Tourist Flow Forecasting Approach

I N Ostapenko[1], R S Usenko[1]

[1]Institute of Economics and Management (structural unit), FGAOU VO "Crimean Federal University named after IN AND. Vernadsky ", Simferopol st. Sevastopol 21/4, 295011, Russian Federation

E-mail: i.n.ostapenko@mail.ru, r_usenko@rambler

Abstract. This article discusses the issues of using methods of analyzing and forecasting seasonal time series to model the tourist flow as one of the most significant parameters of tourism operations. The authors considered reasonability of using the Holt-Winters approach to model time series, estimated model's coefficients, and proposed a forecasting model. The authors also demonstrated use of the Holt-Winters approach to model parameters of tourism operations for a "short series" - the tourist flow to the Republic of Crimea at this stage of the peninsula's development.

1. Introduction

Tourism is a very important sphere of economic activity for the development of any country around the world. For the Republic of Crimea, it is a top priority budget revenue-generating sector, which is why it is necessary to use as precise as possible forecasts of the main indicators, tourist flow in particular, to plan the strategy of development. Let us consider the fluctuation of the number of tourists in the peninsula in 2003-2018 (yearly; Fig. 1) and in 2015-2019 (monthly; Fig. 2).
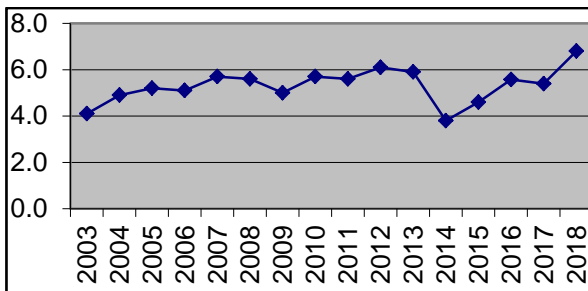

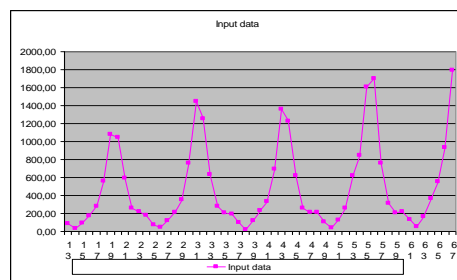**Figure 1.** Number of tourists in Crimea in 2003-2018 [1].

Figure 2. Monthly fluctuation of tourists in Crimea in 2015-2019 [1].

Having analyzed the time series given in Figure 1, we may note how the trend changed in 2014 due to execution of the Treaty of accession between the Russian Federation and the Republic of Crimea on March 18, 2014, which provided for establishment of two new subjects of the Russian Federation - the Republic of Crimea and the federal city of Sevastopol. The effect of a structural shift in the Crimean economy in 2018 on the tourist flow trend change due to the change in the passenger throughput by

means of transport caused by commissioning of the new airport terminal in the City of Simferopol and the architecturally and technically unique Crimean Bridge is clear. These changes lead to the need in particular attention to modeling a time series trend of the number of tourists in the peninsula in the context of structural changes.

Comparability of levels of the input data time series is one of the main conditions required for correct representation of the real tourist flow fluctuation. Results of a fluctuation study with incomparable data will be incorrect.

Application of the G. Chow test [2] to the input data indicates lack of structural stability, and statistical significance of differences in assessing equation parameters for the 2003-2013 and 2014-2018 samples obtained by dividing the whole range of observations into two parts, each whereof may be linearly approximated. As $F_{actual} = 23.94 > F_{table} = 3.88$, the events of 2014 significantly affected the input data trend change.

In this context, the issue of modeling and forecasting the tourist flow to Crimea in the conditions of a "short times series" deserves particular attention. Given a justified structural shift in the time series, we may assume that to generate an adequate model of forecasting the tourist flow to the peninsula, only the data collected after the Republic of Crimea joined the Russian Federation in 2014 may be used.

So long as tourism operations in Crimea remain seasonal, the monthly number of tourists in the Republic of Crimea features a clearly manifested seasonal pattern peaking in summer (see Figure 2).

2. Review of tourist flow modeling methods

The processes taking place in the sphere of tourism are highly uncertain and characterized by multicriteriality and seasonality; this dictates the need in using various methods and models for analyzing and forecasting [3].

Works of numerous foreign and Russian researchers are dedicated to the issues of modeling and forecasting the tourist flow, as well as to analyzing its connection with other factors. English researchers S.F. Witt and C.A. Martin developed econometric models for forecasting international tourism demand on the basis of tourists visits from West Germany and the United Kingdom to assist in forecasting [4]. In the course of their research, T. Garín-Muñoz and T. Pérez Amaral determined the impact of the economic determinants of the international demand for tourist services. The researchers used a panel data set of seventeen countries over the period from 1985 to 1995 to determine the effects of a range of determinants, such as real income per capita, foreign exchange rates, and real prices, on the demand for Spanish tourist services [5]. The applied thematic study performed by L. Botti et al. helps to review tourist demand in France over the periof from 1975 to 2003 using an econometric model. The authors substantiated positive relationship between tourist expenditures and GDP of the visited country, as well as a negative relation between tourist expenditures and prices [6]. Researchers Y. Yang and K. Wong evaluated an econometric approach to modeling tourist flows to investigate and estimate the spillover effects in inbound and domestic tourist flows to 341 in mainland China. Study results confirm the fact that spillover effects resulting from various internal and external factors impact inbound and domestic tourism [7].

Works of such Russian researchers as N.A. Ziuliaev [8], M.A. Bednova and T.A. Ratnikova [9], T.P. Nikolaeva and E.S. Oreshkina [10], etc., are also dedicated to tourist flow modeling. However, different regions are often characterized by unique conditions, which is why the set goal is rather urgent.

In order to analyze and generate an adequate model of forecasting the monthly number of tourists in the Republic of Crimea, it appears reasonable to use seasonal time series forecasting methods, because the tourist flow has not yet become evenly spread over the year. Various recent studies of tourist flow modeling and forecasting employed different time series models and methods, such as Naive I and Naive II [11], the seasonal autoregressive moving average (SARIMA) [12-14], and neural networks [15, 16].

Seasonal time series modeling methods also include the Holt-Winters model [17, 18]. This model is rather widely used to model and forecast the tourist flow [19-21].

Most statistical forecasting methods require time series to be "sufficiently long." For instance, it is desirable to have input data for at least 3-5 full periods when studying seasonal time series. That is why in order to model the tourist flow to the peninsula in the context of a "short time series", we will calculate parameters of the Holt-Winters model, which allows modeling seasonal time series even if the input data are insufficient. The monthly statistical data for 2014-2018 and 7 months of 2019 will be used as the input data.

3. Holt-Winters's triple exponential smoothing model

3.1. Calculations of the model's characteristics

The model originated as a combination of the Holt's two-parameter linear growth model and the Winters's seasonal model. It takes the seasonal multiplier factor and the additive trend into account.

When updating coefficients for different periods, the values of the exponentially smoothened series, trend-cycle component, and seasonal coefficient for the current period may be calculated as follows:

$$S_t = \alpha \frac{Y_t}{K_{t-L}} + (1-\alpha)(S_{t-1} + b_{t-1}). \tag{1}$$

$$b_t = \beta(S_t - S_{t-1}) + (1-\beta)b_{t-1}. \tag{2}$$

$$K_t = \gamma \frac{Y_t}{S_t} + (1-\gamma)K_{t-L}. \tag{3}$$

where α – series smoothing constant; Y_t – current time series value; K_{t-L} — seasonal coefficient of the previous period (for the same month of the previous year); S_{t-1} – smoothened value of the exponentially smoothened series for the previous period; b_{t-1} – trend-cycle component for the previous period; β – trend smoothing constant; γ – seasonality smoothing constant.

Model's values can be calculated as follows:

$$\hat{Y}_t = (S_{t-1} + b_{t-1}) \cdot K_{t-L}. \tag{4}$$

The corresponding predicted values can be calculated as follows:

$$\hat{Y}_{t+\tau} = (S_t + \tau \cdot b_t) \cdot K_{t-L+\tau}. \tag{5}$$

where $\hat{Y}_{t+\tau}$ - the values predicted using the Holt-Winters model for τ following periods.

The choice of smoothening constants α, β, and γ deserves particular attention in this model, as they determine accuracy and adaptivity of predicted values. Predicted values are more sensitive to the latter values of the input time series when α is the largest of the constants, and more sensitive to "fresher" trends when β is the largest one [22]. According to some researchers, smoothening constants shall be in the range from 0.1 to 0.3 [23].

Values of smoothening constants cannot be chosen in advance. Based on the recommendations provided in the available publications, we may set the initial values of smoothening constants at, for instance, 0.3, and then optimize the model using the lowest value of the mean square error (MSE).

The optimal model is calculated with the following smoothening constants: α = 0.1302, β = 0, and γ = 0. Model's characteristics:

$$\hat{Y}_{t+\tau} = (357,32 + 4,6113 \cdot \tau) \cdot K_{t-L+\tau} \tag{6}$$

See the input data of the time series from 2015 through 7 months of 2019 and the calculated predicted values in Figure 3.

3.2. Evaluation of the model's accuracy and adequacy

The coefficient of determination may be used to evaluate the model's quality; this coefficient must be as close to 1 as possible [24]. The main approximation error may be used to substantiate accuracy [25, 26].

Good predicting properties are characterized by a high coefficient of determination ($R^2 = 0{,}989$) and the mean approximation error $MAPE = 9{,}54\%$ within the 10% range.

The graph of the autocorrelation function (Figure 4) may be used to check model residuals for independence. The autocorrelation function of residuals remains within the confidence intervals; this means that the residuals are independent.
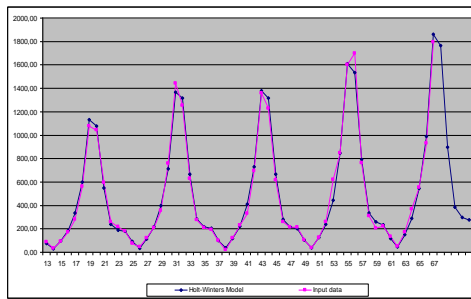

Figure 3. The input time series from 2015 through 7 months of 2019 and the calculated predicted values.
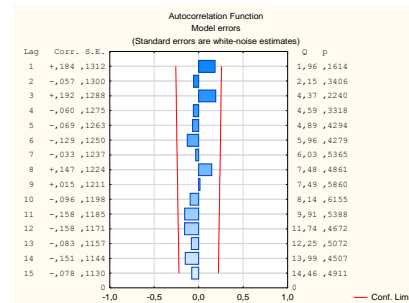

Figure 4. Autocorrelation function of model residuals.

The two-sided David-Hartley-Pearson test (R/S test) may be used to check residuals for normality. In this case, the test's calculated value is as follows:

$$U = \frac{R}{S} = \frac{e_{\max} - e_{\min}}{\sqrt{\sum_{i=1}^{n}(e_i - \overline{e})^2 / (n-1)}} = \frac{176{,}28 - (-88{,}88)}{\sqrt{126156{,}69(55-1)}} = 5{,}49 \tag{7}$$

The Gaussian hypothesis for the residuals is accepted, because this test's calculated level falls within the range given in the table (3.79-5.63).

3.3. Forecasting

Table 1 provides a Holt-Winters model-based forecast for 2019. Comparison of the results obtained by modeling and the available statistical data for the first seven months of 2019 indicates that this modeling method provides an adequate result.

Table 1. The values for 2019 predicted using the Holt-Winters model.

	Months					
	01	*02*	*03*	*04*	*05*	*06*
Model-based forecast	117.74	44.10	149.24	287.29	543.83	989.94
	Months					
	07	*08*	*09*	*10*	*11*	*12*
Model-based forecast	1,862.25	1,765.96	897.27	383.95	297.60	277.86

For instance, according to the statistical data for January-July 2019, the total number of tourists amounted to 4,004.9 thousand people, while the model-predicted value was only slightly smaller - 3,994.39 thousand people.

4. Conclusion

In order to analyze and forecast the tourist flow to the Republic of Crimea, the authors assessed use of the adaptive short-term forecasting models for socioeconomic processes intended to generate self-correcting models that help to timely react to changing circumstances [27]. We believe that the use of adaptive methods is the most promising area of modeling and forecasting small time series, because it is the information for the latter periods that is the most valuable for forecasting indicators of tourism operations.

As long as tourism operations in Crimea are characterized by cyclic seasonal phenomena and the time series reflecting indicators of tourism operations involve cyclic seasonal fluctuations, in order to model indicators, such as the monthly number of tourists in the Republic of Crimea, the authors employed the Holt-Winters model, the prognostic accuracy whereof is as high as the one demonstrated by more complex seasonal time series behavior models.

However, use of only one type of models is not sufficient to forecast indicators of tourism operations accurately enough; different models need to be combined and used as a system according to the study object's peculiarities [28]. Furthermore, use of the analytical and forecasting techniques based on not only a set of quantitative methods of analysis and forecasting, but also their combined use with qualitative techniques will help to competently predict indicators of tourism operations.

References

[1]   Monthly reference information on the number of tourists visiting the Republic of Crimea. Website of the Ministry of Resorts and Tourism of the Republic of Crimea https://mtur.rk.gov.ru/ru/structure/14 (08/15/19

[2]   Chow Gregory C 1960 Tests of Equality Between Sets of Coefficients in Two Linear Regressions *Econometrica* **28(3)** 591-605

[3]   Ostapenko I N 2016 Methods of analysis and forecasting in tourism activities Analysis, modeling, management, development of socio-economic systems: Collection of scientific papers of the X International Symposium School AMUR-2016 pp 305-308

[4]   Witt Stephen & A Martin Christine Econometric Models for Forecasting International Tourism *Journal of Travel Research* **25** pp 23-30

[5]   Garin-Munoz T, Perez Amaral T 2000 An Econometric Model for International Tourism Flows to Spain *Applied Economics Letters* vol 7 525-529

[6]   Botti L, Peypoch N, Randriamboarison R, Solonandrasana B 2007 An Econometric Model of Tourism Demand in France *Tourismos: an International Multidisciplinary Journal of Tourism* **2(1)** 115-126

[7]   Yang Y, Wong K F 2012 A Spatial Econometric Approach to Model Spillover Effects in Tourism Flows *Journal of Travel Research* **51(6)** 768-778

[8]   Zyulyaev N A 2014 The tourist market of Russia: economic aspect *Russian Journal of Entrepreneurship* **2** 226-234

[9]   Bednova M A, Ratnikova T A 2011 Econometric analysis of demand for inbound tourism in Russia *Applied Econometrics* **1** pp 97-113

[10]   Nikolaeva T P, Oreshkina E S 2016 Determinants of demand for inbound tourism (for example, countries of Europe and the CIS) *Service in Russia and abroad* **8** pp 17-28

[11]   Martin A and Witt F 1989 Forecasting tourism demand: A comparison of the accuracy of several quantitative methods *International Journal of Forecasting* **5** 1-13

[12]   Lim C and McAleer M 1999 A seasonal analysis of Malaysian tourist arrivals to Australia *Mathematics and Computers in Simulation* **48(6)** 573-583

[13]   Chen C F and Chang Y H and Chang Y W 2009 Seasonal ARIMA fore- casting of inbound air travel arrivals to Taiwan *Transportmetrica* **5(2)** 125-140

[14]   Apergis N A, Mervar A and Payne J E 2017 Forecasting disaggregated tourist arrivals in Croatia: Evidence from seasonal univariate time series models *Tourism Economics* **23(1)** 78-98

[15]   Law R and Au N 1999 A neural network model to forecast Japanese demand for travel to Hong

Kong *Tourism Management* 20(1) 89-97

[16] Law R 2000 Back-propagation learning in improving the accuracy of neural network-based tourism demand forecasting *Tourism Management* **21(4)** 331-340

[17] Holt C C 1957 Forecasting seasonals and trends by exponentially weighted moving averages *Pittsburgh PA: Carnegie Institute of Technology*

[18] Winters P R 1960 Forecasting sales by exponentially weighted moving averages *Management Science* **6(3)** 324-342

[19] Mishra P K & B Rout, Himanshu & Pradhan, Bibhuti B 2018 Seasonality in Tourism and Forecasting Foreign Tourist Arrivals in India *Iranian Journal of Management Studies*

[20] Premila C & Ponnusamy Ramalingam 2017 Evaluating and Forecasting Room Demand in Tourist Spot using Holt-Winters Method *International Journal of Computer Applications*

[21] Ostapenko I N 2017 The Holt-Winters method in predicting indicators of tourism activity *Scientific Review: theory and practice* **10** pp 76-89

[22] Handanhal V R 2013 Determining the Optimal Values of Exponential Smoothing Constants Does Solver Really Work? *American Journal of Business Education* **6(3)** 347-360

[23] Schroeder R, Rungtusanatham M J, & Goldsteing S 2013 Operations Management in the Supply Chain: Decisions and Cases *UK: McGraw-Hill Higher Education*

[24] 2007 Econometrics: a textbook II Eliseeva S V, Kurysheva T V Kosteev and others under the editorship of I I Eliseeva (M .: Fin. and Statistics) 576 p

[25] 2003 Workshop on Econometrics: Textbook. allowance I I Eliseeva, S V Kurysheva, N M Gordeenko et al (M.: Finance and Statistics) 192 p

[26] Mamula M 2015 Modeling and Forecasting International Tourism Demand - Evaluation of Forecasting Performance *International Journal of Business Administration* **6(3)** 102-112

[27] Sazhin Yu V 2013 Analysis of time series and forecasting: textbook (Saransk: Publishing House of Mordov. University) 192 p

[28] Ostapenko I N 2017 On expert methods for modeling indicators of tourism activity *Management of economic systems: an electronic scientific journal* **9(103)** p 26