´s economic and social perspective. The time horizon for these data is from year 2010 to year 2017, inclusive, i.e. 8 years and 32 quarters.

Regarding the economic data, Portugal´s GDP (annual and quarterly), the annual percentage of direct and indirect taxes applied to Portugal´s GDP and the industrial production index (annual and quarterly) were taken into account.

On the other hand, the social statistics used in the work were: tourism receipts generated by foreign tourists (annual), total number of guests (considering both national and foreign residents) (annual) and overnight stays (annual), total number of trips (quarterly), total number of leisure trips (quarterly), total number of trips with a duration of 1 to 3 nights (quarterly), total number of trips lasting 4 or more nights (quarterly), total number of guests in hotel establishments (quarterly), total number of overnight stays in hotel establishments (quarterly), total income of hotel establishments (quarterly), room income of hotel establishments (quarterly) and total number of trips made by tourists using a terrestrial mean of transport (quarterly).

3.4. Data Limitations

Although EL´s data and consequently the other data collected allow to achieve the objectives outlined for this work in a plausible and realistic way, since they cover a significant time horizon (8 years) and a considerable dimension (a total of 59293 vehicles in the case of EL´s data), thus allowing to obtain a credible set of results, it would be preferable to have access to a data set that presented a larger horizon/time span in order to be able to model it better and consequently obtain more precise results and models.

Therefore, it would be more appropriate to build both causal (regression analysis based) and non-causal (time series based) forecasting models with a greater number of observations of the variable to be predicted, i.e. EL´s volume of vehicles exits, in terms of an annual (8 observations) and quarterly (32 observations) horizon.

On the other hand, since the data made available by ARAC was only expressed in an annual time span, it was not possible to consider any Portuguese car rental market statistics in the quarterly forecasting models, which means that these were only used in the annual forecasting models.

4. Case Study

This chapter describes the application of causal and non-causal forecasting models to the set of data presented.

4.1 Non-causal Forecasting Models

The non-causal forecasting models will only be developed from the time series of the EL´s quarterly volume of vehicles exits.
The time series of the variable EL´s monthly volume of vehicles exits will not be used for this analysis, since as can be observed in Figure 6, this series does not show any type of stationarity, tendency or seasonality, being random and eminently unpredictable. Therefore, the random nature of this time series makes it quite difficult to model.
It will also not be considered in this analysis the time series of the EL´s annual volume of vehicle exits, because it presents a small number of observations, which does not allow to proceed with an appropriate and statistically sustained modelling of the same.
On the other hand, from the analysis of Figure 7 we can infer that the time series of the EL´s quarterly volume of vehicles exits shows a decreasing trend from the year 2010 to 2012, inclusive. However, since 2013 this time series started to present different characteristics, showing seasonal characteristics and a growing trend, as it was already mentioned.
We can somehow perceive that the characteristics evidenced by this time series between 2010 and 2013 were a direct reflection of the Portuguese economic scenario lived in that years (Portuguese Crisis) and that does not depict the current situation of the Portuguese economy and of EL.
Thus, only the time series of this variable since the year 2013, as represented in Figure 8, will be taken into account in this non-causal forecasting approach.

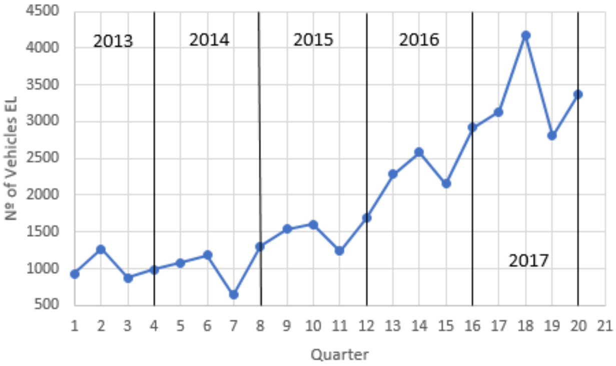

Figure 8 – Quarterly evolution of EL´s volume of vehicle exits from 2013 to 2017

Given the characteristics of the time series under study and taking into account the decision tree of forecast methods developed by Rojas et al. (2015), the proposed solution approaches were the Holt-Winters´s additive and multiplicative models and the additive and multiplicative time series decomposition models, assuming a seasonal cycle duration equal to 4 (s = 4).

4.1.1 Holt-Winters´s Additive Method

In Figure 9 it is possible to visualize the application of the developed model and how it compares with the original time series under study.

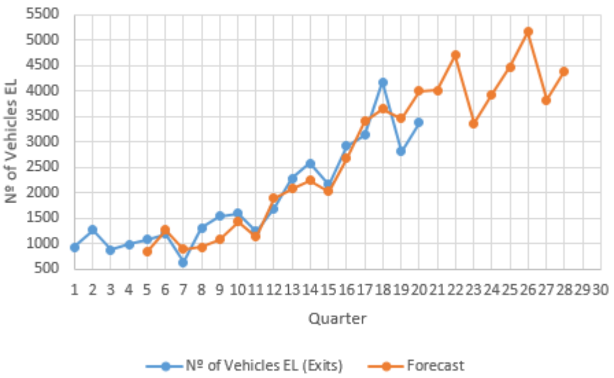

Figure 9 – Holt-Winters´s additive model

Table 1 shows the values of the parameters α, β and γ used (optimal) and of the error measure MSE.

Table 1: Values of the Holt-Winters´s additive method smoothing constants and MSE

α	β	γ	MSE
0.198	1	0.837	81307.2

4.1.2 Holt-Winters´s Multiplicative Method

In Figure 10 it is possible to visualize the application of this model and the way it compares with the original time series.

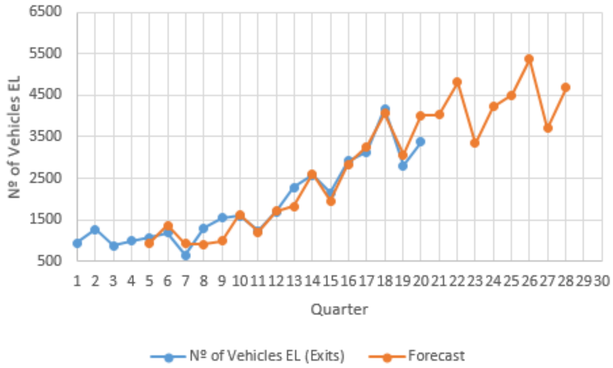

Figure 10 – Holt-Winters´s multiplicative model

Table 2 shows the values of the parameters α, β and γ used (optimal) and of the error measure MSE.

Table 2: Values of the Holt-Winters´s multiplicative method smoothing constants and MSE

α	β	γ	MSE
0.2185	1	0.2084	65830.99

4.1.3 Additive Time Series Decomposition Method

In Figure 11 it is possible to visualize the application of the model and how it compares with the original time series.

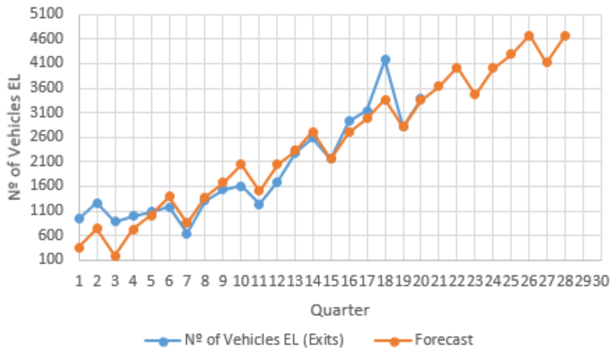

Figure 11 – Additive time series decomposition model

Table 3 shows the values of the seasonal indices used and of the error measure MSE.

Table 3: Values of the additive time series decomposition method seasonal indices and MSE

S'_1	S'_2	S'_3	S'_4	MSE
99.546	317.64	-395.49	-21.7	120535.6

4.1.4 Multiplicative Time Series Decomposition Method

In Figure 12 it is possible to visualize the application of the model and how it compares with the original time series.

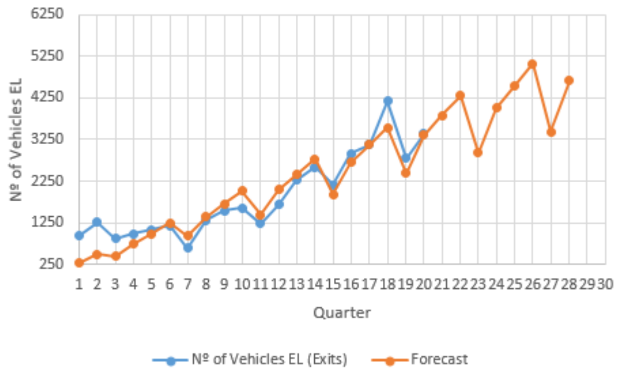

Figure 12 – Multiplicative time series decomposition model

Table 4 shows the values of the seasonal indices used and of the error measure MSE.

Table 4: Values of the multiplicative time series decomposition method seasonal indices and MSE

S'_1	S'_2	S'_3	S'_4	MSE
1.0835	1.1626	0.76	0.9938	122323.4

4.1.5. Conclusion

The four models developed were also subjected to some statistical analysis procedures in order to test if their residuals: are normally distributed, homoscedastic (constant variance), uncorrelated (independent) and have zero mean.

Table 5 shows the performance of the different models with regard to the residual analysis tests considered.

Table 5: Results of the residuals analysis tests

Model|Test	Normal.	Homos.	Zero Aver.	Independ.
Holt-Wint. Addit.	✓	✓	×	✓
Holt-Wint. Mult.	✓	✓	×	✓
Addit. Decomp.	✓	×	×	×
Mult. Decomp.	✓	×	×	×

As can be concluded from the previous sections, the model that best predicts and forecast the original series of the EL´s quarterly volume of vehicles exits is the Holt-Winters´s multiplicative method (higher MSE). This model, together with the Holt-Winters´s additive model, are the ones that presented the best overall performance in the residuals analysis tests, satisfying 3 of the 4 assumptions considered, as can be observed in Table 5.

Therefore, the Holt-Winters´s multiplicative method seems to be the most adequate and reliable model (of the 4 considered) in terms of modelling the time series under analysis.

As can be seen in Figure 8, there appear to be 2 outliers in the entire sample of observations (20), these being the third quarter of 2014 and the second quarter of 2017.

In the third quarter of 2014, there was an excessive decrease in the number of vehicles exits from EL, given the extent of the "fall" observed in the third quarter of the remaining years. On the other hand, in the second quarter of 2017 there is an excessive growth of the variable under study, compared to the increase registered in the second quarter of the remaining years.

Thus, although the results obtained by the 4 models were satisfactory, it is important to note that these 2 outliers contribute negatively to their MSE.

4.2 Causal Forecasting Models

The linear regression models (multiple and simple) to be developed have either a yearly or a quarterly focus. The annual models will be developed from the annual data described in section 3. On the other hand, the quarterly models will be based on the quarterly data presented in section 3.

Both types of models will be developed as dynamic econometric models (as introduced by Carmona-Benitez et al. (2017) in section 2), thus presenting a time lag between the instants of the observations of the response variable and the explanatory variables in question. The use of EDMs is justified by the fact that this type of models facilitate the forecasting activity of a company, in this case EL, since in order to estimate the response variable for a period t, it is sufficient to have access to the values of the explanatory variables in the instants t_i, and the parameter i can present several values according to the values defined for the model in question.

In the annual linear regression models, a value of $i = 1$ was used for all independent variables, since the total number of available observations for the dependent variable and the formers is not very high (8 observations) and it was considered to be important to have a sample of observations with a considerable size, and also because the Pearson correlation coefficients obtained between the dependent and independent variables for this degree of temporal mismatch were quite high. On the other hand, in the quarterly linear regression models, the value of i used for each explanatory variable was the one that maximized the Pearson correlation coefficient between the former and the dependent variable. The considered values for i were the following: 1, 2, 3 and 4.

In order to obtain the best possible models, the variable selection methods used were the following: stepwise, backward and forward.

4.2 Quarterly Models

On Table 6 it´s possible to compare the 2 quarterly models obtained/developed.

Table 6: Quarterly causal forecasting models

Ind.var.	Regression Coefficient	P-Value	R_a^2	MSE	Name
Intercep.	-14723.02	0	0.898	81543.25	Model 1
Nº of trips 1-3 nights	0.0005214	0.005			
Room inc. hotel. estab.	0.000827	0.02			
GDP	0.3258	0			
Intercep.	-12001.61	0	0.913	69606.23	Model 2
GDP	0.2859	0			
Nº of trips 1-3 nights	0.000742	0.001			
Nº of guests hotel estab.	-0.000837	0.024			
Terr. transp. tourism	0.000367	0.035			
Total inc. hotel estab.	0.006288	0.01			

On Figure 13 and Figure 14 it is possible to visualize the application of the Models 1 and 2 and how they compare with the original time series, respectively.

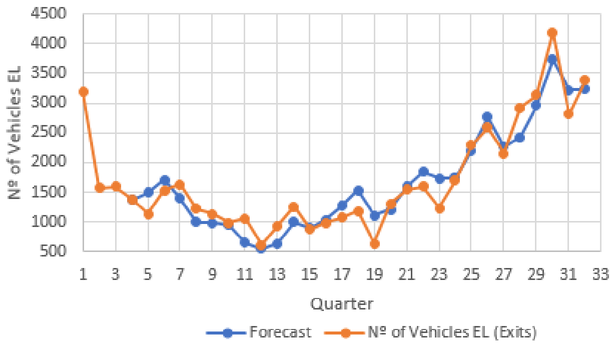

Figure 13 – Quarterly Model 1

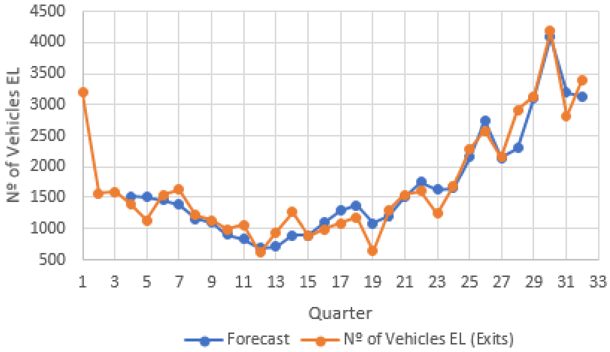

Figure 14 – Quarterly Model 2

Both models were also subjected to several statistical tests in order to ascertain if they comply with the following linear regression assumptions: linear relationship between the dependent and independent variables, no or little multicollinearity, residuals with zero average, homoscedasticity, no auto-correlation and normally distributed residuals. Table 7 shows the performance of the two models with regard to the statistical tests considered.

Table 7: Results of the statistical tests for the quarterly models

Test|Model	Model 1	Model 2
Linearity	✓	✓
No multicollinearity	✓	×
Zero average	✓	✓
Homoscedasticity	✓	✓
No auto-correlation	✓	✓
Normality	✓	✓

From the analysis of Table 7 we can conclude that Model 1 seems to satisfy all the assumptions analysed for the linear regression models, while Model 2 appears not to satisfy only the assumption of independence of the explanatory variables.

Regarding the performance in the tests of significance to which they were subjected, both models translated as being globally significant. Regarding the individual significance tests, all the coefficients of Models 1 and 2 were concluded to be significant.

Also taking into account the fact that both models have a good explanatory capacity for the quarterly volume of vehicles exits from EL (89.8% and 91.3%, respectively), both are statistically supported, especially Model 1, to be used for forecasting purposes.

4.2 Annual Models

On Table 8 it´s possible to compare the 6 annual models obtained/developed.

Table 8: Annual causal forecasting models

Ind.var.	Regression Coefficient	P-Value	R_a^2	MSE	Name
Intercep.	– 21839	0.002	0.909	1209018	Model 1
Average daily billing	193.8613	0.0005			
Intercep.	-2664.363	0.3802	0.9630	493436	Model 2
Total nº of Guests	1396.876	0.0003			
Direct Taxes	-1330.872	0.015			
Intercep.	-84103.63	0.005	0.982	232138	Model 3
Overnight stays	-1720.755	0.067			
GDP	0.4917	0.009			
Tourism receipts	8.561768	0.045			
Intercep	-4839.32	0.1443	0.742	3446312	Model 4
Investment	19.56094	0.0079			
Intercep.	-27945	0.0017	0.953	633237	Model 5
Average daily billing	144.895	0.0062			
Average fleet	244.1714	0.078			
Intercep.	-56387	0.005	0.952	635188	Model 6
GDP	0.287	0.01			
Tourism receipts	1.286	0.003			

On Figures 15, 16, 17, 18, 19 and 20 it is possible to visualize the application of the Models 1, 2, 3, 4, 5 and 6 and how they compare with the original time series, respectively.

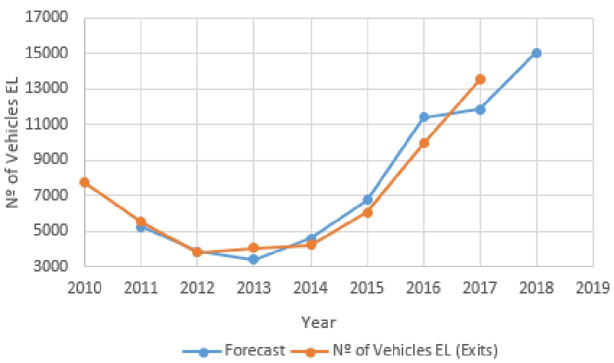

Figure 15 – Annual Model 1

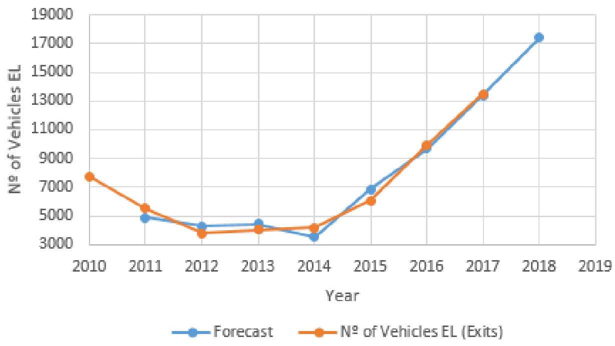

Figure 16 – Annual Model 2

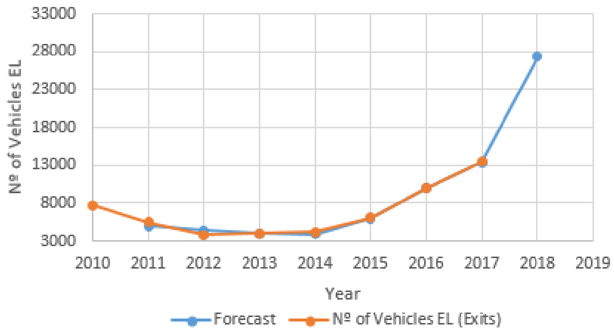

Figure 17 – Annual Model 3

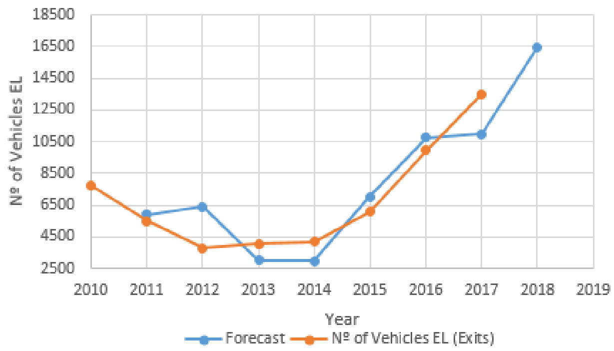

Figure 18 – Annual Model 4

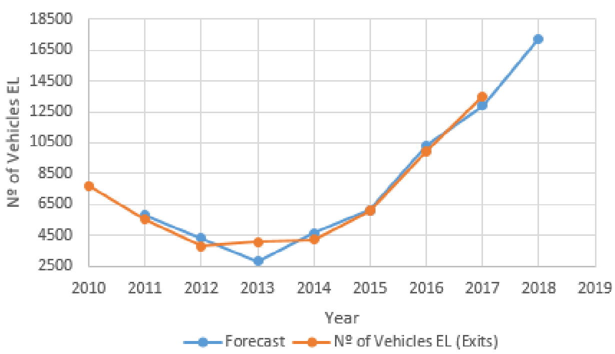

Figure 19 – Annual Model 5

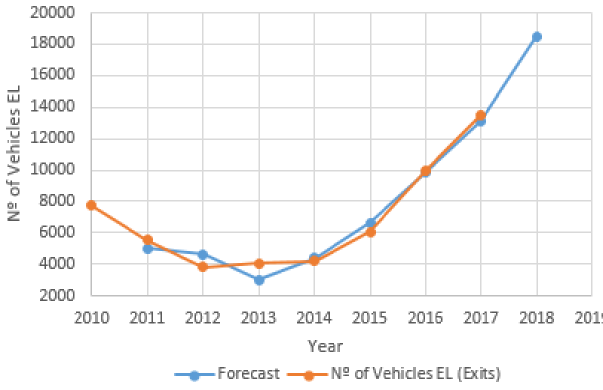

Figure 20 – Annual Model 6

All models were subjected to several statistical tests in order to conclude if they comply with the following linear regression assumptions: linear relationship between the dependent and independent variables, no or little multicollinearity, residuals with zero average, homoscedasticity, no auto-correlation and normally distributed residuals. Table 9 shows the performance of the six models regarding the statistical tests considered.

Table 9: Results of the statistical tests for the annual models

Test|Model	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Linearity	✓	✓	✓	✓	✓	✓
No multicollineari.	✓	✓	×	✓	✓	✓
Zero average	✓	✓	✓	✓	✓	✓
Homoscedasti.	×	✓	✓	✓	✓	✓
No auto-correlation	✓	✓	✓	✓	✓	✓
Normality	✓	✓	✓	✓	✓	✓

As can be concluded from the analysis of Table 9, only Models 1 and 3 do not satisfy all the assumptions considered.

Regarding the performance in the tests of significance to which they were subjected, all models translated as being globally significant. Regarding the individual significance test, all the coefficients of the models were found to be significant, at least for a 90% confidence interval.

Although the 6 models have a high explanatory capacity of the annual volume of vehicles exits from EL, Models 1 and 4 are not as good in this aspect as the other models, especially Model 4.

Taking this into account, we can assume that Models 2, 3, 5 and 6 seem to be the most adequate to be used for forecasting the time series under analysis, especially Models 2 and 3.

Although Model 3 presents a significantly lower MSE value than the other models, since it does not satisfy the assumption of independence of the explanatory variables and appears to present a very "optimistic" estimate for the year 2018 and, in principle , with regard to the near future, it is considered that Model 2, which presents a not so much higher MSE value, is the most adequate to model the series of the annual volume of vehicles exits from EL.

5. Final remarks and recommendations

The objective of this work arose from the need felt by EL to improve its forecasting procedure and also to account for the impact of the Portuguese rent-a-car industry, which, according to the experience of the company, has a strong influence on its volume of vehicles exits.

In order to achieve this objective, the present work was developed over several phases. In a first phase, a literature review and an attempt to identify a set of forecasting models that could be applied to the case study were made. From this analysis it was concluded that the best resolution approaches would be the linear regression models (causal forecasting models), the Holt-Winters´s additive and multiplicative models, and the additive and multiplicative time series decomposition methods (non-causal forecasting models).

Subsequently, numerous statistics that could be significantly correlated with the variable to be predicted (annual and quarterly volume of vehicles exits from EL) were considered, these being related to the rent-a-car industry and the Portuguese socio-economic aspect.

Having defined the variables to be used in this work, quantitative forecasting models (causal and non-causal) were developed and implemented in order to identify the type of model that was the most adequate in forecasting the volume of exits of vehicles from EL (quarterly and annual). Therefore, a number of statistical tests were carried out, which allowed a continuous validation of the results implementation strategy.

Finally, the performance results, in terms of the value of MSE obtained, of the various models developed (both annual and quarterly) was compared, as can be seen in Figure 21.

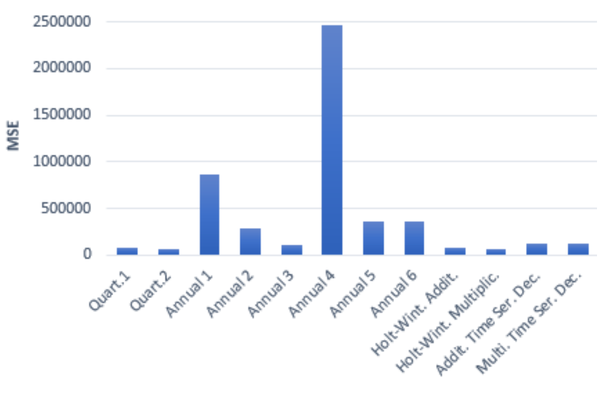

Figure 21 – MSE value of the forecasting models developed

7In addition to the analysis of the MSE values obtained, the performance of the twelve models in the statistical tests to which they were subjected was also taken into account in the selection process of the most suitable quarterly and annual models.

Regarding an annual approach, as already mentioned, it was concluded that Model 2 proved to be the most adequate to model the series of the annual volume of vehicles exits from EL.

On the other hand, in a quarterly perspective the multiplicative Holt-Winters model was the one that demonstrated to be the most precise (smaller MSE), satisfying all the tests of residual analysis considered. Therefore, this model proved to be the most suitable to be used for forecasting purposes.

Thus, the models used and reviewed in the literature turned out to be as adequate as expected, leading to reduced MSE values.

Taking into account the results obtained, the multiplicative Holt-Winters model and Model 2 show to be the most suitable ones to be implemented in EL's management system, the first one from a quarterly planning perspective and the last one from an annual standpoint.

5.1. Conclusions and future work

As future developments of the present work, different aspects can be identified.

A first one relates to the tools used throughout the work. Non-causal forecasting models were developed solely from the use of Excel tools. However, due to the complexity and volume of the data being analysed, it would be interesting and advisable to use other computer tools, which would allow a faster and simpler analysis of data and the analysis of a larger data set and, consequently, a greater validation of the results obtained.

However, this aspect does not apply to the development process used in the causal forecasting models, since these were developed, in addition to Excel, from the statistical software's SPSS and NCSS, which are specialized in this subject, therefore allowing to undergo on a more extensive analysis.

In addition, it would be advisable to have access to a longer time horizon, and consequently to a greater number of observations of EL´s data and of the remaining variables, in order to guarantee a greater reliability of the obtained models.

Finally, although the models identified in the literature showed an interesting performance and results, a more in-depth study should be carried out, where scientific developments on the models identified as more promising could be explored.

References

[1] Archer, B. (1994). Demand Forecasting and Estimation, in Travel, Tourism and Hospitality Research. Second edition, John Wiley & Sons, Inc. New York.
[2] Banco de Portugal (2017). Boletim Estatístico – Maio de 2017. Obtained from: https://www.bportugal.pt/sites/default/files/anexos/bemai17.pdf.
[3] Carmona-Benitez, R.B. (2012). "The Design of a Large Scale Airline Network". Thesis. Delft University of Technology, Delft.
[4] Carmona-Benitez, R.B., Nieto, M.R. and Miranda, D. (2017). An Econometric Dynamic Model to estimate passenger demand for air transport industry. Transportation Research Procedia. 25: 17–29.
[5] Cobacho, M.B. (2000). Contrastes de hipótesis en datos de panel, XIII Jornadas de ASEPUMA. 20: 2-7.
[6] Diário de Notícas (2018). Rent a car vale 700 milhões de euros à boleia dos turistas. Obtained from: https://www.dn.pt/dinheiro/interior/rent-a-car-vale-700-milhoes-de-euros-a-boleia-dos-turistas-9116224.html.
[7] Fleet Magazine (2015). Rent-a-Car Portugal. Obtained from: http://fleetmagazine.pt/2015/06/24/rent-a-car-portugal/.
[8] Jornal Económico (2018). Rent-a-Car Contraria Setor Automóvel e Entra no Novo Ano a Adquirir Mais Viaturas. Obtained from: http://www.jornaleconomico.sapo.pt/noticias/rent-a-car-contraria-setor-automovel-e-entra-no-novo-ano-a-adquirir-mais-viaturas-270396.
[9] Mahajan, M. (2015). Consumerism: A Globalization Concept. International Journal of Multidisciplinary Research and Development. 9(2): 245-248.
[10] Observador (2018). Portugal com 15 mil milhões de euros de receitas turísticas em 2017. Obtained from: https://observador.pt/2018/08/28/portugal-com-15-mil-milhoes-de-euros-de-receitas-turisticas-em-2017/.
[11] Observador (2019). Receitas turísticas atingem novo recorde de 166 mil milhões de euros em 2018. Obtained from: https://observador.pt/2019/02/20/receitas-turisticas-atingem-novo-recorde-de-166-mil-milhoes-de-euros-em-2018/.
[12] Ribeiro, A.G.M.O. (2016). Algoritmos heurísticos para a afetação de veículos a reservas em empresas de "rent-a-car". Tese de Mestrado em Engenharia Eletrotécnica e de Computadores. Faculdade de Engenharia da Universidade do Porto, Porto, Portugal.
[13] Rojas, J.P.S., Rojas, F., González, H.V. e Martín, M.S. (2015). Forecasting Models Selection Mechanism for Supply Chain Demand Estimation. 55: 1060-1068.
[14] Tsekeris, T. e Tsekeris, C. (2011). Demand Forecasting in Transport: Overview and Modeling Advances. Economic Research - Ekonomska Istraživanja, 24(1): 82-94.
[15] Viegas, J.M. (2000). A Utilização de Modelos Matemáticos para a Estimação da Procura de Transportes.
[16] Wang, L., Yang, Z. e Xue, C. (2016). Regional Logistics Demand Forecasting: A Method Based on Improved Grey Neural Network. 13.1-13.6.