Erciyes Üniversitesi
İktisadi ve İdari Bilimler Fakültesi Dergisi

Araştırma Makalesi | Research Article
Yayın Tarihi: 30 Aralık 2024 | Published: December 30, 2024
doi: 10.18070/erciyesiibd.1581119

ISSN: 1301-3688/e-ISSN: 2630-6409

MODELING AND FORECASTING OF TOURISM INCOME: THE CASE OF TURKIYE

BEGÜM KARA GÜLAY[1*]

[1] *Arş. Gör. Dr., Dokuz Eylül Üniversitesi Fen Fakültesi İstatistik Bölümü, begum.kara@deu.edu.tr, https://orcid.org/0000-0003-2926-2699.*

ÖZ

This study aims to identify the optimal forecasting model for predicting Türkiye's tourism income, a crucial factor for economic planning and development. This study employs different forecasting techniques, including the seasonal Autoregressive Integrated Moving Average (SARIMA), the additive and multiplicative Holt-Winters methods, the Exponential Smoothing State Space (ETS), Artificial Neural Networks (ANNs) and seasonal-trend decomposition procedure based on the loess (STL)-ANN hybrid model and evaluates their performance. The methodology involves analyzing monthly tourism income data from January 2012 to December 2023, incorporating additional economic indicators such as the economic confidence index, number of visitors, consumer price index, industrial production index, and USD exchange rate, which serve as input for ANN models. The findings reveal that ANNs, particularly the model that incorporates tourism income alongside other economic indicators, outperform traditional models with the lowest Mean Absolute Scaled Error (MASE) and Root Mean Squared Scaled Error (RMSSE). Specifically, the ANN model with additional predictors demonstrates the highest forecasting accuracy. These results suggest that advanced machine learning techniques provide superior predictive capabilities compared to traditional linear models. The study underscores the importance of integrating complex models to achieve more accurate forecasting, offering valuable insights for policymakers and practitioners in the tourism sector.

Keywords: Tourism Income, Forecasting, Artificial Neural Networks, Time Series Analysis

Editör / Editor:
Gökhan ÇOBANOĞULLARI, Erciyes Üniversitesi, Türkiye

****Sorumlu Yazar/ Corresponding Author:***
Begüm KARA GÜLAY, begum.kara@deu.edu.tr

JEL:
C18, C45, C52

Geliş: *7 Kasım 2024*
Received: *November 7, 2024*
Kabul: *6 Aralık 2024*
Accepted: *December 6, 2024*
Yayın: *30 Aralık 2024*
Published: *December 30, 2024*

Atıf / Cited as (APA):
Kara Gülay, B. (2024), Modeling and Forecasting of Tourism Income: The Case of Turkey, Erciyes Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi, 69, 251-257, doi: 10.18070/erciyesiibd.1581119

TURİZM GELİRLERİNİN MODELLENMESİ VE TAHMİN EDİLMESİ: TÜRKİYE ÖRNEĞİ

ABSTRACT

Bu çalışmanın amacı, ekonomik planlama ve kalkınma için önemli bir faktör olan Türkiye'nin turizm gelirini tahmin etmek için optimum tahmin modelini belirlemektir. Mevsimsel Otoregresif Entegre Hareketli Ortalama (SARIMA), Holt-Winters yöntemleri (eklemeli ve çarpımsal), Durum Uzay Modelleri (ETS), Yapay Sinir Ağları (YSA) ve mevsimsel eğilim ayrıştırma (STL) -ANN hibrit modeli dahil olmak üzere tahmin tekniklerini kullanmak ve bunların performansını değerlendirmektedir. Metodoloji, Ocak 2012'den Aralık 2023'e kadar aylık turizm geliri verilerinin analiz edilmesini ve YSA modeli için ziyaretçi sayıları, ekonomik güven endeksi, tüketici fiyat endeksi, endüstriyel üretim endeksi ve ABD doları döviz kuru gibi ek ekonomik göstergelerin dahil edilmesini içermektedir. Bulgular, özellikle turizm gelirini diğer ekonomik göstergelerle birlikte içeren model olan YSA'ların, en düşük Ortalama Mutlak Ölçekli Hata ve Kök Ortalama Karesel Ölçekli Hata ile geleneksel modellerden daha iyi performans gösterdiğini ortaya koymaktadır. Özellikle, ek öngörücülere sahip YSA modeli en yüksek tahmin doğruluğunu göstermiştir. Bu sonuçlar, gelişmiş makine öğrenme tekniklerinin geleneksel doğrusal modellere kıyasla üstün tahmin yetenekleri sağladığını göstermektedir. Çalışma, daha doğru tahminler için karmaşık modellerin entegre edilmesinin önemini vurgulayarak, turizm sektöründeki politika yapıcılar ve uygulayıcılar için değerli sonuçlar sunmaktadır.

Anahtar Kelimeler: Turizm Geliri, Öngörümleme, Yapay Sinir Ağları, Zaman Serisi Analizi

INTRODUCTION

The tourism industry is a significant contributor to economic growth worldwide and serves as a primary revenue source for many nations. Accurate estimation of tourism income is crucial for companies in the sector, policymakers, and economists, given the key role of tourism in economic development. Tourism income also contributes significantly to local economies and regional development, enhancing infrastructure development and creating job opportunities in tourist- visited regions (Dritsakis, 2012).

In 2024, the World Travel & Tourism Council (WTTC) projects the global economic contribution of travel and tourism to reach an unprecedented $11.1 trillion, accounting for approximately 10% of global GDP. The sector supports nearly 348 million jobs worldwide, an increase of over 13 million since its previous peak in 2019. International visitor spending is expected to approach $1.9 trillion, highlighting its importance for economies reliant on inbound tourism (WTTC, 2024). These statistics underline the sector's resilience and its role in driving infrastructure development, creating employment, and enhancing the quality of life in many regions globally.

Accurate estimation of tourism income is vital for policymakers, industry players, and other stakeholders as it facilitates strategic economic planning, efficient resource allocation, and effective risk management. Reliable income estimates are crucial for developing sustainable tourism strategies, supporting regional development, and ensuring the long-term competitiveness of the tourism sector. Given the role of tourism in economic planning and national growth, understanding the methodologies used to predict tourism income and factors affecting revenue is essential for future planning. The present study seeks to determine the optimal forecasting model for estimating monthly tourism revenues in Türkiye, contributing to more effective tourism management and development. Specifically, the research question addressed in this study is: Which forecasting model provides the highest accuracy for predicting Türkiye's monthly tourism revenues?

I. TOURISM SECTOR IN TURKIYE

Türkiye is globally recognized as an important destination for tourism income. Türkiye is a prominent global tourism destination, known for its rich cultural heritage, natural beauty, and diverse attractions that appeal to a wide range of tourists. The country has gained recognition for its unique blend of historical sites, natural wonders, and affordable travel experiences. From the UNESCO World Heritage Sites such as Ephesus, Cappadocia, and the historic areas of Istanbul, to the scenic Mediterranean beaches of Antalya and Bodrum, Türkiye offers a broad range of attractions that drive its tourism sector.

In 2023, the tourism sector accounted for 12% of Türkiye's GDP, equating to approximately TRY 3.11 trillion, a substantial increase from 11% in 2019 (WTTC, 2023). This growth not only reflects the sector's resilience post-pandemic but also highlights its critical role in supporting the socio-economic development of regions dependent on tourism. For example, international visitor spending in Türkiye increased by 38.2% compared to pre-pandemic levels, signaling a strong rebound and significant contribution to local economies (WTTC, 2024).

The tourism industry also supports millions of jobs, directly and indirectly employing approximately 3.23 million people as of 2023 (WTTC, 2023). This substantial workforce underlines the sector's importance in providing livelihoods and fostering regional development. Moreover, the government's Tourism Strategy of Türkiye - 2023 reflects a commitment to sustaining and growing the sector through initiatives such as diversifying tourism offerings, improving infrastructure, and expanding niche markets, including eco-tourism, cultural tourism, and health tourism (Republic of Türkiye Ministry of Culture and Tourism, 2023).

Given the sector's pivotal role in Türkiye's economy, understanding the factors influencing tourism revenue and refining forecasting models are essential for ensuring sustainable development and long-term growth. Türkiye's tourism industry exemplifies the transformative potential of tourism in driving economic growth, improving infrastructure, and enhancing quality of life, making it a cornerstone of the country's economic framework.

II. EMPIRICAL LITERATURE REVIEW

Tourism income not only boosts the national economy but also facilitates regional development by fostering infrastructure projects and creating job opportunities (Dritsakis, 2012). However, the sustainable growth of the sector depends on accurate forecasting and effective management of tourism income. Estimation methodologies are critical for strategic planning, investment decisions, competitiveness assessments, efficient resource allocation, and risk management (Sonmez & Sirakaya, 2002; Song et al., 2003). Accurate predictions also provide a strategic advantage, enabling policymakers and stakeholders to navigate challenges and seize opportunities in the dynamic global tourism landscape.

The estimation and forecasting of tourism revenues in Türkiye have been investigated in a few empirical studies, employing a variety of methods to improve accuracy and robustness. Akal (2004) used the ARMAX model to estimate Türkiye's tourism revenues and forecast the future of international tourism revenues in the post-2001 economic crisis period. Cuhadar (2020) examined different model performances using artificial neural networks (ANNs), focusing on modeling and forecasting Türkiye's tourism revenues. Aydın (2016) modeled these tourism revenues using the panel data method and examined periodic effects. Kayakuş et al. (2023) estimated the share of Türkiye's tourism revenues in total export revenues using multiple linear regression and ANNs. Koyuncu et al. (2016) analyzed health tourism revenues for 2002-2015, providing future predictions. Findings indicate a rising trend in Türkiye's health and overall tourism revenues. Tuncsiper (2023) modeled and estimated Türkiye's tourism revenues using a deep learning network, based on factors such as tourist numbers, oil prices, and exchange rates. Kayral et al. (2023) employed several statistical methods to estimate Türkiye's tourist arrival volumes and tourism revenues, aiming to determine the best forecast model, evaluating the impact of various determinants, and assessing the Russia-Ukraine war's effects on arrivals and revenues. Yenişehirlioğlu et al. (2020) examined the relationship between tourism revenues and economic growth from 1995-2017 applying symmetric, asymmetric, and moving window regressions.

In addition to these studies focusing on Türkiye, the global literature provides valuable insights. Song and Li (2008) applied advanced econometric models in international tourism demand forecasting, emphasizing the importance of structural breaks and demand elasticity. Using cointegration techniques, Dritsakis (2012) demonstrated the long-run relationship between tourism development and economic growth. Furthermore, a study by Wong et al. (2020) demonstrated the effectiveness of Bayesian methods in tourism demand forecasting and proved that these techniques yield results with high accuracy rates. The inclusion of external determinants, such as exchange rates, oil prices, and geopolitical risks, has also been a common thread in literature. For example, Goh and Law (2012) evaluated the effects of external shocks on tourism demand using stochastic and intervention analysis methods and they showed that these factors are decisive, especially during crisis periods.

This study builds on the existing literature by addressing the gap in monthly tourism revenue forecasting for Türkiye, incorporating a range of predictors, and evaluating the performance of different models in terms of accuracy and reliability. This study aims to determine the optimal forecasting model by comparing the forecast accuracy of Autoregressive Integrated Moving Average, Holt-Winters, Error-Trend-Seasonality, ANNs and STL-ANN for Türkiye's monthly tourism revenues. It aims to provide a basis for tourism development plans, facilitate decision-making in the monthly plans of operators, and contribute to tourism literature more broadly.

III. MATERIAL AND METHODS

In this section, we outline the data sources, variables, and methodologies used to forecast Türkiye's tourism income. We start by describing the dataset, and then present the various forecasting methods applied in this study.

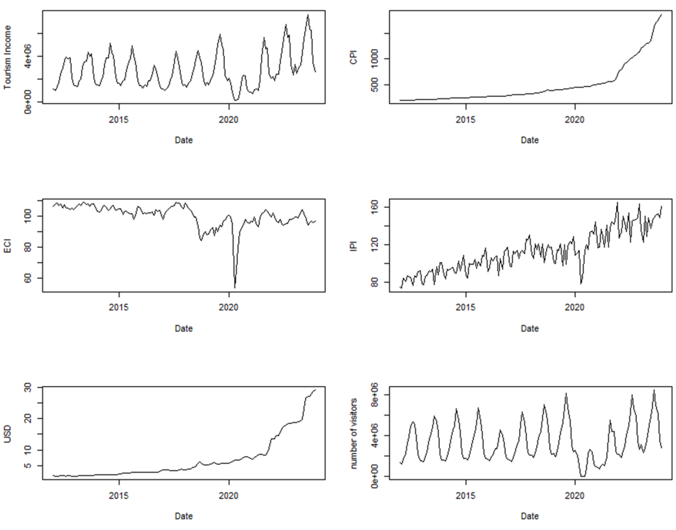
I. DATASET

This study aims to forecast Türkiye's tourism income using univariate time series and machine learning models. For this purpose, in addition to monthly tourism income (thousand $) data, the number of visitors, economic confidence index (ECI), consumer price index (CPI), industrial production index (IPI) and the dollar exchange rate (USD) are used as inputs in the ANN model. Data from January 2012 to December 2023 are used. While the Central Bank of the Republic of Türkiye website provides the dollar exchange rate data, other variables are obtained from the Turkish Statistical Institute website. The descriptive statistics of all variables are summarized in Table 1.

TABLE 1 | Descriptive Statistics

	Tourism Income (Thousand $)	CPI	ECI	IPI	USD	# of Visitors
Minimum	175638	202	53.34	73.39	1.751	727126
Maximum	7616302	1859.4	108.72	165.56	29.023	8488026
Mean	2746789	482.8	99.97	112.38	6.819	3403031
Median	2366457	329.1	101.7	110.96	3.813	2800281
Std. Dev.	1506302	373.691	7.621	22.091	6.641	1813137
Skewness	0.8	1.98	-2.54	0.39	1.76	0.7
Kurtosis	0.1	3.25	10.98	-0.7	2.29	-0.41
# of Missing Value	0	0	0	0	0	3

FIGURE 1 | Time Series Plots of the Variables



According to Figure 1, tourism income and the number of visitors show clear seasonal peaks, indicating higher activity during holiday periods. Both variables have grown over the years, reflecting an expanding tourism sector. CPI exhibits a strong upward trend, with accelerated inflation in recent years. ECI remains stable but drops significantly around 2020 due to COVID-19, with partial recovery afterward. IPI steadily increases, indicating growth in industrial production despite short-term fluctuations. The USD exchange rate sharply increased, especially since 2018, showing exponential growth.

II. METHODS

In this study, we utilize various forecasting techniques that are classified into linear and non-linear models. For predicting tourism revenues, we employ linear models such as autoregressive integrated moving average (ARIMA), seasonal additive Holt-Winters (HW-A), seasonal multiplicative Holt-Winters (HW-M), error-trend-seasonality (ETS) and seasonal- trend decomposition procedure based on Loess-artificial neural network (STL-ANN). These methods are created using only tourism income and its lags. Conversely, we utilize the artificial neural network (ANN) model in two configurations: one using only tourism income and its lags, and the other uses using tourism income as the dependent variable with the number of visitors, ECI, CPI, IPI, and USD as independent variables.

1. Seasonal Autoregressive Integrated Moving Average (SARIMA)

The Autoregressive Moving Average (ARMA) model is commonly employed in time series forecasting. It consists of two components: the Autoregressive (AR) model, where the output variable is a linear function of its past values, and the Moving Average (MA) model, where the output variable is a linear function of the current and past values of a stochastic term. ARIMA models, or Integrated Autoregressive Moving Average processes, extend ARMA models to handle non-stationary data. When the data are non-stationary, differencing (d) is applied to achieve stationarity, transforming the model into an ARIMA model. The general form of an ARIMA model is denoted as ARIMA (p, d, q), where p represents the order of the autoregressive part, d denotes the degree of differencing, and q indicates the order of the moving average part. The ARIMA (p, d, q) model is expressed in Equation (1):

$$(1-\phi_1 B-\phi_2 B^2-\cdots-\phi_p B^p)(1-B)^d Y_t = c+(1-\theta_1 B-\theta_2 B^2-\cdots-\theta_q B^q)e_t \qquad (1)$$

where Y_t represents the values for interest at time t, e_t denotes a series of random errors, and B is the backward shift operator.

The SARIMA (p, d, q) (P, D, Q) s model is represented in Equation (2):

$$(1-\phi_1 B-\cdots-\phi_p B^p)\phi_p B^S(1-B)^d(1-B^S)^D Y_t = (1-\theta_1 B-\cdots-\theta_q B^q)\theta_Q(B^S)e_t \qquad (2)$$

where p, d, and q represent the non-seasonal orders, while P, D, and Q denote the seasonal orders. S is the time span of the repeating seasonal pattern (Hossen et al., 2022).

SARIMA is particularly suited for capturing seasonal patterns and trends within time series data, making it highly relevant for tourism revenue forecasting, where seasonality plays a critical role. SARIMA can model both non-seasonal and seasonal components simultaneously while accounting for potential autocorrelations and stationarity issues. In this study, we utilize the SARIMA model, which incorporates additional seasonal components into the ARIMA model, as our variable is monthly tourism income data.

2. Holt-winters (HW) Method

Holt-Winters models are designed to capture structural features such as trends and seasonal patterns in time series, such as tourism revenues. When a series shows a trend, seasonality, or both, the HW method addresses these features using three smoothing equations: Level (*Lt*), Trend (*Bt*) and Seasonality (*St*). The HW method can be expressed in two forms as additive and multiplicative based on how seasonality is integrated into the model. In this study, we apply both versions of the HW model to forecast tourism revenues.

Multiplicative Form

$$Lt = \alpha\left(\frac{Y_t}{S_{t-m}}\right)+(1-\alpha)(L_{t-1}+B_{t-1}) \qquad (3)$$

$$Bt = \beta(L_t-L_{t-1})+(1-\beta)B_{t-1} \qquad (4)$$

$$St = \gamma(Y_t / L_t)+(1-\gamma)S_{t-M} \qquad (5)$$

$$\hat{Y}_{t+1} = (L_t+k.B_t)S_{t-M-k} \qquad (6)$$

Additive Form

$$L_t = \alpha(Y_t-S_{t-M})+(1-\alpha)(L_{t-1}+B_{t-1}) \qquad (7)$$

$$B_t = \beta(L_t-L_{t-1})+(1-\beta)B_{t-1} \qquad (8)$$

$$S_t = \gamma(Y_t-L_t)+(1-\gamma)S_{t-M} \qquad (9)$$

$$\hat{Y}_{t+1} = (L_t+k.B_t)S_{t-M-k} \qquad (10)$$

where M represents the seasonal period (*e.g.,* M =12 for monthly data) and k denotes the forecast horizon (Chatfield et al., 2001).

3. Error-trend-seasonality (ETS)

Exponential smoothing state-space methods encompass a range of techniques for univariate time series forecasting. Each model comprises a measurement equation that describes the observed data and state equations that explain how unobserved components (level, trend, seasonal) evolve over time. For this reason, they are called state-space models. There are different types of state-space models to address various specific situations, and these are generally called level, trend and seasonality (Hyndman et al., 2008).

The models contain different components to adapt to the characteristics and structures of different time series data. Each state-space model, consisting of Error, Trend, and Seasonal components, is denoted as ETS (*E, T, S*). In this notation, *E* represents the error type, which can be additive (*A*) or multiplicative (*M*); *T* represents the trend type, which can be none (*N*), additive (*A*), additive damped (*Ad*), multiplicative (*M*), or multiplicative damped (*Md*); and S represents the seasonal type, which can be none (*N*), additive (*A*), or multiplicative (*M*). For example, ETS (*M, Ad, N*) indicates a model with multiplicative error, additive damped trend, and no seasonality. ETS models include state equations that describe the evolution of these unobserved components over time. The 30 possible model combinations within the ETS framework are listed in Table 2.

TABLE 2 | ETS Models

ETS (M, M, N)	ETS (A, M, A)	ETS (M, M, M)
ETS (M, A, N)	ETS (A, Md, N)	ETS (M, A, A)
ETS (A, M, N)	ETS (A, Md, M)	ETS (A, A, A)
ETS (A, A, N)	ETS (A, N, A)	ETS (M, Ad, A)
ETS (M, A, M)	ETS (M, Ad, M)	ETS (A, Ad, A)
ETS (M, N, N)	ETS (M, Md, N)	ETS (M, Md, A)
ETS (A, M, M)	ETS (A, Ad, N)	ETS (A, Md, N)
ETS (A, A, M)	ETS (M, Md, M)	ETS (A, Md, A)

Source: *Hyndman et al., 2008.*

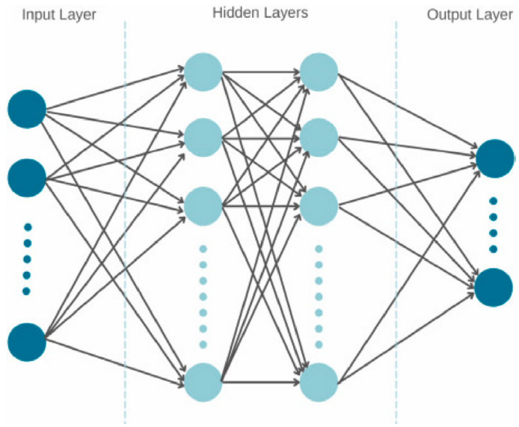
ETS explicitly decomposes time series into error, trend, and seasonal components. This decomposition provides a flexible and transparent approach to handling data with complex seasonal or trend structures, such as tourism revenue. Moreover, ETS can adaptively select the best combination of additive or multiplicative error, trend, and seasonality components, ensuring that the model is tailored to the underlying data dynamics (Hyndman et al., 2008). This flexibility makes it particularly effective for datasets with non-linear patterns or varying seasonal effects, which are common in tourism data. By including ETS in this study, we aim to leverage its strengths in capturing intricate patterns and providing accurate, interpretable forecasts.

4. Artificial Neural Network (ANN)

Artificial Neural Networks (ANNs) consist of multiple interconnected processing nodes, called "neurons," arranged in layers. Input data is fed through the network layer by layer, with each neuron connection modifying the data, until it reaches the output layer where a prediction or classification is generated.

The initial stage of developing an ANN involves making key decisions, such as determining the number of layers and neurons per layer, defining the input and output variables, and selecting training, validation, and test datasets. During the training phase, the neural network learns to identify the relationships between the input and output variables. This iterative process continues until the network minimizes error. In the validation phase, adjustments are made to the connection weights using validation data to ensure the model's robustness and generalization. Finally, the network's performance is evaluated using separate test data to evaluate its effectiveness in real-world applications.

FIGURE 2 | ANN with Two Hidden Layers



Source: *Kontogianni et al., 2022*

The theoretical findings indicate that using two hidden layers in the ANN model is enough (Kecman, 2001; Kůrková, 1992; Thomas et al. ,2017). Figure 2 illustrates the architecture of an ANN which includes two hidden layers. Using two hidden layers not only increases the efficiency of activation functions but also provides better generalization with fewer parameters. This approach requires fewer computational resources and can reduce the risk of overfitting.

ANN's ability to model both linear and nonlinear components without requiring explicit assumptions provides a distinct advantage when dealing with high-dimensional and complex datasets. By including ANNs in this study, we aim to leverage their strengths in handling complex data structures and evaluate their performance against more traditional models.

5. Seasonal-trend Decomposition Procedure Based on Loess (STL) -ANN Hybrid Model

STL applies iterative Loess smoothing to decompose a time series into its seasonal, trend, and residual components. The decomposition starts by approximating the trend line of the time series, typically using a polynomial model. The series is then de-trended by subtracting the original data from the approximated trend line. The seasonal component is subsequently identified by examining the entire dataset. De-seasonalization is performed by subtracting the seasonal data from the original time series (Apaydın et al., 2021). Finally, the residuals, representing the remaining variations after removing both the seasonal and trend components, are computed. These extracted components are modeled using an ANN, and the final forecasts are generated by summing the individual forecasts of the components.

6. Assessing Performance

Evaluating the performance of a model is a crucial step in determining its reliability and effectiveness in making accurate predictions or classifications. Without proper evaluation, it is impossible to determine whether a model effectively captures underlying patterns in the data or generalizes well to unseen situations. It also facilitates fair comparisons between different models or algorithms, allowing researchers and practitioners to choose the most appropriate approach for a given problem. Using appropriate metrics, the strengths and weaknesses of a model can be quantified, ensuring informed decision-making. To this end, this study employed Mean Absolute Scaled Error (MASE) and Root Mean Squared Scaled Error (RMSSE). The definitions of MASE and RMSSE are provided in the Equations (11) and (12), respectively:

$$MASE = \frac{\frac{1}{h}\sum_{t=n+1}^{n+h}\left|y_t - \hat{y}_t\right|}{\frac{1}{n-1}\sum_{t=2}^{n}\left|y_t - y_{t-1}\right|} \tag{11}$$

$$RMSSE = \sqrt{\frac{\frac{1}{h}\sum_{t=n+1}^{n+h}(y_t - \hat{y}_t)^2}{\frac{1}{n-1}\sum_{t=2}^{n}(y_t - \hat{y}_{t-1})^2}} \tag{12}$$

where y_t represents the actual value at time t, $\hat{y}_t$ denotes the forecasted value at time t, n is the number of observations in the training sample and h is the forecasting horizon (Muhaimin et al., 2021).

III. RESULTS

In this study, various forecasting models were employed to estimate Türkiye's tourism income. These models include linear and non-linear approaches, specifically the SARIMA, HW-A, HW-M, and ETS models. Additionally, ANNs were utilized to capture complex relationships within the data. All analyses were performed with the R 4.1.2 programming language (R Foundation for Statistical Computing, Vienna, Austria).

The SARIMA, HW-A, HW-M, ETS and STL-ANN models were constructed using only tourism income data and its lags. These models are well-suited for capturing linear relationships and patterns within the tourism income series, leveraging past values to forecast future outcomes.

In contrast, we employed the ANN model in two distinct configurations to account for both linear and nonlinear relationships. The first configuration utilized only tourism income data and its lags. The second configuration, shown in Equation (13), incorporated tourism income as the dependent variable, with the number of visitors, ECI, CPI, IPI, and USD as independent variables.

$$Y_t = f(V_t, ECI_t, CPI_t, IPI_t, USD_t) \quad (13)$$

where Y_t represents the tourism income, V_t denotes the number of visitors, ECI_t is the Economic Confidence Index, CPI_t is the Consumer Price Index, IPI_t is the Industrial Production Index, USD_t is the exchange rate at time t. This approach allowed the ANN to consider a broader range of economic indicators, potentially enhancing its predictive accuracy by capturing the multifaceted nature of the factors influencing tourism income.

In forecasting, the univariate time series data was split into two parts: a training set and a test set. The training set includes 132 observations, and the test set comprises the last 12 observations of the data. For the ANN, the data were divided into three distinct partitions: the training set, validation set and test set. The training set contains 120 observations, and both the validation and test sets consist of 12 observations each.

To compare the forecasting performance of the models, we utilize two accuracy measures: MASE and RMSSE. The results of these model comparisons are presented in Table 3. The different values of p, d, q, P, D, Q and S were tested to identify the most effective model for optimal forecasting. The Akaike information criterion (AIC) was used to determine the degree of differencing and the appropriate seasonal orders for selecting the best SARIMA model. Based on the results, SARIMA (2,0,0) $(0,1,1)_{12}$ was identified as the best model. ANN models were built using the hyperparameters outlined in Table 4. The ANN model with a lag of 12 indicates that the model uses data from the past 12 periods (months) to make predictions and is shown as ANN_1 in Table 3. The selected network has two hidden layers with 8 and 10 neurons, respectively. The other ANN model is defined as ANN_2 in Table 3 uses tourism income as a function of USD and the number of visitors. The selected model architecture includes a lag of 12 and two hidden layers with 12 and 6 neurons. This selection was made based on the lowest MSE among 31 different ANN models, using the validation set, derived from combinations of five input variables.

TABLE 3 | Model Comparisons

Model	MASE	RMSSE
SARIMA	0.455	0.178
HW_M	0.447	0.214
HW_A	0.355	0.118
ETS(A,N,A)	0.499	0.245
ANN_1	0.374	0.143
ANN_2	**0.279**	**0.093**
STL-ANN	0.727	0.451

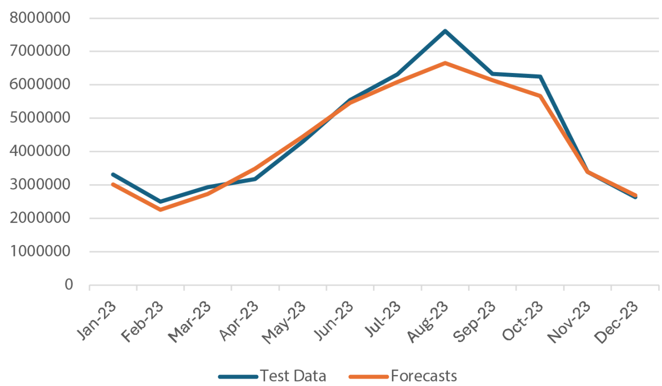
The SARIMA model yields a MASE of 0.455 and an RMSSE of 2.801. While its MASE

indicates moderate forecasting accuracy, the relatively high RMSSE compared to other models suggests it tends to produce larger errors. The HW_M model shows a MASE of 0.447 and an RMSSE of 2.687, indicating slightly better accuracy compared to SARIMA. However, it still exhibits a notable error magnitude. The HW_A model outperforms both SARIMA and HW_M with a MASE of 0.355, signifying improved accuracy. However, its RMSSE of 2.873 is high among the models tested. The ETS model has a MASE of 0.499, the highest among the models, indicating low forecast accuracy, and the RMSSE is the highest at 2.927, reflecting poorer performance compared to others. The ANN model with a lag=12 and c(8,10) achieves a MASE of 0.374 and an RMSSE of 2.680. This model delivers better forecasts compared to the univariate time series models in terms of MASE and RMSSE. The ANN model considering tourism income, USD exchange rate and the number of visitors performs the best with a MASE of 0.279 and an RMSSE of 2.266. This model not only demonstrates the highest accuracy with the lowest MASE, but also has the lowest RMSSE value. In addition to Table 3, Figure 3 shows the plot of real values and best forecasts calculated by ANN using specified inputs. The STL-ANN model, which combines decomposition of time series with ANN, has the highest MASE of 0.727 and an RMSSE of 0.451. These values indicate that this model performs worse than the other models in terms of both accuracy and error magnitude.

TABLE 4 | Hyperparameters of ANN models

Parameter	Learning Rate	Batch Size	Epochs	Optimizer	Activation Function	Loss Function
Value	0.001	30	200	Adam	Tanh	MSE

FIGURE 3 | Real and Forecasted Tourism Income Calculated by ANN with the Explanatory Variables



IV. DISCUSSION AND CONCLUSION

Tourism revenue is a significant component of many countries' economies, and its accurate forecasting plays a key role in economic planning, resource allocation, and sustainable development. This study highlights the critical role advanced forecasting models can play in accurately predicting Türkiye's tourism income. By evaluating a range of forecasting techniques, including SARIMA, HW models, ETS, STL-ANN and ANNs, the research underscores the effectiveness of incorporating complex, nonlinear models for improved accuracy.

The findings reveal that while traditional models such as SARIMA and Holt-Winters offer valuable insights, their predictive accuracy is often limited by their inability to capture intricate patterns and relationships in the data. In contrast, ANN, particularly the model incorporating additional indicators like the number of visitors and USD exchange rate, demonstrates superior performance with the lowest MASE and RMSSE. This indicates that ANNs can effectively leverage complex, multidimensional data to produce more accurate forecasts.

Traditional time series methods are widely used in tourism-related studies, but many studies emphasize the superior performance of artificial neural networks (ANNs). For example, Zorlutuna and Bircan (2019) compared time series analysis and ANN methods in predicting the number of tourists visiting Türkiye and concluded that ANN provides higher predictive accuracy. ANN models can learn nonlinear relationships and hidden patterns in data. Similarly, Çuhadar et al. (2009) compared the prediction of foreign tourism demand for Antalya province using ANN and time series methods, finding that ANN achieved greater accuracy. Karahan (2015) also evaluated the predictive performance of ANNs and successfully estimated monthly tourism demand for future periods.

The effectiveness of ANNs in tourism demand prediction is further highlighted in international literature. Law and Au (1999) used ANN to predict the number of tourists visiting Hong Kong, achieving better performance than traditional methods. Chen et al. (2007) successfully predicted tourism demand in China using ANN. Additionally, Uysal and El Roubi (1999) described ANN as a flexible and powerful tool for tourism demand prediction. Palmer et al. (2006) proposed a step-by-step methodology for designing neural networks that predicted tourism demand in Spain with high accuracy. Cho (2003) achieved successful results in tourism demand prediction using ANN in Korea, while Burger et al. (2001) reported similar success with ANN in predicting tourism demand in South Africa. Furthermore, Li et al. (2021) reviewed articles on tourism forecasting using internet data and found that, although time series and econometric forecasting models remain dominant, artificial intelligence methods are rapidly developing.

In this study, the superior performance of ANN models in predicting Türkiye's tourism revenues aligns with similar findings in the literature. ANN's ability to capture nonlinear relationships and complex data patterns allows it to generate more accurate predictions with lower error rates than traditional methods.

These results offer significant implications for decision-makers and practitioners in the tourism sector. Adopting advanced forecasting techniques can improve decision-making processes, enabling more effective allocation of resources and support for strategic planning. Considering the importance of tourism revenues in national economies, such methods can facilitate the development of more informed and effective policies.

From the perspective of policymakers, accurate forecasting methods greatly contribute to economic planning processes. Predicting fluctuations in tourism revenues allows for strategic decisions in areas ranging from budget allocation to infrastructure investments. For example, advanced forecasting models enable resources to be distributed not only to popular destinations but also to underdeveloped tourism regions. This approach can reduce economic imbalances between regions and support sustainable growth in the tourism sector. Additionally, these methods are critical during crises. The integration of data from social media analyses or travel trend monitoring allows for the rapid development of policy designs. In unforeseen situations, such as pandemics, they provide data-driven solutions for subsidies or incentive programs to support the tourism sector.

These implications are equally valuable for industry practitioners. Businesses such as hotels, airlines, and travel agencies can better analyze customer demand and adjust their strategies accordingly using such forecasting models. Forecast results help optimize pricing strategies by predicting increases in tourist demand during specific periods. Furthermore, businesses can plan their marketing campaigns more effectively and target specific audience segments. Supporting forecasting models with social media data provides insights into customers' changing expectations and preferences, leading to increased customer satisfaction and a competitive edge.

One of the most important contributions to these results is their support for sustainable development. Accurate forecasting models not only enhance economic gains but also enable better management of the environmental and social impacts of tourism activities. They provide data to prevent the negative effects of over-tourism and to promote regions with lower demand. The adoption of these models by both policymakers and industry practitioners can maximize the economic benefits of tourism while ensuring the sector's long-term sustainability. Therefore, these results are highly valuable for understanding the dynamics of the tourism sector, and for making the sector more resilient and sustainable.

Future research could enhance the findings of this study with new applications in tourism forecasting. Data such as social media trends, web search activity, and sentiment analysis could be integrated into forecasting models. These data streams have the potential to capture dynamic changes in tourist behavior and preferences, leading to more accurate and responsive forecasts.

The development of hybrid models that combine traditional econometric techniques with advanced machine learning methods could also provide deeper insights into the complex interdependencies between tourism revenue and economic indicators. For instance, integrating ARIMA with neural networks or other deep learning architectures could improve both short-term and long-term forecasting performance.

Another potential research direction involves applying these advanced forecasting techniques to specific subsectors within the tourism industry, such as ecotourism, cultural tourism, or luxury tourism. Analyzing these subsectors could uncover unique factors influencing tourism demand and offer tailored strategies for stakeholders.

Additionally, future studies could focus on cross-country comparisons to evaluate the generalizability of these forecasting models across different regions and economies. Such research could provide a more global perspective on tourism forecasting by identifying how various cultural, economic, and geopolitical factors influence the effectiveness of these models.

REFERENCES

[1] Akal, M. (2004). Forecasting Türkiye's tourism revenues by ARMAX model. *Tourism Management,* 25(5), 565-580.

[2] Apaydin, H., Sattari, M. T., Falsafian, K., & Prasad, R. (2021). Artificial intelligence modelling integrated with Singular Spectral analysis and Seasonal-Trend decomposition using Loess approaches for streamflow predictions. *Journal of Hydrology*, *600*, 126506.

[3] Aydin, O. (2016). Tourism Income of Türkiye: A panel data approach. *Procedia economics and finance, 38*, 245-256.

[4] Burger, C. J. S. C., Dohnal, M., Kathrada, M., & Law, R. (2001). A practitioners guide to time-series methods for tourism demand forecasting—a case study of Durban, South Africa. *Tourism management*, *22*(4), 403-409.

[5] Chatfield, C., Koehler, A. B., Ord, J. K., & Snyder, R. D. (2001). A new look at models for exponential smoothing. *Journal of the Royal Statistical Society: Series D (The Statistician), 50*(2), 147-159.

[6] Chen, K. Y., & Wang, C. H. (2007). Support vector regression with genetic algorithms in forecasting tourism demand. *Tourism management, 28*(1), 215-226.

[7] Cho, V. (2003). A comparison of three different approaches to tourist arrival forecasting. *Tourism management, 24*(3), 323-330.

[8] Çuhadar, M., Güngör, P., & Göksu, Y. (2009). Turizm talebinin yapay sinir ağları ile tahmini ve zaman serisi yöntemleri ile karşılaştırmalı analizi: Antalya iline yönelik bir uygulama [Forecasting tourism demand using artificial neural networks and comparative analysis with time series methods: An application for Antalya]. S*üleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi, 14*(1), 99–114.

[9] Çuhadar, M. (2020). A comparative study on modelling and forecasting tourism revenues: The case of Türkiye. *Advances in Hospitality and Tourism Research (AHTR), 8*(2), 235-255.

[10] Dritsakis, N. (2012). Tourism development and economic growth in seven Mediterranean countries: A panel data approach. *Tourism Economics, 18*(4), 801-816.

[11] Goh, C., & Law, R. (2002). Modeling and forecasting tourism demand for arrivals with stochastic nonstationary seasonality and intervention. *Tourism Management, 33*(4), 819-829.

[12] Hyndman, R., Koehler, A. B., Ord, J. K., & Snyder, R. D. (2008). *Forecasting with exponential smoothing: the state space approach.* Springer Science & Business Media.

[13] Hossen, S. M., Ismail, M. T., Tabash, M. I., & Anagreh, S. (2022). Do tourist arrivals in Bangladesh depend on seasonality in humidity? A SARIMA and SANCOVA approach. *Geo Journal of Tourism and Geosites, 41*(2), 606-613.

[14] Karahan, M. (2015). Turizm talebinin yapay sinir ağları yöntemiyle tahmin edilmesi [Forecasting tourism demand using the artificial neural networks method]. *Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi, 20*(2), 195–209.

[15] Kayakuş, M., Erdoğan, D., & Terzioğlu, M. (2023). Predicting the share of tourism revenues in total exports. *Alphanumeric Journal, 11*(1), 17-30.

[16] Kayral, İ. E., Sarı, T., & Tandoğan Aktepe, N. Ş. (2023). Forecasting the Tourist Arrival Volumes and Tourism Income with Combined ANN Architecture in the Post COVID-19 Period: The Case of Türkiye. *Sustainability, 15*(22), 15924.

[17] Kecman, V. (2001). *Learning and soft computing: support vector machines, neural networks, and fuzzy logic models.* MIT press.

[18] Kontogianni, A., Alepis, E., & Patsakis, C. (2022). Promoting smart tourism personalised services via a combination of deep learning techniques. *Expert Systems with Applications, 187,* 115964.

[19] Koyuncu, O., Gozlu, M., & Atici, K. B. (2016). Analysis and forecasts on the healthcare tourism income of Türkiye. *Journal of Economics Finance and Accounting, 3*(3), 222-233.

[20] Kůrková, V. (1992). Kolmogorov's theorem and multilayer neural networks. *Neural networks, 5*(3), 501-506.

[21] Law, R., & Au, N. (1999). A neural network model to forecast Japanese demand for travel to Hong Kong. *Tourism Management, 20*(1), 89-97.

[22] Li, X., Law, R., Xie, G., & Wang, S. (2021). Review of tourism forecasting research with internet data. *Tourism Management, 83,* 104245.

[23] Muhaimin, A., Prastyo, D. D., & Lu, H. H. S. (2021, January). Forecasting with recurrent neural network in intermittent demand data. *In 2021 11th International Conference on Cloud Computing, Data Science & Engineering (Confluence)* (pp. 802-809). IEEE.

[24] Palmer, A., Montano, J. J., & Sesé, A. (2006). Designing an artificial neural network for forecasting tourism time series. *Tourism management, 27*(5), 781-790.

[25] Republic of Türkiye Ministry of Culture and Tourism. (2023). *Türkiye tourism strategy.* Retrieved from https://www.ktb.gov.tr/TR-96696/turkiye-turizm-stratejisi.html

[26] Song, H., & Li, G. (2008). Tourism demand modelling and forecasting-A review of recent research. *Tourism Management, 29*(2), 203-220.

[27] Song, H., Witt, S. F., & Jensen, T. C. (2003). Tourism forecasting: accuracy of alternative econometric models. *International Journal of Forecasting, 19*(1), 123-141.

[28] Sönmez, S., & Sirakaya, E. (2002). A distorted destination image? The case of Türkiye. *Journal of travel research, 41*(2), 185-196.

[29] Tuncsiper, C. (2023). Modelling the Tourism Revenue of Türkiye Using Deep Learning Networks. *Open Access Indonesia Journal of Social Sciences, 6*(1), 888-897.

[30] Thomas, A. J., Petridis, M., Walters, S. D., Gheytassi, S. M., & Morgan, R. E. (2017). Two hidden layers are usually better than one. In *Engineering Applications of Neural Networks: 18th International Conference, EANN 2017, Athens, Greece, August 25–27, 2017, Proceedings* (pp. 279-290). Springer International Publishing.

[31] Uysal, M., & El Roubi, M. S. (1999). Artificial neural networks versus multiple regression in tourism demand analysis. *Journal of Travel Research, 38*(2), 111-118.

[32] Yenişehirlioğlu, E., Taşar, İ., & Bayat, T. (2020). Tourism Revenue and Economic Growth Relation in Türkiye: Evidence of Symmetrical, Asymmetrical and the Rolling Window Regressions. *Journal of Economic Cooperation & Development, 41*(2).

[33] Wong, K. K. F., Song, H., & Chon, K. K. S. (2006). Bayesian models for tourism demand forecasting. *Tourism Management, 27*(5), 773-780.

[34] World Travel & Tourism Council (WTTC). (2023). *Türkiye (Turkey) travel & tourism economic impact factsheet.* Retrieved from https://cdn.prod.website-files.com/6329bc97af73223b575983ac/66695b1694d436e5830238da_Turkiye2024_.pdf

[35] World Travel & Tourism Council (WTTC). (2024). *Travel & tourism set to break all records in 2024*. Retrieved from https://wttc.org/news-article/travel-and-tourism-set-to-break-all-records-in-2024-reveals-wttc

[36] Zorlutuna, Ş., & Bircan, H. (2019). Türkiye'ye gelen turist sayısı tahmininde zaman serileri analizi ve yapay sinir ağları yöntemlerinin karşılaştırılması [Comparison of time series analysis and artificial neural network methods in forecasting the number of tourists visiting Türkiye]. *Cumhuriyet Üniversitesi İktisadi ve İdari Bilimler Dergisi, 20*(2), 164–185