JOURNAL OF TOURISM AND SERVICES
Issue 21, volume 11, ISSN 1804-5650 (Online)
www.jots.cz

Modelling and Forecasting Inbound Tourism Demand to Croatia using Artificial Neural Networks: A Comparative Study

Murat Çuhadar
Süleyman Demirel University, Isparta, Turkey

Received: 19 June 2020. **Revision received**: 18 August 2020. **Accepted**: 10 October 2020

Abstract
Tourism demand is the basis on which all commercial decisions concerning tourism ultimately depend. Accurate estimation of tourism demand is essential for the tourism industry because it can help reduce risk and uncertainty as well as effectively provide basic information for better tourism planning. The purpose of this study is to develop the optimal forecasting model that yields the highest accuracy when compared to the forecast performances of three different methods, namely Artificial Neural Network (ANN), Exponential Smoothing, and Box-Jenkins methods for forecasting monthly inbound tourist flows to Croatia. Prior studies have been applied to forecast tourism demand to Croatia based on time series models and casual methods. However, the monthly and comparative tourism demand forecasting studies using ANNs are still limited, and this paper aims to fill this gap. The number of monthly foreign tourist arrivals to Croatia covers the period between January 2005-December 2019 data were used to build optimal forecasting models. Forecasting performances of the models were measured by Mean Absolute Percentage Error (MAPE) statistics. As a result of the experiments carried out, when compared to the forecasting performances of various models, 12 lagged ANN models, which have [4-3-1] architecture, were seen to perform best among all models applied in this study. Considering both the empirical findings obtained from this study and previous studies on tourism forecasting, it can be seen that ANN models that do not have any negativities (such as over-training, faulty architecture, etc.) produce successful forecasting results when compared with results generated by conventional statistical methods.

Key Words: Modelling, Forecasting, Tourism Demand, ANN's

JEL Classification: C53, L83

Reference: Çuhadar, M. (2020). Modelling and Forecasting Inbound Tourism Demand to Croatia using Artificial Neural Networks: A Comparative Study. *Journal of Tourism and Services*, 21(11), 55-70. doi:10.29036/jots.v11i21.171

1. Introduction

Travel and tourism, which is one of the world's largest economic sectors, create jobs, leads to exports, and creates prosperity worldwide. In the annual analysis of the "Global Economic Impact of Travel & Tourism" published by World Travel and Tourism Council (WTTC), the sector was shown to constitute 10.3 % of global GDP and 330 million jobs or 10 % of total employment in 2019. According to the results of research conducted by WTTC and Oxford Economics, the travel and tourism industry grew by 3.5% in 2019, showing that it has surpassed the global economy (2.5%) for the ninth consecutive year. In the past five years, one of four new jobs in the world has been created by the tourism industry. This has made the tourism industry the best partner for governments to create jobs (WTTC, 2020). With its advantageous location and natural beauty, Croatia has become an important tourism destination in international tourism. The wars and conflicts that took place in the former Yugoslavia in the 90s seriously affected both the demand for international tourism and the tourism infrastructure, but in the last 20 years, tourism has started to rise again. Incoming tourism is an important economic activity for Croatia and is the main source of export revenues. International tourists' spending in Croatia accounts for almost 20% of GDP - the largest share in the EU. Croatia has a typical "sea, sand, sun" tourism that concentrates on the coastal regions during the summer months (Orsini & Pletikosa, 2019).

Under increasingly competitive circumstances, government officials, administrators and practitioners in tourism sector are faced with the requirement of forecasting tourist flows for effectual planning. Tourism demand forms the starting point of all business decisions regarding the tourism sector. Forecasting future tourism flows constitute an essential ingredient in tourism planning. Identifying, modelling and forecasting future tourist flows are critical aspects for tourism decision makers to ensure the short and long-term success of tourism services and destinations. Tourism demand forecasts are one of the factors that tourism businesses need in planning activities to be carried out by local and general public administrations. As in other economic sectors, reliable and accurate forecasts are needed in the tourism sector. According to Song and Turner (2006), the success of tourism enterprises depends to a large extent on the state of tourism demand. One of the most important problems that business managers face is the failure of meeting the market demand. Lack of inventory capability in tourism services makes demand forecasts in this sector more important. This lack of ability to store services is due to the perishability of services and their simultaneous production and consumption. Because the instability and seasonal fluctuations in demand make it difficult for businesses operating in the tourism sector to make an ideal balance between supply and demand. In many other sectors of the economy, this balance can be established by storing and keeping the products (Saayman & Saayman, 2010). However, this is not possible for tourism in the services sector. For example, it is not possible to store a hotel room or an airplane seat which cannot be sold for later sale. Due to the fact that tourism demand plays a key role as the determinant of business profitability, demand forecasts for the future play a very important role in planning operations (Goh & Law, 2011). Tourism demand modelling and forecasting studies have gained importance since the inception of tourism studies as an academic field and with the understanding of the effects of tourism on national economies (Dogru & Sirakaya-Turk, 2018). Therefore, researchers, practitioners and governments recognize the need for demand forecasts in tourism. Recognizing the importance of accurate forecasts for the tourism industry, both short- and long-term forecasting techniques have been widely applied in international tourism flows reported in empirical studies published with empirical findings. Tourism demand forecasting uses various measures. Examples of these are the number of tourists arriving the country, the number of visitors staying in the accommodation establishments, the number of overnight stays, tourist expenditure and the occupancy rates. However, the most frequently used measure is tourist arrivals (Fretchling, 2001; Lim, 2006; Song et al. 2010; Höpken et al. 2020). According to many authors such as Law & Pine (2004); Kon & Turner (2005); Song & Li (2008); Peng et al. (2014); Önder (2017); Höpken et al. (2020) and Zhang et al. (2020), in the tourism forecasts, there is no single method that produces the best results in all cases; the results will vary according to the method and model to be used, the characteristics of the data set, the number of observations and the estimated period, and that many researchers working in this field have agreed on. In other words, there is no magic model or method that can produce the best result in every situation. Based on this approach, the purpose of this study is to develop the optimal forecasting model that yields the highest accuracy when compared the forecast performances of three different methods namely Artificial Neural Network (ANN), Exponential Smoothing and Box-Jenkins methods for forecasting monthly inbound tourist flows to Croatia. Prior studies have been applied to forecast tourism demand to Croatia based on time series models, such as exponential smoothing and extrapolative methods (Baldigara, 2013), ARIMA (Baldigara & Mamula, 2015), ARAR (Apergis et al., 2015) and ARFIMA (Bahovec & Erjavec, 1999), ARDL method (Mervar & Payne, 2007), casual methods, such as regression, VAR and GARCH models (Tica & Kožić, 2015). However, the monthly and comparative tourism demand forecasting studies using ANNs are still limited and this paper aims to fill this gap. In light of the existing forecasting literature with respect to tourism in Croatia, it can be said that this study is the one of the limited studies to build and compare ANN models in modelling and forecasting tourism demand to Croatia. Having introduced the emphasis of demand forecasting in tourism and research objectives, the remainder of this paper is organized as follows. Next section describes the research methodology. Section 3 describes the data employed in the study as well as research design. Thereafter, forecasting performances of alternative methods and models are compared and the empirical findings are presented. In the last section, the implications and contributions of the research and suggestions for future research are included.

2. Data and Method

In the following subtitles, explanations about data analysis and forecasting methods used in the study are included.

2.1. Data Analysis

In the study, international tourist arrivals were taken as the measure of inbound tourism demand and the number of monthly foreign tourist arrivals to Croatia cover the period between January 2005-December 2019 data were used to build optimal forecasting models. Data were obtained from the Croatian Bureau of Statistics web site. The Croatian Bureau of Statistics takes over the data on tourist traffic (number of incoming tourists and number of nights) and accommodation capacities from the Croatian National Tourist Board, extracting them from the e-Visitor system, statistically processing and publishing them on the institution's website. Tourist arrival is the number of persons (tourists) who arrived and registered their stay in an accommodation establishment. (Croatian Bureau of Statistics, 2020). In the study, monthly data was preferred to be analysed in more detail by considering seasonal, trend and other time series features. In the study firstly, the time series characteristics of the data used were analysed and the time series components that were effective on the data were determined. In the following stage, analyses were carried out to create forecasting models suitable for the structure of the data from the Exponential Smoothing, Box-Jenkins and ANN methods. In the process of determining the appropriate model, the smoothing constants that minimize the mean of error squares for exponential smoothing models, Autocorrelation (ACF) and Partial Autocorrelation functions (PACF) and Bayes Information Criteria (BIC) for Box-Jenkins models were taken into consideration. When constructing appropriate ANN models, different time delays in the input layer and the instantaneous data values are used in the output layer. For ANN models, parameters and parameter values that will minimize the prediction error produced by the model were used. It was determined which method produced more successful results by comparing the estimated values produced by the models with the number of foreign tourist arrivals realized. Forecasting performances of the applied methods were measured by the "Mean Absolute Percentage Error (MAPE)" statistic. The mathematical expression of MAPE is as follows:

$$MAPE = \frac{\sum_{t=1}^{n} \frac{|e_t|}{y_t}}{n} 100(\%)$$

where:

$(e_t = y_t - \hat{y}_t)$

y_t = Value of the observation at time *t*,

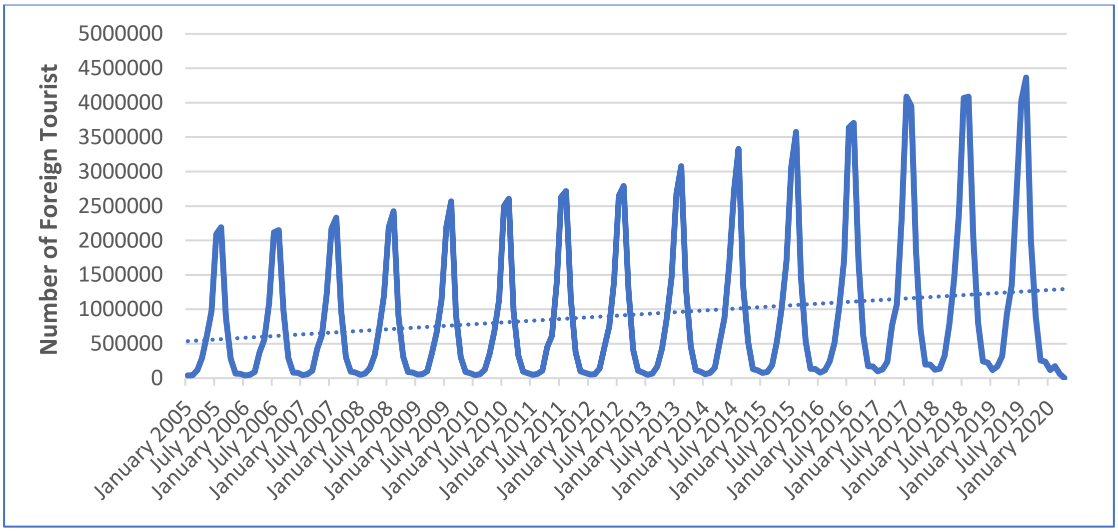
$\hat{y}_t =$ Fitted value for the observation at time *t*,
n is the length of forecast horizon.

In the evaluation of alternative models, if the MAPE statistic value is below 10 %, it is accepted as a high accuracy, good accuracy between 10-20 %, reasonable accuracy between 20-50 % and an unsuccessful estimate over 50% (Lewis, 1982).

2.2. Analysis of Time Series Properties of Data

The time graph of the foreign tourist arrivals data from January 2005 to December 2019 is given in Figure 1. When the graph is analysed, it is seen that the data is influenced by seasonal fluctuations together with a positive trend. In general, seasonal fluctuations reached the highest values in August of consecutive years, while the lowest in January.

Figure 1. **Time Series Graph of Monthly Foreign Tourist Arrivals in Croatia**



Source: Own calculation on the basis of data from Croatian Bureau of Statistics.

Table 1. **Model Summary of Trend Analysis**

Equation	Model Summary				Parameter Estimates		
	R Square	F	Df	Sig.	Constant	b1	b2
Linear	.059	11,066	1	.001	476784,929	5047,364	
Logarithmic	.041	7,591	1	.006	-40152,091	(ln) 231151,554	
Quadratic	.064	6,007	2	.003	651585,334	-715,286	31,838
Exponential	.059	11,189	1	.001	229411,639	.007	

Source: Own calculations.

In order to determine the forecasting models suitable for the structure of the data used in the study, trend analyses in different forms (linear and non-linear) were performed. As a result of the applied trend analyses, the data showed a positive trend. It was observed that the F tests performed to test the statistical significance of various linear and non-linear (Logarithmic, Quadratic, and Exponential) trend analyses were statistically significant at the significance level of 0.05. The model summary of the trend analyses are given in Table 1.

Similarly, in order to determine the seasonal effects in the data used in the study, seasonal decomposition analysis was applied by using the "ratio to moving average" method, also called X-12 ARIMA in the literature. The seasonal factor values given in Table 2 reveal that the data are influenced by the seasonal component that repeats every twelve months.

Table 2. **Seasonal Factor Values of the Data**

Period	Seasonal Factor Values (%)	Period	Seasonal Factor Values (%)
January	7.0	July	309.8
February	8.6	August	333.2
March	17.6	September	139.9
April	53.2	October	48.0
May	91.6	November	13.1
June	166.8	December	11.3

Source: Own calculations

As can be seen in Table 2, the seasonal pattern is formed by reaching the lowest value in January of each year and the highest value in August.

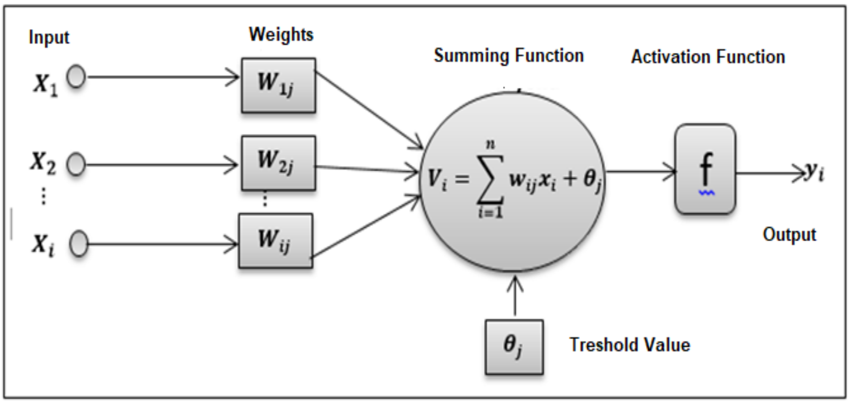
2.3. Artificial Neural Networks (ANNs)

Experiments and researches in the field of science, natural and engineering sciences regarding the brain structure have opened a horizon in modelling the information processing processes of the human brain and have led to the development of systems that mimic the working principles of the human brain. The smallest element of the human brain is nerve cells. Nerve cells perform their information processing activities in a group, not alone. Recently, ANN-based models have been one of the most important non-parametric methods proposed for time series forecasting. ANN is defined as a structure consisting of several interconnected units or artificial nerve cells that mimic biological neural networks (Chen et al., 2007). According to another definition, ANNs are complex systems that are created in the form of artificial nerve cells (neurons), which are compared to biological neurons in the human brain, to form connections with each other at different effect levels. Artificial neurons come together to form ANN. Neurons are composed of five main elements: input connections, weights, addition function, activation function and output connections. The basic element of an ANN is the neuron as shown in Figure 2.

In Figure 2, (x_i) represents inputs, (w_{ij}) weights, (θ_j) threshold value, (*f*) activation function and (y_i) output. In the information processing continuum, information input from the outside world to the neuron takes place first through the input connections $(x_1, x_2, \dots . x_i)$. Information entry to a neuron from other neurons is done by multiplying by a weight$(w_{1j}, w_{2j}, \dots , w_{ij})$ with a value between (-1) and (+1). In the next step, the information multiplied by weights is passed through an activation function (*f*) to obtain the output (y_i) (Golam & Hasin, 2013). The activation function is briefly the mechanism for converting input signals to output signals for each process element. There are generally three layers in an ANN: input layer with interactive neurons, output layer and hidden layer(s). As can be seen in Figure 3, the first layer in the ANN is the input (also called the input) layer, which functions to receive external data from the artificial neural network.

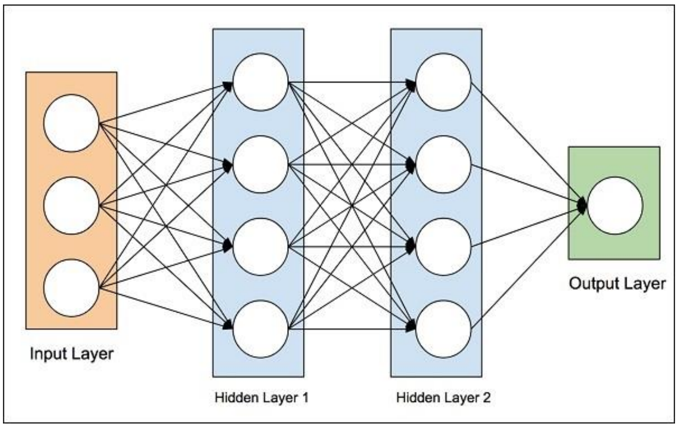
JOURNAL OF TOURISM AND SERVICES
Issue 21, volume 11, ISSN 1804-5650 (Online)
www.jots.cz

Figure 2. **Basic Elements of an Artificial Neuron**



Source: Golam & Hasin, (2013)

Figure 3. **An ANN model**



Source: Çuhadar et al. (2013).

The input layer is the first layer in ANN, and serves to network the data from the outside world. The input layer consists of data from variables that are assumed to be related to the problem being solved. Therefore, the number and quality of neurons in the input layer are shaped according to the variables that affect the problem. The last layer in the ANN is the output layer that performs the function of exporting the processed information. The layer between the input and output layers is called the hidden layer. Neurons in the hidden layer do not interact with the outside world, they take the task of receiving the signals from the input layer and transmit them to the output layer. ANNs are trained by learning a problem through the datasets shown to it. Training of ANNs, in other words the learning of the network is the process of updating the connection weights to fulfil a function expected by the user. The most used method in the literature for updating ANN parameters is the "Back Propagation" algorithm. Over the years, ANN architectures and models such as MLP, RBF, Hopfield, Jordan, Elman, Kohonen SOM, which are suitable for use in different fields, have been developed. The most widely used ANN architecture for predictive purposes in the social sciences and tourism forecasting field is the back propagation-feed forward Multilayer Perceptron (MLP) model (Wong et al. 2000; Zhang & Qi, 2005; Song & Li, 2008; Moreno et al. 2011; Teixeira & Fernandes, 2012; Bayramoğlu & Başarir, 2018). MLP model is very popular because it is widely applicable in solving business related problems such as prediction, classification and modelling (Smith, 2002). In the use of MLP models for forecasting, determining the network architecture is especially important issue. In the literature, it is stated that network architectures with a single hidden layer are sufficient in estimating time series. The inputs to be presented to the network are among the lagged observations $(y_{t-1}, y_{t-3}, y_{t-12} \dots . y_{t-N})$ of the data set used; output consists of original (y_t) observation values. The relationship between the output value and the inputs is as follows:

$$y_t = w_0 + \sum_{j=1}^{p} w_j f\left(w_{0j} + \sum_{i=1}^{N} v_{ij} y_{t-i}\right) + e_t$$

In this equation, w_j, v_{ij} represents the weight values between neurons, *p* represents the number of neurons in the hidden layer and *f* is the activation function used in the hidden layer. The most used activation functions are sigmoid $f(x) = \frac{1}{1+e^{-x}}$ and hyperbolic tangent $f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$ functions. (Egrioglu et al., 2017).

The basic decisions to be taken regarding an ANN model for forecasting and modelling are data preparation, input variable selection, network type and architecture selection, transfer function and determination of training algorithm, model selection, evaluation and verification. In modelling the data used in the study with ANN, the forecasting performances of alternative ANN architectures created using different data sets were analysed. 160 observations of the current 180-month data are grouped as training and 20 observations are grouped as test data. In the input layer of each network established for trial purposes, different time delayed tourist arrival observations $(y_{t-1}, y_{t-3}, y_{t-12} \dots\dots y_{t-k})$ recommended for the prediction of monthly time series were used, and the original observation values (y_t) were used in the output layer. The data was normalized in the range [0:1] before keying into the computer using the equation below,

$$x_n = \frac{x_0 - x_{\min}}{x_{\max} - x_{\min}}$$

Where:

x_0 = original values,

x_n = normalized values,

x_{min} = the lowest number included in the data, and

x_{max} = the biggest number included in the data.

In order for Multilayer Perceptron (MLP) networks to be used for forecasting, the structure of the network must be determined. The process of determining the network structure briefly includes the processes of determining how many layers the network will consist of, how many processing elements will be in each layer, and which transfer function these process elements will have. Within the scope of the study, model alternatives created with 3, 4 and 5 layer architectures were tried. The learning process of MLP network was realized by presenting the training data to the network. The implementation of the method was carried out with the Neural Network Toolbox of the "MATLAB – Simulink (version: r2019a)" computer program. The stop limit of the training process of the established networks was determined as 10,000 iterations for each trial. In the ANN application performed, the predictive performances of candidate ANN models with 12 lagged datasets in the input layer and various number of neurons (1-7) in the hidden layer were examined. Since there is no generally accepted rule in the literature to determine the number of neurons in hidden layers, the network has been trained with different neuron number alternatives. In the process of modelling the data with alternative multi-layered ANN structures, all the different ANN architectures were tested using the data determined subsequently for testing purposes. The forecasting performances of models for different architectures were examined by comparing the generated forecast values as a result of the test process with real observation values.

As a consequence of several attempts, it has been observed that 12 lagged Multilayer Perceptron (MLP) model which has [4-3-1] architecture has showed best forecasting performance. During the training of the ANN model, the connection weights were updated with the "back propagation" method. The "Logarithmic Sigmoid" algorithm was used as the activation function and the "Levenberg-Marquardt" algorithm was used as the training function.

2.4. Exponential Smoothing

The exponential smoothing method consists of different application options in which the estimates are constantly updated, considering the recent changes in the data. In the Exponential smoothing method, the weighted averages of past period values are calculated and included in the models as the estimated value of the future periods. The working principle of this method is that more recent data and observations are more important than the effect of very old observations or data (Yaffee, & McGee, 2000). Holt-Winter's seasonal exponential smoothing method is suitable for modelling and estimating data containing trend and seasonal effects. The equations used in the calculation of the Multiplicative-Seasonal Holt-Winter's method are given below (Makridakis et al., 1998).

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

Seasonal: $S_t = \gamma_t(Y_t - L_t) + (1-\gamma)S_{t-s}$

Forecast: $F_{t+m} = L_t + S_{t-s+m}$

In the given equations;

$S =$ Season length,

$L_t =$ The general level of the series in the t period,

$b_t =$ Trend component,

$S_t =$ Seasonal component,

$F_{t+m} = m$ forecasting value for the next period,

α (level), β (trend), γ (season) smoothing constants.

Since the data used in the study are under the influence of seasonal fluctuation and seasonal patterns have different magnitudes in periods, "Multiplicative-Seasonal" models were applied in practice, and additive models were not included in the analyses. In order to determine the appropriate model, statistical significance of t-tests belonging to parameter estimates of the models, error squares and MAPE values of the "Simple Seasonal, Multiplicative-Seasonal Exponential Trend, Multiplicative-Seasonal, Damped Trend and Multiplicative-Seasonal Linear Trend (Holt-Winter's)" models were examined. As a result of the evaluations, it was seen that the optimum exponential smoothing model is the "Multiplicative-Seasonal Holt-Winter's" model for the series. In the model, season factor values obtained by seasonal decomposition are used as seasonal factors. Model summary is given in Table 3.

Table 3. **Exponential Smoothing Model Parameters**

Parameter	Estimate	SE	t	Sig.
Alpha (Level)	.041	.011	3.674	.000
Gamma (Trend)	.005	.001	4.083	.000
Delta (Season)	.998	.074	13.498	.000

Source: Own calculations.

JOURNAL OF TOURISM AND SERVICES
Issue 21, volume 11, ISSN 1804-5650 (Online)
www.jots.cz

Initial values of the model are calculated as follows:
L_S = 608,437.68155 (Level)
B_S = 4,814.28919 (Trend)

2.5. Box-Jenkins Method

The Box-Jenkins approach is a well-established and commonly used method, especially in short and medium-term time series forecasting. It is an assumption of the method that the data set applied in this method, which provides successful results in short and medium term modelling and prediction studies, is a discrete and stationary data set consisting of observation values with equal time intervals. The basic principle of the Box-Jenkins methodology is based on its value in any period of time series, a combination of past observations and error terms (Gaspareniene, 2018). Since Box-Jenkins models can only be used in stationary series, stationary condition is important in determining the appropriate model group (Anvari et al., 2016). Data used in practice, especially financial and economic data, often do not meet stationary conditions. The stationarity of this type of data is disrupted by trend, season and cyclical fluctuations and random factors. In general, the expression of the ARIMA (p.d.q) model is as follows:

$$w_t = \phi_1 w_{t-1} + \phi_2 w_{t-2} + ... + \phi_p w_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - .... - \theta_q \varepsilon_{t-q}$$

The tendency of time series, which consist of values observed in weekly, monthly or quarterly time periods, to reach the highest and lowest level in the same periods of successive years or months, indicates the presence of the seasonal component in the data. The representation of the seasonal ARIMA model is ARIMA (p,d,q)(P,D,Q)$_S$ or simply SARIMA. Here, P represents the degree of seasonal auto regression (SAR), D represents the number of seasonal differences, Q is the order of the seasonal moving average (SMA) model, and S is the seasonal period. Seasonal ARIMA model in P, D and Q orders with back shift processor is expressed as:

$$\Phi_p(B^s)\Delta_s^D y_t = \Theta_Q(B^s)\varepsilon_t$$

In this model, Δ_S represents the seasonal difference operator and S is the seasonal period, and S = 12 for monthly data and S = 4 for quarterly data. In the model, the operator Δ^D specifies the degree (D) of the seasonal difference of the data. Non-stationary data is transformed into stationary data with the differencing process indicated byΔ_S^D. The term Φ_p in the model refers to the seasonal auto regression (SAR) parameter, Θ_Q is the seasonal moving average (SMA) parameter, and y_t is the non-stationary series. The Box-Jenkins method uses a four-step, and iterative process to determine a suitable model among the candidate model options. These stages are briefly; model identification, parameter estimation, diagnostic control and forecasting. The method, combines autoregressive, moving average differencing/integration procedures and finally tests the model for statistical validity. In the model selection, factors such as the fact that the data meet the conditions of stationarity and whether they are affected by the seasonal component are determinants. The basic strategy in the model building process is based on the principle of "parsimony" (Moeeni & Bonakdari, 2016). Therefore, it is essential to determine the time series characteristics of the data used first in the modelling and prediction process with the Box-Jenkins method, and to create candidate models suitable for the characteristics of the data and the most appropriate and statistically significant model among them. As multiple candidate models can be identified and predicted satisfactorily, certain criteria must be applied to select the final Box-Jenkins model. The model criteria selection step includes: (1) statistical significance of predicted parameters, (2) residual sum of squares, (3) Akaike Information Criterion (AIC) or Schwarz Bayesian Information Criteria (SBC or BIC).

In the application of the Box-Jenkins method, firstly, the stationarity of the data was tested. The stationarity analysis of the data was carried out with the Augmented Dickey-Fuller (ADF) test. As a result of the ADF analysis, it was determined that there is a high association between neighbouring seasonal observations in the series and the series is out of stationary. As a result of the analysis, the trend stationarity has been provided by taking the first-order seasonal (D=1) difference. Natural logarithm was taken to ensure variance stationarity. The orders of model processes were determined by analysing the auto correlation (ACF) and partial auto correlation (PACF) functions calculated from the data. As a result of the attempts to establish various models, it was determined that the Box-Jenkins model suitable for the series is "Multiplicative-Seasonal $(0,0,11)(0,1,1)_{12}$" model. The final parameter estimates of the determined model and the general summary of the model are given in Table 4. When the table is examined, it is seen that the t-tests of the parameter estimates in the final SARIMA model are significant at the 0.05 significance level.

Table 4. **Summary of SARIMA Model**

Variable	**Estimate**	**Std. Error**	**t-Statistic**	**Prob.**
Constant	0.074	0.008	9.650	0.0000
MA(11)	-0.550	0.074	-7.422	0.0000
SMA(1)	0.356	0.078	4.552	0.0000
R-squared	0.993	**Schwarz Bayes Criterion**		23.035
F-statistic	81.1478	**Prob. (F-statistic)**		0.000000
Seasonal Differencing: 1	**Non-Seasonal Differencing**: 0	**Transformation:** Natural Logarithm		

Source: Own calculations.

After model determination and parameter estimates, residues of the model were analysed. The Ljung-Box (Q^*) statistics were computed for checking residuals. The Ljung-Box statistics, for the seasonal series calculating as follows:

$$Q^* = n(n+2)\sum \frac{r_k^2}{n-k} \sim \chi^2(k-p-q-P-Q)$$

It is a diagnostic measure of white noise for a time series, assessing whether there are patterns in a group of autocorrelations under the hypotheses:

H_0: ACFs are not significantly different than white noise ACFs (i.e., ACFs = 0).

H_1: ACFs are statistically different than white noise ACFs (i.e., ACFs ≠ 0).

In the calculations made for the 12th, 24th and 36th delays of the model's series of residuals, it was found to be $Q^* < X^2$, therefore, null hypothesis of 0.05 significance level was accepted. The calculated Q^* statistics show that there is no significant auto correlation between the model's residuals; confirms that the series has a random process and therefore the suitability of the determined model.

3. Empirical Findings

The empirical results of the forecasting performances of ANN, Holt-Winter's Exponential Smoothing and Box-Jenkins ARIMA models examined are given in Table 5. The ex-post forecasting performances of different forecasting models indicated that Multilayer Perceptron (MLP)-ANN model outperforms the other models with the smallest MAPE of 6.52. According to the criteria proposed by Lewis, it can be said that all applied models successfully produce high accuracy estimates, since the MAPE values of each model are less than 10%. Low MAPE indicates that the deviations between the predicted values obtained from the model and the actual values are very small.

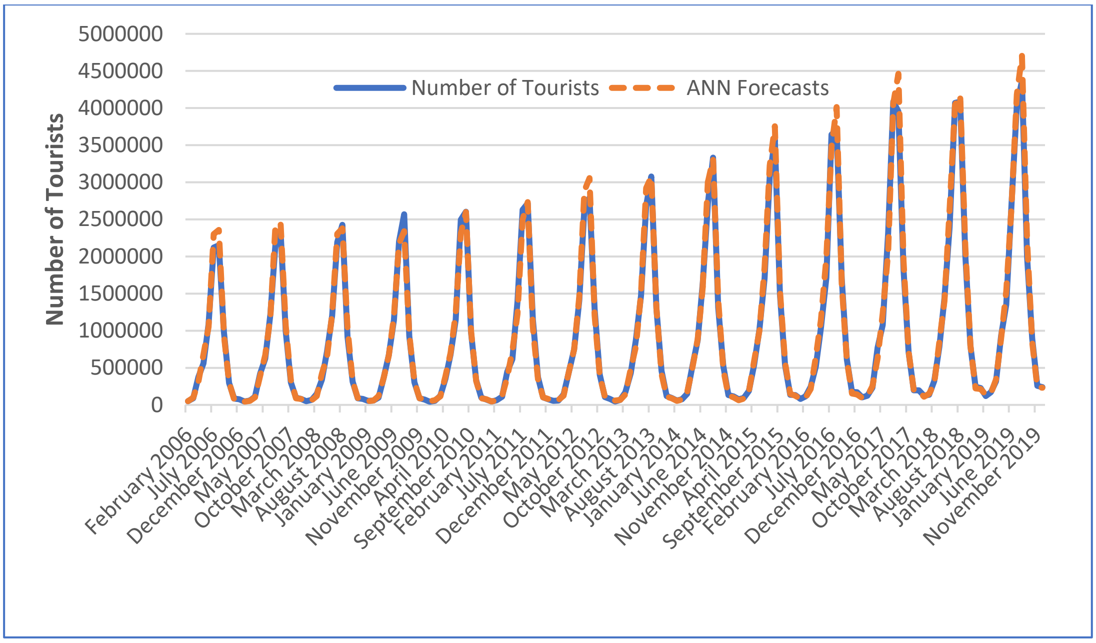
Table 5. **Accuracy Comparisons of Forecasting Models**

Forecast Model	MAPE (%)
[4-3-1] architecture Multilayer Perceptron (MLP)	6.52
SARIMA(0,0,11)(0,1,1)$_{12}$	7.49
Holt-Winter's Multiplicative-Seasonal	7.70

Source: Own calculations

When the results of the forecasting accuracy of each model tested in the study are examined, it is possible to say that all the applied models produce highly successful forecasting results. As stated earlier, in the classification in the prediction modelling literature, models with 50% MAPE statistic value are misidentified models, models with 20% margin of error are models with acceptable accuracy, and models with 10% or less error margin are classified as highly accurate models.

Figure 4. **Time Graph of Inbound Tourist Flows Series and ANN Forecasts**



Source: Own calculations

The Box-Jenkins ARIMA method has been successfully applied in time series forecasting studies due to its advantages such as following a gradual path in achieving the most appropriate solution, controlling the candidate model studied at all levels, providing the opportunity to model according to the characteristics of the data and especially the short and medium term prediction successes have been revealed through the studies conducted. Similarly, Holt-Winter's seasonal exponential smoothing method has also been successfully applied to modelling data under the influence of the trend and seasonal component. However, the model with the lowest margin of error among all tested models is the ANN model. ANNs can learn and generalize the nonlinear patterns in the data and thus create solutions with a reasonable margin of error for the problems that they have not encountered before (Kolkova, 2020; Vrbka, 2020). These findings reveal the ability of ANNs, a machine learning technique, to learn complex and nonlinear models effectively, which is a clear advantage over linear models. For this reason, ANN is also known as a successful method in forecasting time series. Considering both the empirical findings obtained from this study and previous studies on tourism forecasting, it can be seen that ANN models that do not have any negativities (such as over-training, faulty architecture) produce successful forecasting results when compared with results generated by conventional statistical methods.

The time graph of the original foreign tourist arrivals series with the forecasted observation values generated by Multilayer Perceptron (MLP) ANN model is given in Figure 4. When the chart is examined, it is seen that the forecasted values series is in harmony with the real values series and the deviations remain at a very small level. These results demonstrate that constructed Multilayer Perceptron (MLP) model of [4-3-1] architecture can be successfully applied in forecasting future (ex-ante) tourist flows to Croatia.

4. Conclusions and Implications

Travel and tourism, one of the world's major economic sectors, contributes significantly to GDP, creates employment, and contributes to the economic welfare of countries with its invisible export effect. Despite the various crises and adversities experienced, tourism has been among the rapidly growing sectors in the world in recent years. As a result of this growth, the tourism sector has gained a monetary and mass feature today. Therefore, many developed or developing countries attract international visitors in an increasing global competitive environment; providing foreign exchange input, which is a factor in their development and growth; it is in a fierce competition for increasing international tourism revenues and opening new business opportunities. As in many developing countries, the tourism sector is one of the driving forces of the economy in Croatia. Tourism demand forecasts are one of the factors they need for the future planning of the enterprises operating in the tourism sector, governments and local management bodies. In order to accurately administrate production planning, pricing, promotion and strategic marketing programs, labour and capital resources, accurate and reliable demand forecasts are needed. In fact, forecasting is a significant part of the whole planning process in the tourism sector. Accurate estimation of tourism demand is important for the tourism industry because it can help reduce risk and uncertainty as well as effectively provide basic information for better tourism planning. Tour operators, travel agencies, hospitality businesses, airlines and other sectors associated with the tourism industry also need reliable demand forecasts to plan their business strategies and operations. Tourism forecasts are also used for predicting traffic flows, occupancy rates, visitor spending, for new operations and budget planning. Modelling and forecasting studies in tourism has a history of more than half a century and is one of the long-standing areas of research in tourism economics. However, forecasting is often carried out under uncertain conditions, so modelling and forecasting developments in tourism is a difficult task for both practitioners and academics. From a practical forecast perspective in the tourism industry under uncertainty, the need for short-term tourist arrival estimates is becoming increasingly important. This study employs Artificial Neural Networks, Box-Jenkins and Exponential Smoothing methods for forecasting monthly tourist arrivals via the model providing the highest accuracy. The forecasting performances of various models used in the study indicated that Multilayer Perceptron Feedforward (MLP) ANN model outperforms the Box-Jenkins and Exponential Smoothing models with the smallest MAPE. The ANN's has been successfully used for modelling and forecasting demand in a number of fields such as tourism, manufacturing, finance and aviation sectors. ANN is basically a technology used for prediction, clustering, classification and warning against abnormal patterns. The ability to learn examples is one of the most prominent features of neural networks in applications. Considering the results obtained not only with this study but also with many studies, it was observed that problem-free ANN models such as over-training and structural failures generate predictions close to real values. Study results are practically high. Due to the perishable nature of tourism services, accurate and reliable forecasts are of great importance for stakeholders and decision makers in the tourism industry. The contributions of this study to the literature can be summarized as follows. Firstly, the number of studies which modelling and forecasting monthly inbound tourism demand to Croatia by ANNs and conventional methods are still limited and this study aims to fill this gap. Secondly, when the tourism demand forecasts for 2020 are produced by the help of the model developed, it will be enlightening to reveal possible losses due to the Covid-19 outbreak, which continues its impact worldwide.

Forecasting is often carried out under uncertain conditions, so modelling and forecasting developments in tourism is a difficult task for both practitioners and academics. The epidemic of "New Type Coronary Virus (Covid-19)", which emerged in China and spread throughout the world, started to have profound effects on the tourism sector. According to World Tourism Organization (UNWTO, 2020), 209 countries in the world imposed travel restrictions as a precaution against the epidemic between January 2020 and 6 April 2020, this number corresponds to 96% of all destinations in the world. Travel restrictions, which countries have started to implement, have brought tourism and related transportation activities to a halt. In the press release of the UNWTO on 7 May 2020, it was stated that the number of tourists in the world decreased by 22 percent in the first three months of 2020 compared to the same period of the previous year, and this decline may vary between 60 and 80% by the end of 2020. UNWTO made predictions based on three different scenarios prepared for 2020. Accordingly, if international travels start in July, 58% of the international travel market; 70% if it starts in September; It is predicted that it will shrink by 78% in case it is expected to start in December. Therefore, plans and programs prepared by different institutions can be revised in the light of the possible scenarios envisaged by the UNWTO. This study has some limitations, and some areas are open for future research. Firstly the study was carried out for modelling and forecasting inbound tourism demand and it can be extended to focus on both inbound and domestic tourism demand. Second, there are some alternative approaches that can be used in modelling and forecasting tourism demand. For academic studies to be prepared prospectively, comparative studies in which various methods of Artificial Intelligence such as Adaptive Neural Fuzzy Inference System (ANFIS), Support Vector Machines (SVMs), Fuzzy Logic and Genetic Algorithms together with the methods used in this study can be proposed to researchers. Additionally, comparing forecasting accuracy with other forecasting techniques and other accuracy measurement dimensions can yield beneficial results. Witt et al. (1992) argued that domestic tourism demand is less responsive than international (inbound and outbound) tourism demand to external factors such as exchange rate fluctuations, epidemic diseases and international political events, and therefore likely to be less volatile. Therefore, econometric methods that comparatively deal with the relationship of domestic and inbound tourism demand using other variables can be considered as additional extensions of this study. Considering the limited number of studies on modelling and forecasting tourism demand in Croatia, it can be said that the proposed studies will form the basis for the planning activities of the practitioners and public administrators operating in tourism sector and will contribute to filling the gap in field.

References

1. Anvari, S., Tuna, S., Canci, M. & Turkay, M. (2016). Automated Box–Jenkins Forecasting Tool With An Application For Passenger Demand In Urban Rail Systems. *Journal of Advanced Transportation*, 50, 25-49. doi: 10.1002/atr.1332
2. Apergis, N., Mervar, A. & Payne, J.E. (2015). Forecasting Disaggregated Tourist Arrivals in Croatia: Evidence From Seasonal Univariate Time Series Models. *Tourism Economics*, 23(1), 78-98. doi: 10.5367/te.2015.0499
3. Bahovec, V. & Erjavec, N. (1999). ARIMA Model Broja Noćenja Turista u Republici Hrvatskoj. *Ekonomski Pregled*, No. 7-8, 35-70.
4. Baldigara, T. (2013). Forecasting Tourism Demand in Croatia: A Comparison of Different Extrapolative Methods. *Journal of Business Administration Research*, (2)1, 84-92. doi:10.5430/jbar.v2n1p84
5. Baldigara, T. & Mamula, M. (2015). Modelling international tourism demand Using Seasonal ARIMA Models. *Tourism and Hospitality Management*, 21(1), 19-31.
6. Bayramoğlu, M.F. & Başarir, Ç. (2018). International Diversified Portfolio Optimization with Artificial Neural Networks: An Application with Foreign Companies Listed on NYSE. In Dileep

Kumar G. (Ed.), *Machine Learning Techniques for Improved Business Analytics.* (pp. 201-223), Hershey: IGI-Global.
7. Chen, B. F., Wang, H. D. & Chu, C.C. (2007). Wavelet and Artificial Neural Network Analyses of Tide Forecasting and Supplement of Tides Around Taiwan and South China Sea. *Ocean Engineering*, 34(16), 2161-2175. doi: 10.1016/j.oceaneng.2007.04.003
8. Croatian Bureau of Statistics (2020). *Sources and methods of data collection*, Retrieved July 24, 2020, from https://www.dzs.hr/Eng/DBHomepages/Tourism/methodology.htm
9. Cuhadar, M., Güngör, İ. & Göksu, A. (2009). Forecasting Tourism Demand by Artificial Neural Networks and Time Series Methods: A Comparative Analysis in Inbound Tourism Demand to Antalya. *Süleyman Demirel University*, *Faculty of Economics and Administrative Sciences Journal*, 14(1), 99-114.
10. Dogru, T. & Sirakaya-Turk, E. (2018). Modeling Turkish Outbound Tourism Demand Using a Dynamic Panel Data Approach. *Tourism and Hospitality Research*, 18(4), 411–414. doi:10.1177/1467358416663822
11. Egrioglu, E., Aladag, Ç.H., Yolcu, U., Baş, E. & Dalar, A.Z. (2017). A New Neural Network Model with Deterministic Trend and Seasonality Components for Time Series Forecasting. In Aladağ Ç.H. (Ed.) *Advances in Time Series Forecasting*, Vol. 2, (pp. 76-92). Sharjah, UAE: Bentham Science Publishers.
12. Frechtling, D.C. (2001). *Forecasting Tourism Demand: Methods and Strategies*, Butterworth-Heinemann, Oxford.
13. Gaspareniene, L., Remeikiene, R., Sadeckas, A., & Ginevicius, R. (2018). Gold Investment Incentives: An Empirical Identification of the Main Gold Price Determinants and Prognostication of Gold Price Future Trends. *Economics and Sociology*, 11(3), 248-264. doi:10.14254/2071-789X.2018/11-3/15
14. Goh, C. & Law, R. (2011). The Methodological Progress of Tourism Demand Forecasting: A Review of Related Literature. *Journal of Travel & Tourism Marketing*, 28(3), 296-317. doi:10.1080/10548408.2011.562856
15. Golam, K. & Hasin, M.A.A. (2013). Comparative Analysis of Artificial Neural Networks and Neuro-Fuzzy Models for Multicriteria Demand Forecasting. *International Journal of Fuzzy System Applications*, 3(1), 1-24. doi: 10.4018/ijfsa.201301010
16. Höpken, W., Eberle, T., Fuchs, M. & Lexhagen, M. (2020). Improving Tourist Arrival Prediction: A Big Data and Artificial Neural Network Approach. *Journal of Travel Research*, 1–20. doi:10.1177/0047287520921244
17. Kolkova, A. (2020). The Application of Forecasting Sales of Services to Increase Business, Competitiveness. *Journal of Competitiveness*, 12(2), 90–105. doi 10.7441/joc.2020.02.06
18. Kon, S. C. & Turner, W. L.(2005). Neural network forecasting of tourism demand. *Tourism Economics*, 11(3), 301-328. doi: 10.5367/000000005774353006
19. Law, R. & Pine, R. (2004). Tourism Demand Forecasting for the Tourism Industry: A Neural Network Approach. In Zhang, G.P. (ed.) *Neural Networks In Business Forecasting* (pp. 121-141), Idea Group Publishing, Hershey: PA.
20. Lewis, C. D. (1982). *Industrial And Business Forecasting Methods: A Practical Guide To Exponential Smoothing And Curve Fitting.* London: Butterworth Scientific.
21. Lim, C. (2006). A survey of tourism demand modelling practice: issues and implications, In Dwyer, L. and Forsyth, P. (Eds.), *International Handbook on the Economics of Tourism*, (pp. 45-72), UK: Edward Elgar Publishing Inc.
22. Makridakis, S., Wheelwright, S.C. & Hyndman, R.J. (1998). *Forecasting: Methods and Applications* (3rd Edition). New York: John Wiley and Sons.

23. Mervar, A. & Payne, J. E. (2007). Analysis of Foreign Tourism Demand for Croatian Destinations: Long-Run Elasticity Estimates. *Tourism Economics*, 13(3), 407–420. doi:10.5367/000000007781497764
24. Moeeni, H. & Bonakdari, H. (2016). Forecasting Monthly Inflow with Extreme Seasonal Variation Using the Hybrid SARIMA-ANN Model. *Stochastic Environmental Research Risk Assessment*, 31(8), 1997-2010. doi: 10.1007/s00477-016-1273-z
25. Moreno, J.J.M., Poll, A.P. & Gracia, P.M. (2011). Artificial Neural Networks Applied To Forecasting Time Series. *Psicothema*, 23(2), 322-329.
26. Önder, I. (2017). Forecasting Tourism Demand with Google Trends: Accuracy Comparison of Countries versus Cities. *International Journal of Tourism Research* 19, 648–60. doi: 10.1002/jtr.2137
27. Orsini, K. & Pletikosa, M. (2019). *Croatia's Tourism Industry-Part 2: Curse or Blessing?.* European Economy Economic Briefs. Publications Office of the European Union, Luxembourg. Retrieved June 22, 2020, from https://ec.europa.eu/info/publications/croatias-tourism-industry-part-2-curse-or-blessing_en
28. Peng, B., Song, H. & Crouch. G. I. (2014). A Meta-analysis of International Tourism Demand Forecasting and Implications for Practice. *Tourism Management*, 45, 181–93. doi:10.1016/j.tourman.2014.04.005
29. Saayman, A. & Saayman, M. (2010). Forecasting Tourist Arrivals in South Africa. *Acta Commercii*, 10(1), 281-293. doi: 10.4102/ac.v10i1.141
30. Smith, K. A. (2002). Neural Networks for Business: An Introduction. In Smith, K.A & Gupta, J.T.D (Eds.). *Neural Networks in Business: Techniques and Applications*, (pp. 1-25). Idea Group Publishing.
31. Song, H. & Turner, L. (2006). Tourism demand forecasting. In Dwyer, L. & Forsyth, P. (Eds.) *International Handbook on the Economics of Tourism*. (pp. 89-114). Cheltenham: Edward Elgar Publishing Limited.
32. Song, H. & Li, G. (2008). Tourism Demand Modelling and Forecasting: A Review of Recent Research. *Tourism Management*, 29(2), 203-220. doi: 10.1016/j.tourman.2007.07.016
33. Song, H., Li, G., Witt, S. F. & Fei, B. (2010). Tourism Demand Modelling and Forecasting-How Should Demand Be Measured?. *Tourism Economics*, 16(1): 63-81. doi:10.5367/000000010790872213
34. Teixeira, J.P. & Fernandes, P.O. (2012). Tourism Time Series Forecast-Different ANN Architectures with Time Index Input. *Procedia Technology*, (5), 445-454. doi:10.1016/j.protcy.2012.09.049
35. Tica, J. & Kožić, I. (2015). Forecasting Croatian Inbound Tourism Demand. *Economic Research-Ekonomska Istraživanja*, 28(1), 1046-1062. doi: 10.1080/1331677X.2015.1100842
36. UNWTO-World Tourism Organization (2020). COVID–19 Related Travel Restrictions. A Global Review for Tourism. First Report as of 16 April 2020. Retrieved June 26, 2020, from https://www.unwto.org/news/covid-19-response-travel-restrictions
37. Vrbka, J. (2020). The use of neural networks to determine value based drivers for SMEs operating in the rural areas of the Czech Republic. *Oeconomia Copernicana*, 11(2), 325–346. doi: 10.24136/oc.2020.014
38. Witt, S. F., Newbould, G. D. & Watkins, A. J. (1992). Forecasting domestic tourism demand: Application to Las Vegas arrivals data. *Journal of Travel Research*, 31(1), 36-41. doi: 10.1177/004728759203100108
39. Wong, B. K., Jiang, L. & Lam, J. (2000). A bibliography of neural network business application research: 1994-1998. *Computers and Operations Research*, 27(11), 1045-1076. doi: 10.1016/S0305-0548(99)00142-2

40. World Travel and Tourism Council (WTTC) - Oxford Economics (2020). Global Economic Impact of Travel & Tourism, UK. Retrieved June 11, 2020, from https://wttc.org/Research/Economic-Impact
41. Yaffee, R.A. & McGee, M. (2000). *Introduction to Time Series Analysis and Forecasting with Applications of SAS and SPSS*. London: Academic Press.
42. Zhang, G. P. & Qi, M. (2005). Neural Network Forecasting For Seasonal And Trend Time Series, *European Journal of Operational Research*, 160(2), 502-514. doi: 10.1016/j.ejor.2003.08.037
43. Zhang, Y., Li, G., Muskat, B. & Law, R. (2020). Tourism Demand Forecasting: A Decomposed Deep Learning Approach. *Journal of Travel Research*, 1-17. doi: 10.1177/0047287520919522

Brief description of Author/Authors:

Murat ÇUHADAR, PhD
Associate Professor at Süleyman Demirel University, Faculty of Economics and Administrative Sciences, Tourism Management Department. Isparta, Turkey. His research interests are in tourism demand forecasting, modelling tourism time series, econometric analyses, Artificial Neural Networks (ANNs) and marketing research in tourism and hospitality.
e-mail: muratcuhadar@sdu.edu.tr, web page: https://w3.sdu.edu.tr/personel/01495