



Forecasting Türkiye's International Tourism Demand *

Zeynep KURTULAY¹, İsmail KIZILIRMAK²

ARTICLE INFO

Background:

Received: 22/12/2023
Accepted: 09/01/2024
Published: 28/03/2024

Keywords:

Demand Forecasting,
International Tourism,
Time Series, Tourism
Demand.

ABSTRACT

The structure of tourist demand is sensitive since it is very easily affected by the consequences of economic, political, and social crises. Since the limited ability to increase tourism supply, it is crucial to analyze the demand structure and develop suitable strategies. This outcome can only be achieved by an accurate and effective demand prediction. There is no singular approach that ensures success in demand forecasting. Hence, to estimate demand accurately, it is advisable to create many models and choose the one with the lowest error rate. This study aimed to develop the best-performing prediction model by using monthly data of international visitors who visited Türkiye from January 2002 to August 2023 and stayed in Tourism Ministry-certified accommodation establishments. Within this framework, the data was first analyzed to identify the trend and seasonal components. Afterwards, various models were employed including Naive III, simple moving average, double-moving average, seasonal exponential smoothing, and artificial neural networks. The data generated by these models has been analyzed by comparing it with the actual data from the last 24 months, using MAPE and RMSE results. According to the research findings, it has been determined that artificial neural networks produce the most accurate results.

Introduction

Demand is defined as "asking, wishing, demanding from someone to do or not do something" (TDK, 2020). In everyday life, it is commonly used to refer to a wish or desire. In order for these wishes and desires to be considered as demands in economic terms, the person making the request must have purchasing power. From this perspective, demand is characterized as the act of providing a product or service to the market at a specific price and within a specific timeframe, and the quantity that customers desire to purchase of that product or service (Dinler, 2006, p.48-49). The demand for tourism is closely associated with the economic definition. Tourism demand refers to the quantity of tourism products that individuals, who have enough free time and financial resources, desire to purchase while visiting a destination other than their place of residence (Öztaş & Karabulut, 2006, p.54; Ünlüönen et al., 2007, p.45).

The tourism industry is very susceptible to both national and international crises (Göktaş, 2023, p.626). Positive impacts on the sector have a slow effect, negative impacts including economic and political crises, wars, and epidemics have an immediate impact on the industry (Edgell, 1993, p.13). Hence, any problems in these industries might impact the demand for tourism. Seasonality is a significant characteristic of tourism demand. Seasonality is defined as recurring patterns that occur within specific time periods on a yearly basis (Lundberg et al., 1995, p.160). When examining Türkiye's tourism demand, it's easy to see that the highest demand is observed over the period

¹Doktora Öğrencisi, İstanbul Üniversitesi, Sosyal Bilimler Enstitüsü, Turizm İşletmeciliği ABD, zeynepkurtulay@gmail.com ORCID: [0000-0003-1983-362X](https://orcid.org/0000-0003-1983-362X)

²Prof. Dr., İstanbul Üniversitesi, İktisat Fakültesi, Turizm İşletmeciliği Bölümü, ikizilirmak@yahoo.com ORCID: [0000-0001-9141-6420](https://orcid.org/0000-0001-9141-6420)
Cite as: Kurtulay, Z.; Kızılırmak, İ. (2024). Forecasting Türkiye's International Tourism Demand. *Journal of Global Tourism And Technology Research*, 5(1), 12-26.

* Bu makale, İstanbul Üniversitesi, Sosyal Bilimler Enstitüsü, Turizm İşletmeciliği ABD, Doktora Programı, doktora tezinden üretilmiştir

from May to September. The widespread popularity of sea-sand-sun tourism in Türkiye is the real cause of this situation.

Another important aspect is the human factor. While the main reason of participating in tourism activities is to explore new destinations, discover new experiences, or relax, the factors that motivate individuals to travel are different from one another. Hence, tourism demand is very sensitive to the effects of human behavior, desires, and requirements, as well as sociocultural and psychological elements.

Individuals' disposable income also has an impact on the tourism demand. People have essential requirements, including food, a place to live and healthcare, that are essential for their continued existence and well-being. They have a limited amount of time and financial resources to fulfill these essential demands. The discretionary income, which is the amount of funds left after fulfilling the basics, impacts people's luxurious purchases. The rise in obligatory necessities forces individuals to postpone their more luxurious desires, such as traveling, exploring new destinations, joining a social group, and discovering new experiences. This condition causes more flexibility in tourism demand.

This study aimed to design models for forecasting that achieves maximum prediction accuracy by using the monthly number of international tourists staying in Tourism Ministry-certified accommodations establishments from January 2002 to August 2023. Within the given context, the data collected were primarily analyzed according to their trend and seasonal component. The data set was further analyzed using several modeling techniques, including Naive III, simple moving average, double moving average, seasonal exponential smoothing, and artificial neural networks. Upon analyzing the data set, it becomes clear that it has been impacted by many economic, political, sociocultural incidents, as well as the continuous Covid-19 pandemic that has evolved over the years. The presence of frequent breaks caused by different factors within a short period of time causes difficulties to analysis. However, it also shows the significance of studying such breaks, as it helps the development of adaptable models that can adjust to various circumstances.

Literature

In order to succeed in achieving tourism demand forecasting, it is fundamental to comprehend the characteristics of tourism demand and the different factors that are influencing it. The following part of the study encompasses an examination of all the factors that influence the demand for tourism, specifically focusing on Türkiye's international tourism demand. Additionally, it discusses the significance of demand forecasting for the tourism sector.

Factors Affecting Tourism Demand

The structure of tourism demand is complicated and influenced by many different factors. The aforementioned factors significantly influence tourists' travel preferences, destination choices, and travel routines. It is possible to collect the factors affecting individuals' tourism demand under five basic headings.

- ✓ **Economic Factors:** Price level, income level, economic distance, transportation costs and exchange rates (Lim, 1997, p.835).
- ✓ **Social Factors:** Age and gender, education and cultural level, occupation and social class, leisure time, family structure, communication difficulties, and urbanization (Barry & O'Hagan, 1972, p.147; Pamukçu et al., 2015, p.466).
- ✓ **Political and Legal Factors:** Terrorism, legal sanctions, political uncertainty, crime rates (Bayram, 2018, p.33).
- ✓ **Psychological Factors:** Personality, fashion, taste and habits, perception of tourism (Pizam & Milman, 1986, p.30-31).
- ✓ **Other Factors:** Epidemic diseases, advertising and promotion, globalization and technology (Taptik & Keleş, 1998, p.27).

Economic factors have a major impact on the tourism demand. Raising prices of tourism products lowers demand, but raising individuals' income improves disposable income and may causes an increase in demand. In the same way, widening the distance between destinations causes a rise in the cost of transportation and a longer duration of time required to reach there as well. This situation causes a decrease in demand. The exchange rate has an impact on the ability of individuals to purchase goods and services. This situation causes a decrease in demand. The exchange

rate has an impact on the ability of individuals to purchase goods and services. When the currency of the nation where tourists come from increases in value, it causes a rise in demand because tourists have greater disposable income. On the other hand, when the currency decreases in value, demand drops.

Social factors have an important effect on the demand for tourism. Although individuals between the ages of 15 and 24 tend to travel more frequently (Köroğlu & Güleç, 2008, p.42), age-related health issues may cause difficulties in travel and influence the choice of destination. Gender roles sometimes support travel; however, this is affected by social and cultural factors (Barry & O'Hagan, 1972, p.147). Leisure time has a positive effect on tourism demand, but if there are communication difficulties between tourists and local people at the destination, it might lead to a decline in demand.

The primary goal of tourists when participating in tourism activities is to ensure the safety of their lives and belongings. Hence, terrorist attacks, increased crime rates, political uncertainty, or important social incidents in a destination increase the sense of dangers for tourists and might lead to a decline in tourism demand. Moreover, the bureaucracy of the official procedures, such as applying for visas and passports, decreases the demand for tourism.

Psychological factors affect individuals' habits and willingness to travel (Cruz-Milan, 2018, p.50). Advertisements about a destination can lead to increased demand. An increase in advertisements highlighting a destination can cause to an rise in its popularity. And when a place becomes popular, the demand for that destination rises. Moreover, tourism demand is higher in the cultures that have a tourism perception.

The rise in epidemic diseases strengthens the perceived risk by tourists and leads to a decrease in demand. An example of that situation currently is the Covid-19 pandemic. Due to the pandemic, some nations have implemented curfews and suspended flights, and there has been a serious decrease in tourism demand. Hence, minimizing the negative effects of the decline in tourism demand, as a result of the pandemic, has been identified as the main priority for many countries.

Türkiye's International Tourism Demand

After understanding the positive effects of tourism, such as providing job opportunities, decreasing economic inequalities across regions, and leading to foreign exchange revenue, many governments are making attempts to get a share of the tourism market. In response to the rapid growth of globalization throughout the 1980s, several nations attempted to increase their export capabilities. Nevertheless, nations relying on agriculture and industry have failed to achieve their desired levels of economic expansion. Tourism has emerged as a significant alternative, particularly for nations with limited resources.

International tourism demand is showing a consistent upward tendency, especially ever since the 1980s, although experiencing sometimes fluctuations. Since the 2000s, Türkiye is experiencing an upward trend in demand in the global tourism industry. Türkiye became the eighth most visited country in 2017 (UNWTO, 2018), the sixth most visited country in 2019 (UNWTO, 2021) and the fourth most visited country in the world in 2022 (UNWTO, 2023). Alongside the growth in tourist numbers, Türkiye's tourism revenue is also rising. The tourism revenue in 2017 reached over \$27 million, which increased to approximately \$31 million in 2018, and then increased to about \$46 million in 2022, according to the UNWTO (2023). Despite being in the post-pandemic period, the substantial rise in 2022 is noteworthy.

Tourism revenue also has an important part of exports. In the 1960s, the percentage of tourism revenue in Türkiye's exports was between 2-3%. However, by the early 2000s, the percentage had increased by approximately 30% (TÜRSAB, 2021a). Tourism revenues have an essential part in reducing the current account deficit. Short and medium-term growth in industrial and agriculture-based exports is impractical, particularly for developing countries. Hence, the tourism industry plays a crucial role in providing direct foreign exchange revenue. Tourism revenues have an essential role in minimizing the current account deficit of Türkiye. In 2011, at its lowest point, the current account deficit was covered at a rate of 34.14% (TÜRSAB, 2021b).

The Importance of Demand Forecasting in Tourism

Many countries attempt to grow their market share globally because of realizing the impacts of tourism. Nevertheless, tourism demand is very sensitive to the effects of economic, political, and social changes, both locally

and globally. Additionally, the sector's dependency on financial investments in fixed assets and infrastructural prospects requires investors to be cautious. According to Şenel (2007, p.7), the mining companies have a fixed capital investment rate of 64%, while the iron and steel industry has a rate of 50%, and the chemical sector has a rate of 42%. The accommodation enterprises, on the other hand, have the highest rate of fixed capital investment, reaching as high as 94%. Furthermore, the costs that include water, electricity, amortization, rent, and workers' expenditures for accommodation enterprises are not influenced by the rate of room occupancy. In other words, regardless of the low occupancy rate in accommodation enterprises, their fixed expenditures are high. Hence, it is important to make the most suitable financial decision.

To ensure effective investment decisions, it is important to thoroughly analyze and identify the structure and characteristics of tourism demand. Investments accomplished without consideration of the characteristics of demand not only outcome in financial losses, but also have the potential to lead to a variety of environmental, political, or social problems in the region. Several potential problems may arise, including the establishment of construction zones for the accommodation enterprise, surpassing the region's capacity, increasing solid waste, noise pollution, raised population growth, lowering local welfare, and negative effects on tourism perceptions of local people. In order to minimize these problems, it is essential to make a reliable and accurate forecasting of tourism demand. An accurate demand forecasting plays an important role in minimizing the risks assumed by investors in both the private and governmental sectors.

Nevertheless, it is important to keep in mind that there is no generally acceptable approach for making completely accurate demand forecasting. The choice of method should be customized to the particular requirements of the problem and data sets (Soysal & Ömürgönülşen, 2010, p.129). The main purpose of demand forecasting is to select the most suitable method given the available data and ascertain the alternative that shows a minimum difference from the actual. Upon reviewing the relevant literature, studies on demand forecasting may be categorized into qualitative, quantitative, and hybrid approaches (Krajewski, et al, 2010, p.566).

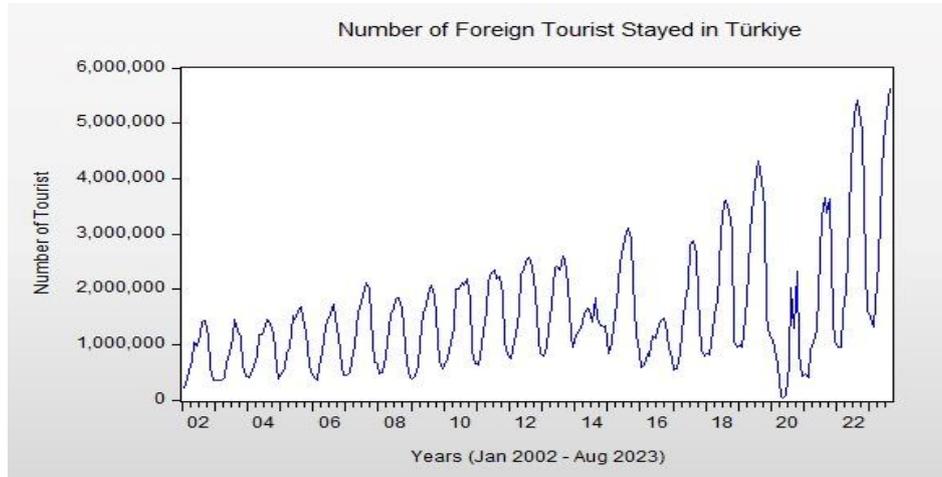
Method

The concept of forecasting tourism demand increased in popularity from 1960 to 2002. A total of 420 studies were published in both national and international literature within this specific time frame. Published studies commonly use many data sources, but the most common method is to use indicators such as the arrival of tourists, tourism revenue, number of overnight stays, or occupancy rates to measure demand (Lim, 1997). In this study, the number of tourists was used as a measurement of tourism demand.

The data set obtained from the monthly numbers of foreign tourists who came to Türkiye and stayed in Culture and Tourism Ministry-Certified accommodation enterprise, between January 2002 and August 2023 was examined in terms of its components. For the following phase, demand forecasting was executed by applying several techniques such as basic and double moving averages with distinct time periods, exponential smoothing approach, and artificial neural networks with diverse topologies. At the final phase, the predicted values were compared to the actual values acquired over the last 24 months, and the prediction accuracy of the models was measured. The graph showing the data for analysis is shown below.

As to Witt and Witt (1995), MAPE values that are less than 10% are very accurate, and values between 10-20% can be classed as accurate models. Alternatively, Lewis (1982) categorizes MAPE values below 10% as highly satisfactory, values between 10-20% as satisfactory, values between 20-50% as acceptable, and values beyond 50% as inaccurate.

Figure 1. Number of Foreign Tourists Stayed in Türkiye Between Jan 2002-Aug 2023



The models' performance was compared using MAPE and RMSE measurements. Formulas for calculating MAPE and RMSE values are given below.

Mean Absolute Percentage Error (MAPE)

$$\frac{1}{n} \sum_{t=1}^n \frac{|e_t|}{y_t} * 100 \tag{1}$$

Root Mean Square Error (RMSE)

$$\sqrt{\frac{1}{n} \sum_{t=1}^n e_t^2} \tag{2}$$

Review of Research Data

When reviewing Figure 1, it's visible that there is an overall upward trend in the number of tourists. However, there are changes in the number of tourists from month to month. In other words, while Türkiye's tourism demand is affected by seasonal factors, it is now experiencing a growing momentum. Tourist numbers typically begin to rise in April and peak between June and August. When examining the details of the graph, it's easy to see that the data between April and September, which showed the peak value in 2014, remained below the overall trend. The cause of this may be related to the controversial political atmosphere at that period, the terrorist attacks committed by ISIS, and the Gezi Park protests that got a great deal of attention in the international media. In 2016, tourism demand stayed to be lower than the overall trend, although seeing a spike in 2015. The decrease in tourism demand in 2016 can be explained to a several factors, including the bomb attacks that were committed by the PKK terrorist organization on popular destinations such as Sultanahmet, Taksim, and Istanbul Airport, as well as the attempted coup that occurred on July 15 and received attention by the global media.

Another major break in the chart was observed in 2019. The cause of this situation can be explained to the Covid-19 epidemic, which has quickly reached a worldwide scale. Despite the Covid-19 actions, the effect of aircraft cancellations, enforcing of curfews, and closure of borders have significantly reduced the demand for tourism. Upon analyzing the 2022 data, it becomes evident that the values during the summer months have the biggest value on the graph. This can be explained by the fluctuating exchange rate of the Turkish Lira in comparison to the euro and the dollar.

Another major break in the chart was observed in 2019. The cause of this situation can be explained to the Covid-19 epidemic, which has quickly reached worldwide scale. Despite the Covid-19 actions, the effect of aircraft cancellations, enforcing of curfews, and closure of borders have significantly reduced the demand for tourism. Upon

analyzing the 2022 data, it becomes evident that the values during the summer months have the biggest value on the graph. This can be explained to the fluctuating exchange rate of the Turkish Lira in comparison to the Euro and the Dollar.

For understanding the path of the time series in upcoming periods, it is essential to determine the trend component. To achieve this goal, trend analysis was used using the least squares method. The generated findings were examined using the F test, and the time variable was analyzed using the t-test. The data was analyzed using the SPSS 25.0 software. All the results are displayed in the table below

Table 1. Results of Trend Analysis

Linear Model			R ² = ,226	Adjusted R ² =,223	Std. Error of the Est.= 932660,718
	Sum of Squares	df	Mean Square	F	Sig.
Regression	6,552E+13	1	6,552E+13	75,323	,000
Residual	2,244E+14	258	8,699+11		
Total	2,899E+14	259			
Coefficient					
	Unstandardized B	Coefficient Std. Error	Standardized Coefficients Beta	t	Sig
Case Sequence	6688,398	77,651	,475	8,679	,000
(Constant)	668105,868	116016,828		5,759	,000
Logarithmic Model		R=,410	R ² = ,168	Adjusted R ² =,165	Std. Error of the Estimate= 967115,409
	Sum of Squares	df	Mean Square	F	Sig.
Regression	4,863E+13	1	4,863E+13	51,996	,000
Residual	2,413E+14	258	9,353E+11		
Total	2,899E+14	259			
Coefficient					
	Unstandardized B	Coefficient Std. Error	Standardized Coefficients Beta	T	Sig
Case Sequence	450470,845	62471,328	,410	7,211	,000
(Constant)	-519922,12	292026,418		-1,780	0,076
Quadratic Model		R=,486	R ² = ,236	Adjusted R ² =,230	Std. Error of the Estimate= 928323,279
	Sum of Squares	df	Mean Square	F	Sig.
Regression	6,846E+13	2	3,423E+13	39,723	,000
Residual	2,215E+14	258	8,618E+11		
Total	2,899E+14	259			
Coefficient					
	Unstandardized B	Coefficient Std. Error	Standardized Coefficients Beta	t	Sig
Case Sequence	1175,841	3079,417	0,84	,382	,703
Case Sequence**2	21,121	11,427	,405	1,848	,066
(Constant)	908820,874	174053,655		5,221	,000

Cubic Model		R=,370	R ² = ,317	Adjusted R ² =,134	Std. Error of the Estimate= 907963,440
	Sum of Squares	df	Mean Square	F	Sig.
Regression	7,890E+13	3	2,630E+13	31,901	,000
Residual	2,110E+14	258	8,244E+11		
Total	2,899E+14	259			
Coefficient					
	Unstandardized B	Coefficient Std. Error	Standardized Coefficients Beta	t	Sig
Case Sequence	25874,979	7568,194	1,839	3,419	,001
Case Sequence**2	-215,007	67,311	-4,118	-3,194	,002
Table 1: Results of Trend Analysis (Continue)					
Case Sequence**3	,603	,170	2,863	3,557	,000
(Constant)	366465,716	228526,082		1,604	,110
Growth Model		R=,370	R ² = ,137	Adjusted R ² =,134	Std. Error of the Estimate= ,681
	Sum of Squares	df	Mean Square	F	Sig.
Regression	19,011	1	19,011	40,954	,000
Residual	119,765	258	,464		
Total	138,776	259			
Coefficient					
	Unstandardized B	Coefficient Std. Error	Standardized Coefficients Beta	t	Sig
Case Sequence	,004	,001	,370	6,399	,000
(Constant)	13,546	,085		159,835	,000
Exponential Model		R=,370	R ² = ,137	Adjusted R ² =,134	Std. Error of the Estimate= ,681
	Sum of Squares	df	Mean Square	F	Sig.
Regression	19,011	1	19,011	40,954	,000
Residual	119,765	258	,464		
Total	138,776	259			
Coefficient					
	Unstandardized B	Coefficient Std. Error	Standardized Coefficients Beta	t	Sig
Case Sequence	,004	,001	,370	6,399	,000
(Constant)	76081,462	64757,760		11,799	,000
Logistic Model		R=,370	R ² = ,137	Adjusted R ² =,134	Std. Error of the Estimate= ,681
	Sum of Squares	df	Mean Square	F	Sig.
Regression	19,011	1	19,011	40,954	,000
Residual	119,765	258	,464		
Total	138,776	259			
Coefficient					
	Unstandardized B	Coefficient Std. Error	Standardized Coefficients Beta	t	Sig
Case Sequence	,996	,001	,691	1776,280	,000
(Constant)	1,309E-6	,000		11,799	,000

When looking at Table 1, it was found that the F test and t test findings, which indicate the overall effectiveness of the models, were statistically significant at the 0.05 significance level. Say it in another way, both the created models and the fixed coefficients are statistically significant, which shows a positive trend in the series. The computer was used to determine the linear trend equation for the generated data set, which is provided below.

$$\hat{y}_t = 219,7x - 8E + 06 \text{ ve } R^2=0,226$$

Then, the seasonal index values of the time series were examined. For this purpose, seasonal decomposition process was applied to the series. As a result of the operations performed (Period+1 – Endpoints Weight by 0.5), it was determined that the series follows a cycle that repeats every 12 months. Seasonal index values and seasonal factors of the series are presented in the table below.

Table 2. *Seasonal Index and Seasonal Factors*

Periods	Months	Seasonal Index	Seasonal Factor
1	January	0,44258	44.3%
2	February	0,47397	47.4%
3	March	0,64811	64.8%
4	April	0,83669	83.7%
5	May	1,18802	118.8%
6	June	1,32350	132.3%
7	July	1,48913	148.9%
8	August	1,63490	163.5%
9	September	1,53016	153.0%
10	October	1,28507	128.5%
11	November	0,66464	66.5%
12	December	0,48324	48.3%

After reviewing Table 2, it's simple to see that the months with the lowest seasonal index are December, January, and February. Additionally, there is an important rise from April, with the maximum level seen in July, August, and September.

Application of Forecasting Methods

This section of the study includes an examination of the performances of several forecasting techniques using monthly data. The performance of the techniques was measured by calculating the Mean Absolute Percentage Error (MAPE) and Root Mean Square Error (RMSE) values of the data generated by the models, using the actual data observed throughout the time period from September 2021 to August 2023.

Naive III Forecast Method

Naive models are sometimes referred to as mechanical forecast methods. This strategy, known as "last period demand" (Çağlar, 2007; Yıldırım, 2019), is considered one of the simplest forecasting methods. This method involves using the last data from a time series as the prediction value for the next period. When reviewing the previously mentioned Figure 1, Table 1, and Table 2, it was found that the series had seasonal characteristics and had an increasing trend. Therefore, the Naive III model was chosen for the series. The formula for the method is given below.

$$y_{t+1} = y_{t-3} + \frac{(y_t - y_{t-1}) + \dots + (y_{t-3} - y_{t-4})}{4} \tag{3}$$

y_t represents the time series at a given time t . y_{t+1} represents the value of the time series in the period $t + 1$. It represents the estimated value at a specific point in time. The analysis generated the data that follows MAPE and RMSE values.

MAPE = 34,41%

RMSE = 179616

Moving Averages Method

The simple moving average approach uses dividing the data in a time series into certain sizes, calculating the average of these groups, and replacing the value in the center of each group with this average (Atlas, 2013, p.143). The researcher determines the size of the group to be generated (n) in this method. When new data is added, the oldest data is excluded from the calculation (Can, 2009, p.24). The approach of calculating the simple moving average was utilized for various periods (n=2, n=3, n=4). The formula used for the application of the method is given below.

$$M_{t+1} = \frac{y_t + y_{t-1} + y_{t-2} + \dots + y_{t-m+1}}{n} \tag{4}$$

M_{t+1} represents the estimated prediction of the observation value for period $t + 1$, whereas y_t represents the observation value in period t . n is the total number of observations included in the moving average calculation. The examination gave the following values for MAPE and RMSE.

	n=2	n=3	n=4
MAPE=	41,09%	55,15%	67,40%
RMSE=	211737	257728	296281

The double-moving average technique is used for time series data that has a trend that is defined as a first or second-degree polynomial. This method involves initially calculating a moving average for the given time series. Subsequently, a second series is generated by calculating the moving average of the previously calculated series (Hanke & Reitsch, 1992, p.137). If the simple moving average is shown as M_t , and the double-moving average is represented as M_t^d , the double moving average is calculated based on the simple moving average of n periods using the formula is given below (Uygur, 2011, p.78-79).

$$M_t^d = \frac{M_t + M_{t-1} + \dots + M_{t-n+1}}{n} \tag{5}$$

The formula used for calculating projected values for the future is as follows:

$$a_t = 2M_t - M_t^d \tag{6}$$

$$b_t = \frac{2}{n-1}(M_t - M_t^d) \tag{7}$$

In the given equations, the symbol a_t represents the constant in the equation, whereas b_t shows the trend. Projected value for time period $t+p$;

$$M_{t+p}^d = a_t + b_{tp} \tag{8}$$

In this context, n represents the number of periods used to calculate the moving average, whereas p represents the number of periods that are to be forecasted (Hanke & Reitsch, 1992, p.137).

The values of a_t and b_t were calculated using equations 6 and 7 mentioned before. However, it was observed that these values became negative over the period from November to April. Given the existence of seasonality in the time series according to tests, an analysis was carried out to see if these values repeated over several periods (n = 1, n = 2, and n = 3). The findings indicate that the values of a_t and b_t once again become negative. Thus, it was determined that the double-moving average approach is unsuitable for this dataset.

Seasonal Exponential Smoothing (Holt-Winters) Method

After analyzing the previously given Figure 1, Table 1, and Table 2, it was determined that the series had seasonal features and showed a to rising trend. The Multiplicative Exponential Smoothing (Holt-Winters) approach was applied because of the time-varying nature of the seasonal movements in the series. The series was analyzed using the seasonal index values provided in Table 2. The formula used for the application of the method is given below.

$$L_t = \alpha \frac{y_t}{s_{t-s}} + (1 - \alpha)(L_{t-1} + b_{t-1}) \tag{9}$$

$$b_t = \beta(L_t - L_{t-1}) + (1 - \beta)b_{t-1} \tag{10}$$

$$s_t = \gamma \frac{y_t}{(L_t)} + (1 - \gamma)s_{t-s} \tag{11}$$

$$F_{t+h} = (L_t + hb_t)s_{t-s+h} \tag{12}$$

The equations have three main variables: L_t , which represents the general level of the series; b_t , which represents the trend; and s_t , which represents the seasonal variables. The smoothing constant, γ , is used to smooth the original series by applying a smoothing technique. The seasonal smoothing constant is used to correct the seasonal component, whereas β is referred to as the linear smoothing constant for correcting the trend. The values of α , β , and γ must all be between 0 and 1. The variable m in the equations represents the length or duration of the seasonal component in the series (Akgül, 1994, p.56; Hyndman & Athanasopoulos, 2018). As a result of the analyses,

$$a = 0,798$$

$$b = 0,001$$

$$\gamma = 0,330$$

These values were used as the smoothing coefficient in the model. The model's beginning value is

$$L_5 = 709997,9167$$

$$b_5 = 456,1217$$

The overall forecasting performance of the model,

$$\text{MAPE} = 22,74\%$$

$$\text{RMSE} = 280399$$

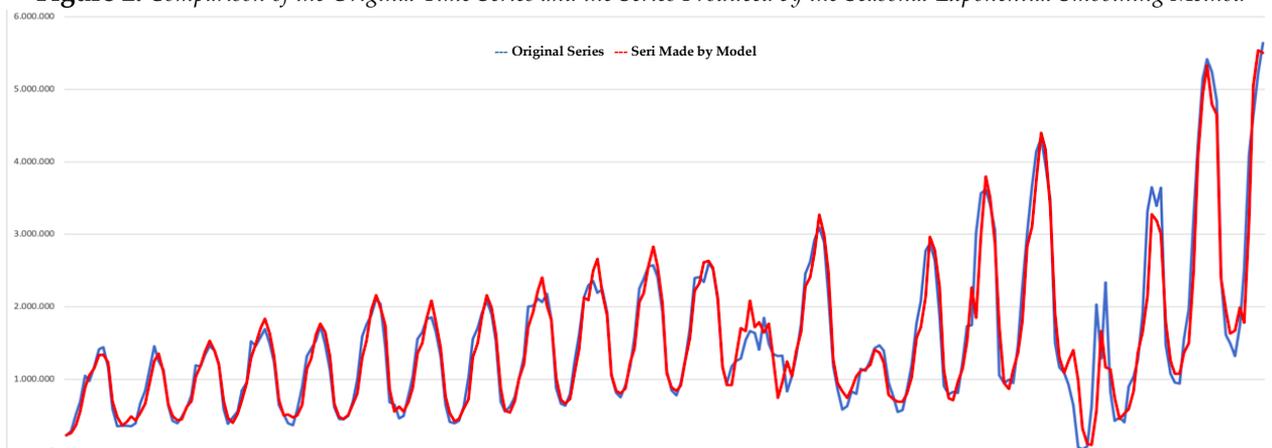
The global spread of the Covid-19 outbreaks in 2019 had major negative impacts on global tourism demand, resulting in major fluctuations in the data set. To minimize the impact of this condition, outlier data was included into the analyses in order to improve the forecasting abilities of the model. The model's overall prediction ability has been raised due to the new results.

$$\text{MAPE} = 21,01\%$$

$$\text{RMSE} = 240057$$

The following graphic illustrates the model's forecasting ability for all observed values.

Figure 2. Comparison of the Original Time Series and the Series Produced by the Seasonal Exponential Smoothing Method



Artificial Neural Networks

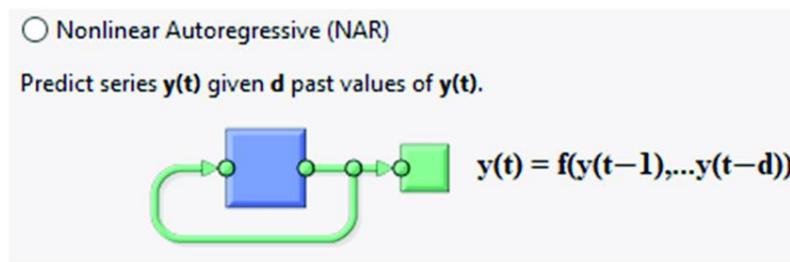
Artificial neural networks are AI systems that simulate the structure and function of the brain's nervous system. Instead of using biological neurons, they use artificial neurons to process and transmit information within the network. This is possible using a mathematical function (Lin, 2009, p.3510). Artificial neural networks minimize the need for determining the correlation between data, compared with traditional methods. To clarify why, it is enough to give the inputs and outputs of a network to the system (Söylemez & Türkmen, 2017, p.276). Artificial neural networks mainly consist of an input layer, weights that give significance to the data, a hidden layer, and an output layer responsible for transmitting the findings either outside the network or to the following neuron (Çalışkan & Acar, 2006, p.55). The network multiplies each data input by its related weight and collects them to calculate the net input of the network. The network's net input is calculated using the following formula.

$$\text{Net Input} = \sum_{i=1}^n I_i W_i \quad (13)$$

The data is transmitted through the network in two different ways. One of these is feedforward networks. Within these networks, neurons are arranged in a particular path from the input to the output, and data is only sent to the next layer. (Çayıroğlu, 2019, p.5). Feedback networks are used in cases when there is an error between the outputs resulting from the network and the actual data. In this feedforward architecture, data is sent from the output layer to the input layer again to minimize error. The data weights are reorganized to generate better results (Baylar, et al., 1999, p.5).

The method was applied using the NAR (Nonlinear Autoregressive Network) networks from the Artificial Neural Networks Time Series Module (ntstool) in the Matlab R2023b application. The program was designed to use the series' past values to create predictions. The NAR model was chosen in the research due to its ability to learn patterns and relationships by reusing the inputs and output values of the previous period to the network through feedback. During the analysis of the observed values in the time series, they were categorized into two distinct groups: 70% for training and 30% for testing, and 80% for training and 20% for testing. This categorization was automatically done by the computer. The networks that have been constructed have a single hidden layer and a single output layer. The delay vector was maintained within the range of 1:1 to 1:24, and the number of neurons in the hidden layer was consistently maintained between 1 and 15. A series of tests were conducted, spanning from 1 to 1000 epochs. The diagram illustrating a fundamental NAR network is depicted below.

Figure 3. Structure of the NAR Network



The series is normalized for easier analysis. As a result, the values were ensured to be inside the range of 0 and 1. The formula used for the normalizing step is shown below.

$$x_n = \frac{x_0 - x_{min}}{x_{max} - x_{min}} \quad (14)$$

x_n represents the normalized new data set. x_0 refers to the RS or SAS series. x_{min} represents the smallest value in the data set, while x_{max} represents the maximum value in the data set. Below is the Sigmoid activation function, which is necessary for the network to calculate the net input from the input layer and provide an output.

$$\psi(\text{NET}) = \frac{1}{1 + e^{-\text{NET}}} \quad (15)$$

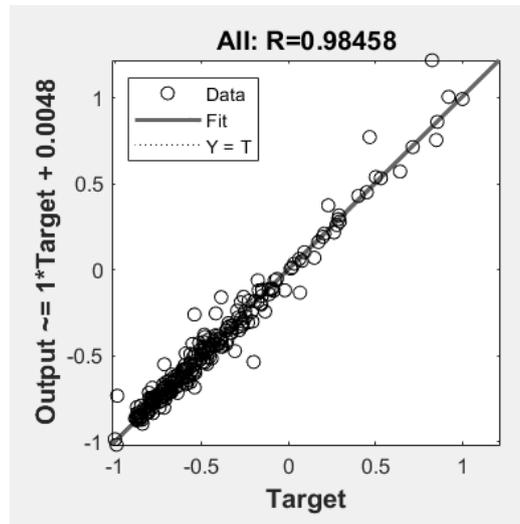
While creating the network, the input layer includes a set of values at different time delays, represented as $(y_{t-1}, y_{t-3}, y_{t-12} \dots \dots y_{t-k})$. The output layer, on the other hand, just includes the current value without any delay, represented as (y_t) . The network is structured as a feedforward architecture and uses error backpropagation. The model uses the logarithmic sigmoid (LogSig) as its transfer function, the Levenberg-Marquardt method (Trainlm) as its training function and mean square error (MSE) as a metric to minimize the network's errors. The results that were obtained were compared with the data from the previous 24 months. Below is a summary of the several networks that have been created.

Table 3. Summary Information about Artificial Neural Networks with Different Architectures

ANN No	Number of Neurons in the Hidden Layer	Number of Delays in the Hidden Layer	Output	MAPE	RMSE	Training (%) Testing (%) Validation (%)
1	10	15	y_t	%10,63	170263	80-10-10
2	12	16	y_t	%17,09	538135	80-10-10
3	13	16	y_t	%18,50	222227	70-15-15
4	11	18	y_t	%18,77	455212	75-15-15
5	10	11	y_t	%19,30	174440	80-10-10
6	4	15	y_t	%21,19	175317	70-15-15

After looking over the table above, it is apparent that the NAR network number 1, consisting of 10 neurons and 15 delays in the hidden layer, displayed the lowest MAPE value. The network assigned 80% of the data for training, 10% for testing, and 10% for validation. The visual representations generated from training the network are displayed below.

Figure 4. Training Results of RS_2



The intention for the R value in network training is to get close to a value of 1. A higher R value shows a greater level of success in network learning. Upon examining the obtained values, it can be concluded that the network has a high learning rate. Below is the table displaying the real values and the values generated by the network.

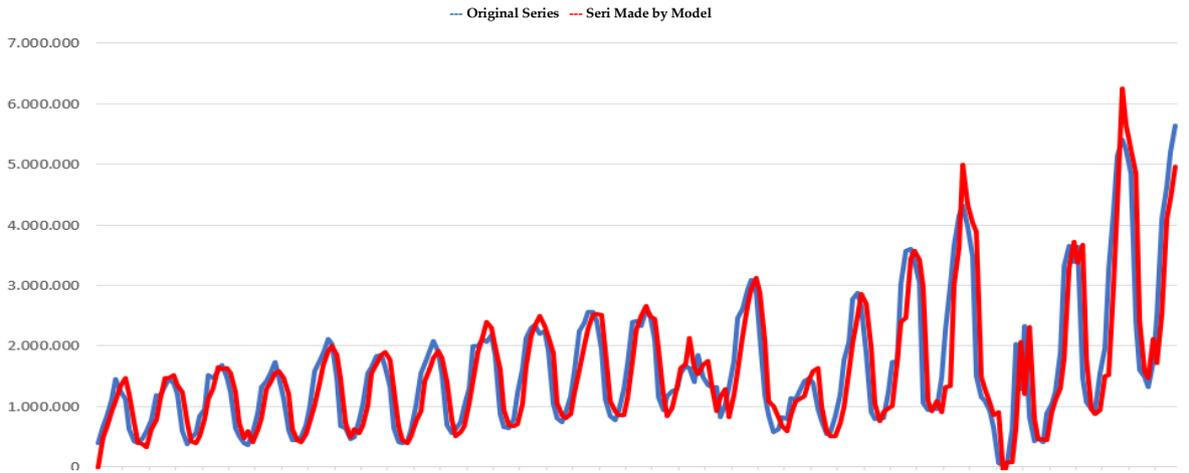
Table 5. Comparison of Network Outputs and Actual Values

Period	Output	Real Value	Period	Output	Real Value
1	3.377.560	3.389.257	13	5.251.373	5.236.797
2	3.663.880	3.641.065	14	4.839.646	4.839.610
3	1.780.370	1.476.375	15	2.409.877	2.401.759
4	962.372	1.074.142	16	1.590.684	1.607.254
5	869.231	954.982	17	1.477.516	1.492.732
6	933.113	942.065	18	2.109.122	1.317.681

7	1.497.038	1.558.729	19	1.717.521	1.740.209
8	1.512.341	1.969.719	20	2.526.170	2.516.309
9	3.034.411	3.251.037	21	4.102.525	4.097.542
10	4.352.123	4.243.484	22	4.440.547	4.636.321
11	6.255.924	5.151.089	23	4.955.020	5.219.266
12	5.659.442	5.417.540	24	5.623.579	5.642.337

The following graphic illustrates the model's forecasting ability for all observed values.

Figure 5. Comparison of the Original Time Series and the Series Produced by the NAR Model



Discussion and Conclusion

Since the beginning of civilization, travels have evolved into massive movements, hence elevating tourism to an important sector. The existence of foreign debts resulting from international transactions has necessitated the acquisition of foreign currencies. Exports have played an essential part in reducing these debts. Nevertheless, the ability to improve supply in the industrial and agricultural sectors is limited in the short and long term. Hence, nations are increasingly relying on the tourist industry to bring in foreign currency revenues by using their natural, cultural, and historical resources, which have little or no economic significance for other industries. Nevertheless, the susceptibility of the tourism industry to national and worldwide social, political, and economic crises, along with its interaction with many sub-sectors, makes the demand for tourism highly flexible. Furthermore, not being able to raise fixed capital investments in the short and medium term, besides the fact that accommodation establishments have fixed expenditures that are not affected by the occupancy rate, underlines the critical importance of making a suitable strategy for tourism investments. To establish a successful investment plan, it is important to evaluate the demand structure and provide a demand forecast with minimum error.

Identifying the negative effects of the Covid-19 pandemic, which started in 2019 and quickly reached as a worldwide scale and creating appropriate solutions have become as an essential priority for both the corporate and public sectors. The implementation of unexpected actions like as closing of borders, cancellations of flights, and strict enforcement of curfews as a result of the Covid-19 pandemic has negatively impacted the tourism industry. Recently, there has been an enormous rise in research focused on researching the demand for tourism. Scientists are trying to obtain results that are most accurate to real-life by creating various models. Nevertheless, it is difficult to address just one method that generates results with the minimum error in demand forecasting. The method used is dependent on the particular characteristics and size of the dataset. To improve the reliability of forecasts, it is essential to use multiple techniques and choose the model with the lowest range of error.

This study aimed to develop forecasting models that reliably forecast Türkiye's international tourism demand based on data collected from January 2002 to August 2023. The models used were compared according to their Mean Absolute Percentage Error (MAPE) and Root Mean Square Error (RMSE) values. Based on the data, it has been discovered that artificial neural networks, which are forms of artificial intelligence, generate better results. Despite the important abnormalities in the data set and the influence of the Covid-19 pandemic, it can be said that artificial

neural networks have better adaptation to unexpected changes in demand. Artificial intelligence models, being increasingly used across different fields due to technological developments, succeed in identifying relationships between variables, creating connections, and generating results. Moreover, they have the ability to produce more accurate forecasts in situations including unpredictability.

Limitations and Future Research

The present study analyzes monthly data dating from January 2002 to August 2023. Hence, the study is limited to the data that was collected within this specific timeframe. For future research, a more comprehensive approach to analyzing tourism demand may include using by pre-2002 statistics or data collected after major crises that had an impact on demand, like as the Covid-19 pandemic. Hybrid models, which combine traditional models and artificial intelligence models, have the potential to produce better results that fit closely with actual data in the future. Moreover, these models may be used to do forecast demand for a particular destination or a certain ethnic group.

Ethics Statement

The writing process of the study named "Forecasting Türkiye's International Tourism Demand" followed to scientific, ethical, and citation criteria. The acquired data was not manipulated, and the paper was not submitted to any other academic journal for review.

References

- Akgül, I. (1994). Time Series Analysis and Forecasting Models. Recommendation: Marmara University Social Sciences Institute Journal, 1(1), 52–69.
- Atlas, M. (2013). "Time Series Analysis", in Statistics II, Şıklar, E.; Özdemir, A. (Ed.), Eskişehir, Anadolu University Publication No: 2806. E-ISBN 978-975-06-3457-4
- Barry, K., & O'Hagan, J. (1972). An Econometric Study of British Tourist Expenditure in Ireland. *Economic and Social Review*, 3(2), 143–161.
- Baylar, A., Emiroğlu, E., & Arslan, A. (1999). Finding the Ratio of Swab That Will Direct to the Lateral Water Intake Structure Using Backpropagation Artificial Neural Network. *Dokuz Eylül University Journal of Engineering Sciences*, 1(2), 1–12.
- Bayram, O. (2018). Estimation of the Tourism Demand Equation for Turkey: Economic, Social and Political Factors [Master's Thesis]. Izmir Katip Çelebi University.
- Can, M. (2009). Forecasting in Business Using Time Series Analysis, Istanbul University, Institute of Social Sciences, Department of Business Administration, Doctoral Thesis.
- Cruz-Milan, O. (2018). Plog's Model of Personality-Based Psychographic Traits in Tourism: A Review of Empirical Research. In *Tourism Planning and Destination Marketing* (pp. 49–74). ISBN: 978-1-78756-292-9, eISBN: 978-1-78756-291-2
- Çağlar, T. (2007). Methods Used in Demand Forecasting and Its Application in a Fencing Wire Production Enterprise, Kırıkkale University, Institute of Science and Technology, Department of Industrial Engineering, Master's Thesis.
- Çalışkan, E., & Acar, H. H. (2006). An Evaluation on the Use of Artificial Intelligence Techniques in Wood Raw Material Production. *Artvin Faculty of Forestry Journal*, 7(1), 51–59.
- Çayıroğlu, I. (2019). Advanced Algorithm Analysis: 5 Artificial Neural Networks. Karabük University Faculty of Engineering. <http://www.ibrahimcayiroglu.com/Dokumanlar/IleriAlgoritmaAnalizi/IleriAlgoritmaAnalizi-5.Hafta-YapaySinirAglari.pdf>
- Dinler, Z. (2006). Introduction to Economics (12th ed.). Ekin Bookstore Publications.
- Edgell, D. L. (1993). World Tourism at the Millennium: An Agenda for Industry, Government, and Education. BASE. Department of Commerce, U.S. Travel and Tourism Administration.

- Göktas, L.S. (2023). The effects of Kahramanmaraş-Centered Earthquakes on the Tourism Sector and Their Response After the Earthquake Recommendations on the Steps Needed. *Journal of Gastronomy, Hospitality and Travel*, 6(2), 624-635.
- Hanke, J. E., Reitsch, A. G. (1992), *Business Forecasting* (4th Ed.), Boston, Allyn and
- Hyndman, R. J., Athanasopoulos, G. (2018) *Forecasting: Principles and Practice* (2nd Edition), (Online) Monash University, Australia. <https://otexts.com/fpp2/>
- Köroğlu, A., & Güleç, B. (2008). Youth Tourism. In N. Hacıoğlu & C. Avciokurt (Eds.), *Touristic Product Diversification*. Nobel Publishing Distribution.
- Krajewski, L. J., Ritzman, L. P., Malhotra, M. K. (2010). *Operations Management: Processes And Supply Chains*. Upper Saddle River, New Jersey: Pearson. ISBN-13: 9780136860631
- Lewis, C. D. (1982). *Industrial and Business Forecasting Methods*, London, Butterworths Publishing. <https://doi.org/10.1002/for.3980020210>
- Lim, C. (1997). Review of International Tourism Demand Models. *Annals of Tourism Research*, 24(4), 835–849. [https://doi.org/10.1016/s0160-7383\(97\)00049-2](https://doi.org/10.1016/s0160-7383(97)00049-2)
- Lin, T.-H. (2009). A Cross Model Study of Corporate Financial Distress Prediction in Taiwan: Multiple Discriminant Analysis, Logit, Probit and Neural Networks Models. *Neurocomputing*, 72, 3507–3516. <https://doi.org/10.1016/j.neucom.2009.02.018>
- Lundberg, D. E., Krishnamoorthy, M., & Stavenga, M. H. (1995). *Tourism Economics*. Wiley. ISBN: 978-0-471-57884-0
- Öztaş, K., & Karabulut, T. (2006). *Tourism Economy General Tourism Information*. Nobel Publishing Distribution.
- Pizam, A., & Milman, A. (1986). The Social Impacts of Tourism. *Tourism Recreation Research*, 11(1), 29–33. <https://doi.org/10.1080/02508281.1986.11014414>
- Soysal, M., & Ömürkünüşen, M. (2010). An Application on Demand Forecasting in the Turkish Tourism Sector. *Anatolia: Journal of Tourism Research*, 21(1), 128–136. <http://dergipark.gov.tr/atad/issue/16799/174505>
- Söylemez, Y., & Türkmen, S. Y. (2017). Financial Failure Prediction with Artificial Neural Networks Model. *Journal of Finance, Economics and Social Research*, 2(4), 270--284. <https://doi.org/10.29106/fesa.364323>
- Şenel, S. A. (2007). Investment Decisions in the Tourism Sector. *Selçuk University Karaman İ.İ.B.F. Journal*, 12(9), 1–12.
- Taptik, Y., & Keleş, Ö. (1998). *Quality War*. Kalender Publications.
- TDK. (2020). "Demand", (Online), Access date 13.05.2020
- TÜRSAB (2021a). "Share of Imports in Exports", (Online), Access date 17.08.2022 <https://www.tursab.org.tr/turkiye-turizmistatistikleri/turizm-geliri>
- TÜRSAB (2021b). "Share in Closing Foreign Trade Deficits", (Online), <https://www.tursab.org.tr/istatistikler-icerik/dis-ticaretaciklarinin-kapatmada-payi>
- UNWTO. (2018). *International Tourism Highlights*, 2018 Edition. <https://doi.org/10.18111/9789284419876>
- UNWTO. (2021). *International Tourism Highlights 2020 Edition*. <https://doi.org/10.18111/9789284422456>
- UNWTO. (2023). *International Tourism Highlights 2023 Edition The Impact of COVID-19 on Tourism (2020-2022)*. <https://doi.org/10.18111/9789284424986>.
- Ünlüönen, K., Tayfun, A., & Kılıçlar, A. (2007). *Tourism Economics* (1st ed.). Nobel Publishing Distribution.
- Witt, S. F., & Witt, C. A. (1995). Forecasting Tourism Demand: a Review of Empirical Research. *International Journal of Forecasting*, 11, 447–475. [https://doi.org/10.1016/0169-2070\(95\)00591-7](https://doi.org/10.1016/0169-2070(95)00591-7)