International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2013): 6.14 | Impact Factor (2014): 5.611

Logistic Growth and Statistical Forecasting Models of Tourist Arrivals in Albania

Valentina Shehu[1], Oltiana Toshkollari[2]

[1]Department of Mathematics, Faculty of Natural Sciences, Tirana University, Tirana, Albania

[2]Department of Computer Science , Faculty of Information Technology, "Aleksander Moisiu" University of Durrës, Durrës, Albania

Abstract: ***Careful prediction of tourist arrivals is a principal factor in arrangement and administration of touristic activities. In the first part we use logistic growth to model and forecast the number of tourist arrivals. Following, an effort has been made to forecast tourists' arrivals applying statistical time series modeling techniques. The results of Albania Tourist Area Life Cycle indicated that the midpoint of Life Cycle is year 2010, the duration of growth time is 13.6 years and the Carrying Capacity is 4,886,858 tourists. The time series analysis showed that the best model was ARIMA(2,1,2), and according to this model the number of tourists is expected to be 3,101,692 for year 2015. The Logistic differential equation, Exponential Smoothing techniques and ARIMA models all helped in the development of a better understanding the overall picture of the tourism story suggesting. This information is particularly useful for new investors and developers as well as longer term planners and government institutions***.

Keywords: Tourism forecast, Logistic model, time series analysis, Albania.

1. Introduction

Tourism to Albania has shown a significant growth during 25 years of democracy. Nowadays, tourism is a promoter of development of economy and an important source of foreign exchange earnings. This primarily depends on the number of tourist arrivals of various groups of tourists. The prediction of tourists‟ arrivals is significant since it would enable the tourism related industries like airlines, hotels and other support business to adequately prepare for any number of tourists at any future date. It is obvious that a lot of direct and sustaining activities, such as the hospitality structures, the transport network, tourist services, catering business are the principal taking advantages of the arrivals of tourists.

If Albania is efficient to make use of its natural beauty for the benefit of tourism than it actually occurs that could bring more profits. Nowadays, Albania is gaining the deserved position, besides in the political and economic status but furthermore as a touristic target. Tourism possesses exceptional opportunity as an accelerator for economic development and is moreover regarded as a crucial field in a global position.

If the tourism business is or is not efficient depends on the balance of the supply and demand of tourism. The forecast of tourist number becomes permanently imperative and has a leading voice in the study of tourism market. The tourism in Albania has increased more and more quickly, accompanied by a high increase of the tourism demand. Dealing such a powerful market, it is necessary to exploit new tourism assets, so a measure of possible destinations is imperative.

It is of interest to the efficient management; a touring schedule which can lead correctly the making use of touristic resources and be established easily. In a touring scheme, the prevision of number of tourist is the crucial moment.

In the first part of the paper we try to forecast the tourist number in the obvious prospective, by means of interlocking the logistic growth S-shaped curve with the environmental saturation level (carrying capacity). The estimation of parameters of logistic S-shaped curve: rate of change and carrying capacity is the crucial moment in the following paper. It has the great priority in improving forecast prevision to apply the logistic growth curve to adapt the evolving tendency of the touristic market and to get forecast.
In the second part of the paper, an effort has been made to make prediction of tourists‟ arrivals using statistical time series modeling techniques such as: Moving Average, Simple Exponential Smoothing, Holt-Winters smoothing exponential and Auto-regressive Integrated Moving Average.

2. Material and Methods

2.1 The Logistic Growth Model (Verhlust Model)

The rule of natural growth during a period of time can be specified by stages of birth, growth, maturity, decline and death for any system. This structure of phases is often named the life cycle of system.

A bell-shaped curve is generally used as a sample to delineate the speed of growth within a period of time, while cumulative growth (amount of units until each point of time) tracks an S-curve. So the S-curve is regarded as the image of cumulative growth. One of the mathematical functions that generate an S-curve is called a logistic.

Origin of the logistic function used as mathematical model of population growth has been firstly introduced by Belgian mathematician Pierre-Francoic Verhulst (1804-1849) in 1983. The logistic equation was a pattern of growth under restriction of a population. Nowadays this equation is often applied to model the evolution of systems socio-economic. The essence of this function is" the rate of growth is proportional to both the number of growth completed and the number of growth remained to complete. Understanding this essence facilitates to answer the question: “Why does the logistic equation possess forecasting power”

The conviction is that forecasting power of S-shaped logistic curve induced due to the fundamental notion of limited resources placed at the base of any growth process. In varying situations, limited resources are called in varying ways: restricted resources (economy), competition (technology), limited environment (tourism) etc.

Applying S-shaped logistic curve, Butler (1980) bring the idea of Tourism Area Life Cycle (TALC). The evolution of a touristic area exhibits an S-curve during its lifetime. That is the growth rate is proportional to the number of tourist arrivals at the time t, T(t) and the number of other people which may visit the destination:

$$\dot{T}(t) = rT(t)[CC - T(t)]$$

where T(t) is number of tourist at time *t*, *r* is the characteristic rate of growth of touristic area, CC is Carrying Capacity or maximum tourist capacity of tourist destination and $\dot{T}$ the derivation of tourists number with respect to time. The particular solution of this differential equation is logistic function:

$$T(t) = \frac{CC}{1 + e^{-(at+b)}}$$

This function generates the well known S-shaped curve. Three parameters *a*, *b*, and CC wholly specify the curve. In many cases, it is useful taking other parameters, which describe better the curve: t_m is the time, point of inflection of S- curve at which 50% of saturation level is achieved; Δt is the characteristic time of logistic S-curve, the interval of time when T(t) takes values from 10% to 90% of CC. Three parameters CC, t_m, and Δt determine another kind of logistic function:

$$T(t) = \frac{CC}{1 + e^{-\frac{\ln 81}{\Delta t}(t - t_m)}}$$

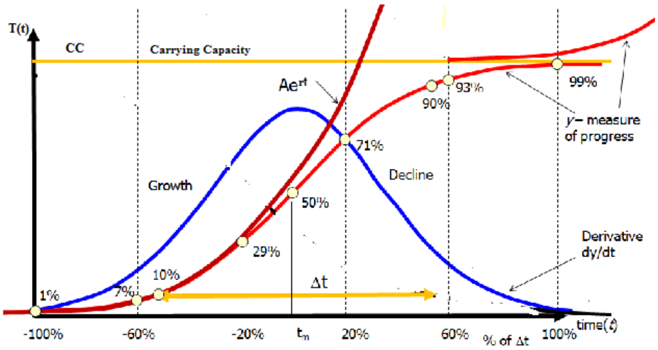

Figure 1: The characteristic parameters of logistic growth curve

The evolution of tourist area‟s stages including: Exploration, Involvement, Development, Consolidation and Stagnation over time (Figure 2). The first stage of exploration tourism is not recognized as an economic activity in that only a few people travel to the destination. The involvement stage is a time period in which tourist numbers increase mainly due to an increased awareness of the destination as a tourist base. By the start of the development stage, the destination‟s tourism facilities in both public and private areas become well developed while in the fourth consolidation stage, the tourists number continue to increase but the destination now becomes well known and visited by many; thus not listed now as a priority for potential tourists in that the rate of growth of the tourist numbers gradually declines until finally, a stagnation stage is reached when potentially all tourists know the destination well including the facilities on offer. In addition, Butler (1980) argued two possibilities after stagnation stage is reached, namely; rejuvenation or decline.

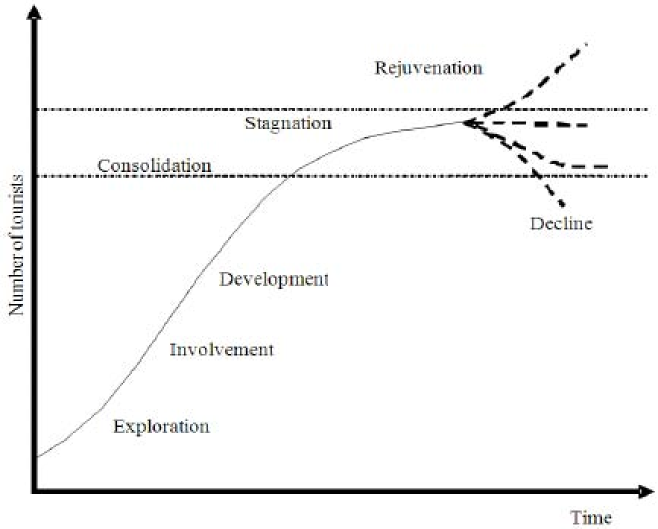

Figure 2: The logistic growth curve of Tourist Area Life Cycle (Adapted from Nejad and Tularam, 2010)

2.2 Determination of a Technique Appropriate For Smoothing the Data

If the data have a visible trend, it is necessary to use the techniques of data smoothing. The techniques used are: Moving average, Simple Exponential Smoothing, Double Exponential Smoothing, (Shumway and Stoer, 2006).

Moving Average

The model which predicts the value of Y at time *t*+1 based on the data up to time t is:

$$\hat{Y}_{t+1} = \frac{Y_t + Y_{t-1} + ...Y_{t-m+1}}{m}$$

where m represents the number of recent observations. Each of m recent observations has weight 1/m on average formula. When m increases, each recent observation takes little weight. The value of m is selected such that Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAEP) to be minimum.

Simple Exponential Smoothing

In the given series $\{Y_t\}$ is a sequence of data. Forecasted value for Y_{t+1} given by F_{t+1}, which is the output, is taken by exponential smoothing algorithm. The estimated value is based on the smoothing constant α, 0 < α < 1. The model is given by:

$$F_1 = Y_0$$
$$F_{t+1} = \alpha Y_{t-1} + (1-\alpha)F_{t-1}, t > 1$$

The best value of α is the value that corresponds to the minimum Mean Square Error (MSE), Mean Absolute Percentage Error (MAPE) and Root Mean Squared Error

(RMSE). Closer to 1 to be α, the greater is the impact of observations which are away from the base forecast.

Double Exponential Smoothing

The model is given by:

The prediction equation: $F_{t+h} = L_t + hb_t$

Level equation: $L_t = \alpha Y_t + (1-\alpha)(L_{t-1} + b_{t-1})$

Trend equation: $b_t = \beta(L_t - L_{t-1}) + (1-\beta)b_{t-1}$

and α, β are smoothing coefficient corresponding to L_t and b_t. (α, β = 0.1,0.2,…..0.9). Initial values mainly are taken: $L_0 = Y_1$ or $L_0 = Y_2 - Y_1$ and $b_0 = 0$ or $b_0 = (Y_n - Y_1)/(n-1)$

The values of α and β are obtained from minimum MSE.

Auto-Regressive Integrated Moving Average

The model ARIMA(p,d,q) is a discrete linear equation with noise:

$$(1-\sum_{k=1}^{p}\alpha_k L^k)(1-L)^d X_t = (1+\sum_{k=1}^{q}\beta_k L^k)\varepsilon_t$$

The parameter p indicates the number of autoregressive terms, d indicates the seasonal differences (the operator of differentiation) and shows the order of integration, and q is the number of moving-average terms. ARIMA model is a special case of ARMA model, but with a special structure.

Let take $Y_t = (1+L)^d X_t$ then Y_t is a ARMA(p,q) model. ARIMA(0,1,0) is given by: $X_t = X_{t-1} + \varepsilon_t$, ARIMA(0,2,2) is given by: $X_t = X_{t-1} + X_{t-2} + (\alpha+\beta-2)\varepsilon_{t-1} + (1-\alpha)\varepsilon_{t-2} + \varepsilon_t$.

The best model is the model that has the smaller RMSE and MAPE. ARIMA model is given by:

$$(1-\sum_{k=1}^{p}\alpha_k L^k)Y_t = (1+\sum_{k=1}^{q}\beta_k L^k)\varepsilon_t$$

where X_t is gained by Y_t as they are carried d successive integration.

The data about the number of tourists in Albania, for the period 2000-2014, are taken from the database of Albanian Institute of Statistics. According to Song and Li (2008), tourist arrivals or tourist population at a particular time is the most common variable to measure tourism demand and usually given by the total number of tourist arrivals from an origin to a destination. The Loglet Lab software is used to obtain the parameters of the logistic model and the predicted values from logistic function and R software for the time series analysis.

3. Results

The data in table 1, clearly shows that the number of tourists increased significantly, mainly after 2009. From 2000 to 2007, the number of tourists has increased by about 255%, or about 4 times more than in 2000. From the year 2007 to 2010 we have an increase of about 114%, or about 2.1 times more tourists than in 2007. From 2010 to 2014 tourists have increased by about 52%, or about 1.5 times more tourists than in 2010. In total, from 2000 to 2014 the number of tourists is increased approximately 11 times. The data provided by INSTAT shows that in year 2014, around 93% of the total number of tourists came from Europe.

Table 1: The Number of Tourists in Albania from 2000 to 2014

Year	Tourist Arrivals	Year	Tourist Arrivals	Year	Tourist Arrivals
2000	317,149	2005	747,837	2010	2,417,337
2001	354,000	2006	937,038	2011	2,932,132
2002	470,574	2007	1,126,514	2012	3,513,666
2003	558,057	2008	1,419,191	2013	3,255,988
2004	645,409	2009	1,855,634	2014	3,672,591

3.1 Logistic Growth Model

Using Loglet Lab software, logistic fits are generated for foreign arrivals of tourists in Albania. The number of tourist arrivals in Albania is modeled using the logistic growth as shown in figure 3:

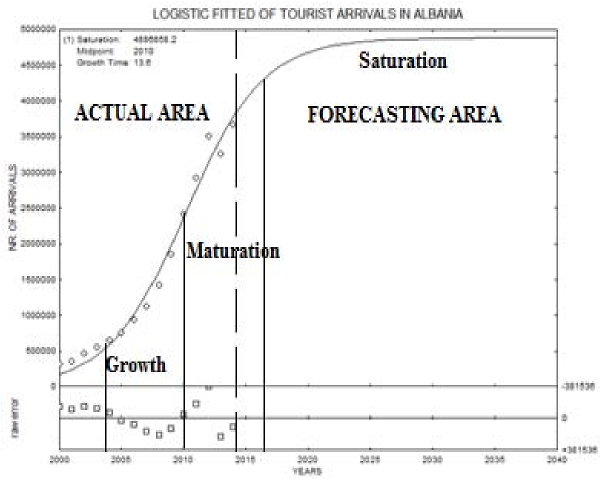

Figure 3: The Logistic Growth Curve of Tourist Arrivals In Albania

From figure 3, the substantial data about Albania Tourist Area Life Cycle may be obtained. The midpoint of Life Cycle is 2010; the duration of growth time is 13.6 years; the Carrying Capacity is 4,886,858. The increase of growth rate of tourist arrivals in 2010, can mark that the logistic growth passes the inflection point. The growth stage is the period 2003 – 2010 and the maturity stage is from 2010 to 2017.

3.2 Time Series Analysis

The results obtained by the methods of exponential smoothing are shown in the table 2.

Table 2: Comparison of Exponential Smoothing Methods

Method	RMSE	MAPE
Simple Exponential Smoothing	348717.3	26.77959
Holt‟s Exponential Smoothing	248052.1	13.22077
Exponential smoothing with “dampening” of the trend	235656.4	14.74788
Exponential Smoothing with Multiplicative Error	310022.5	18.72373

The best model based on the values of RMSE and MAPE is Holt‟s Exponential Smoothing method. The model is given by:

$$L_t = 0.94Y_t + 0.06(L_{t-1} + b_{t-1})$$
$$b_t = 0.34(L_t - L_{t-1}) + 0.66b_{t-1}$$
$$F_{t+m} = L_t + b_t m$$

Using the Holt‟s method, the number of tourists is expected to be 3,952,015 for year 2015, 4,238,527 for year 2016 and 4,525,039 for year 2017. The figure 4 indicates the actual and the predicted number of tourists for the period 2000-2017.

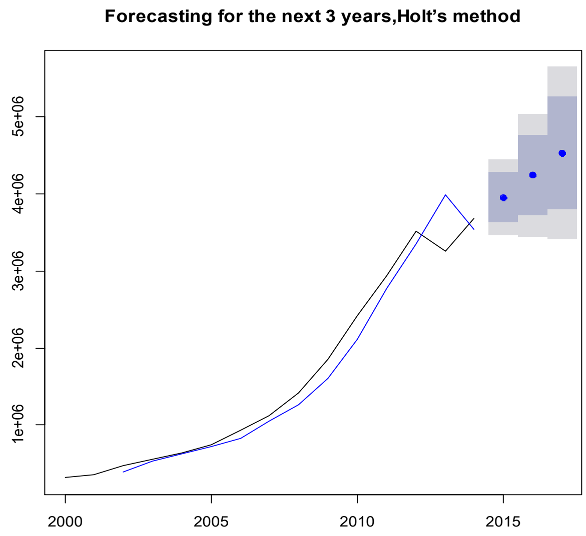

Figure 4: Actual And Predicted Number Of Tourists Using Holt‟s Method

The data series is not a stationary series and it has the trend component. This series can become a stationary series using the first differences. Based on Akaike Information Criteria (AIC), the best ARIMA model is ARIMA (2,1,2). The equation is given by:

$$Y_t = -0.1583Y_{t-1} + 0.5428T_{t-2} + 0.4266\varepsilon_{t-1} + 0.9997\varepsilon_{t-2} + \varepsilon_t$$

Using the ARIMA (2,1,2) method, the number of tourists is expected to be 3,101,692 for year 2015, 3,565,361 for year 2016 and 3,182,100 for year 2017.

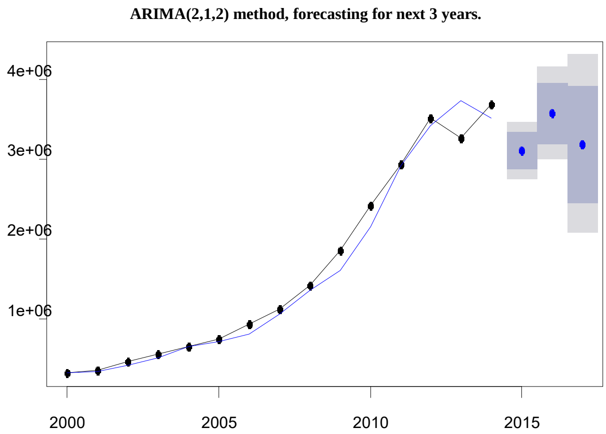

Figure 5: Actual and Predicted Number of Tourists Using ARIMA(2,1,2) Method

Based on values of RMSE and MAPE for Holt‟s Exponential Smoothing and ARIMA methods, ARIMA (2,1,2) is the best method to predict the number of tourists in Albania.

4. Conclusions

The number of tourists has significant impacts on Albania economy. These impacts can be counted as both of negative and positive effects. The most important effect of tourism on economy can be known as number of changes on supply and demand chain in the destination which is the host of tourists. Tourists‟ demand or simply tourist consumption contributes to GDP, increasing the employment rate, making new source of revenue for local people, private and public sectors and destination‟s government and so on.

From 2000 to 2014 the number of tourists arrived in Albania is increased approximately 11 times. In 2014, around 93% of the total number of tourists came from Europe Countries. The results of Albania Tourist Area Life Cycle indicated that the midpoint of Life Cycle is year 2010, the duration of growth time is 13.6 years and the Carrying Capacity is 4,886,858 tourists. The time series analysis showed that the best model was ARIMA(2,1,2), and according to this model the number of tourists is expected to be 3,101,692 for year 2015.

The Logistic differential equation, Exponential Smoothing techniques and ARIMA models all helped in the development of a better understanding the overall picture of the tourism story in Albania. Such information is particularly useful for new investors and developers as well as longer term planners and government departments.

References

[1] Butler, R.W. (1980). The Concept of a Tourism Area Cycle of Evolution: Implications for Management Resources. *The Canadian Geographer*, 24(1), 5-16.

[2] Shobeiri Nejad, S.A.H. and Tularam, G.A. (2010). Modeling tourist arrivals in destination countries: an application to australian tourism. *J. Math. Stat.*, 6, 431-441.

[3] Shumway, R. and Stoer, D. (2006). Time Series Analysis and Its Applications with R Examples, Springer.

[4] Song, H. and G. Li, 2008. Tourism demand modeling and forecasting. Tourism Management, 29: 203-220.

[5] D.Kusharavy, Roand De Guio. Application of S-Shaped Curves. TRIZ-Future Conference 2007: Current Scientific and Industrial Reality,Nov. 2007, Frankfurt,Germany.pp.81-88.

[6] Th. Modis. Strength and weaknesses of S-curves: Technological Forecasting & Social Change 74 (2007) 866-872

[7] A.Saayman,M.Saayman. Forecasting Tourist Arrivals in South Africa: Acta Commercii 2010

[8] S.Singh. Tourism Area „ Life Cycle‟: A clarification: Annals of Tourism Research,Vol 38,No. 3pp. 1185-1187, 2011