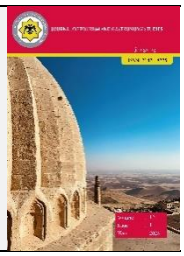
JOURNAL OF TOURISM AND GASTRONOMY STUDIES

ISSN: 2147 – 8775

Journal homepage: www.jotags.net



What if the Covid-19 Pandemic Never Happened? Estimation of the Tourist Arrivals for 2020 Via Levenberg-Marquardt Optimization and K-Star (K*) Machine Learning Algorithms

*** Ferhat ŞEKER** [a], **Alper BOZKURT** [a]

[a] Adana Alparslan Türkeş Science and Technology University, Faculty of Business Administration, Department of Tourism Management, Adana/Türkiye

Article History

Received: 29.02.2024

Accepted: 27.03.2024

Keywords

Artificial intelligence

Machine learning

Levenberg-Marquardt Optimization

K* algorithm

Article Type

Research Article

Abstract

Based on the probability theory, which is used to model uncertainty and randomness in real-world situations, the study aims to understand the impact of uncertain conditions, such as the Covid-19 pandemic, on the accuracy of the algorithms and the resulting losses to a country's tourism industry. The contribution of this paper to the international body of knowledge is twofold: firstly, it advances theoretical understanding of the use of probability theory in modelling real-world problems; and secondly, it offers a methodological approach for estimating tourist arrivals that accounts for the impact of extreme events. To achieve these aims, the Levenberg-Marquardt optimization was first applied to determine the optimal coefficients of the exponential function for estimating tourist arrivals from 1950 to 2020. Next, the K-Star machine learning algorithm was applied to the dataset with and without Covid-19 cases to estimate tourist arrivals.

***** *Corresponding Author*
E-mail: fseker@atu.edu.tr (F. Şeker)

DOI: 10.21325/jotags.2024.1402

INTRODUCTION

Estimating inbound tourist numbers can aid in optimizing capacity utilization, reduce negative impacts on host countries, and enhance tourists' participation and experience. However, this sector can be affected by many unforeseen situations, cases or diseases namely outbreaks or pandemics etc. One of the outbreaks called Covid-19 differs from any previous epidemics up to now. Actually, SARS-CoV-2 is a type of virus, a serious life-threatening disease, known as Covid-19. It was first noted initially perceived to occur in November 2019, in Wuhan/China (Zhu et al., 2020). Firstly, the World Health Organization (WHO) declared the premier case on 31st December 2019 (WHO, 2020). The outbreak was announced as a global pandemic on 11th March 2020 (WHO, 2020). Preventative precautions are provided in order to cut down the transmission of disease by tracing the patients, washing their hands, using face masks, and lockdown dispensable activities, which are run by governments (Yang et al., 2022). Owing to the lockdowns and precautions to decrease the transmission risks, considerable changes have occurred in daily lives and business manners surrounding the globe (Yang et al., 2022). Especially, offices, airports, shopping malls, entertainment sectors such as cinemas, theatres, concerts etc., colleges, schools, coffees, and railway stations are closed down so as to prevent interactions and forgathering of humans.

As a result of lockdowns and social distancing policies, the number of incoming tourists decreased in order to ensure their responsibilities against the host country. That's to say, the Covid-19 pandemic had a significant impact on the tourism industry (Hüsser, 2023), leading to a decrease in the number of tourists arriving in host countries. Many countries implemented travel restrictions and closed their borders to prevent the spread of the virus, while others imposed quarantine requirements on travellers. Additionally, many people were hesitant to travel due to health concerns and economic uncertainty. As a result, the tourism industry experienced a significant decline in revenue and many businesses were forced to close. According to a report by the World Tourism Organization (UNWTO), international tourist arrivals declined by 73% in 2020 compared to the previous year, resulting in a loss of $1.3 trillion in international tourism receipts. It has been a hard hit for many countries such as Türkiye that rely heavily on tourism.

There are several ways in which unwanted situations, such as the Covid-19 pandemic, can affect the estimation of tourist arrivals using machine learning. Some of these include lack of data, non-stationarity, unexpected events, complex relationships, overfitting, and lack of representativeness. Initially, data collection may be disrupted, leading to a lack of data on tourist arrivals, making it difficult to train and test machine learning models during a pandemic. On the other hand, The Covid-19 pandemic has led to changes in travel patterns and restrictions (Ling et al., 2023), which can make the data non-stationary. This means that the relationships between variables may change over time, making it difficult to predict future tourist arrivals. As mentioned before, the pandemic has led to unexpected events, such as lockdowns and travel bans, that can make it difficult to predict future tourist arrivals. Taking everything into account, factors that influence tourist arrivals during the pandemic are complex and may be difficult to model (Heidari, 2022). For example, the number of Covid-19 cases and travel restrictions may interact in unexpected ways, making it difficult to accurately predict future tourist arrivals. In addition to this, the model may be overfitting the data, which means it will perform well on the training data, but poorly on unseen data. As a result, the predictions may not generalize well to new situations. Hence, the data available might not be representative of the current situation, leading to poor predictions and unreliable estimates of tourist arrivals. To overcome these issues, this paper presents the modelling and estimation of the tourist arrivals into Türkiye with and without pandemic cases. It is also

important to note that the uncertainty associated with predictions made by machine learning models should be tackled by researchers and scientists.

Although artificial intelligence-based methods such as machine learning have been increasingly used in the tourism literature recently (William et al., 2019; Li et al., 2021; Bi et al., 2022; Sangkaew and Zhu, 2022; Oh and Kim, 2022; Chen et al., 2021; Yu et al., 2021) to improve various aspects of the industry, including customer service, marketing and pricing, and demand forecasting, this paper provides a comprehensive overview of the novel method in predicting tourist arrivals. The purpose of this research is to examine the effectiveness of the Levenberg-Marquardt optimization and the K-Star (K*) algorithm in predicting tourist arrivals during a pandemic. This research offers a unique perspective on the limitations of traditional machine learning algorithms, such as the Levenberg-Marquardt optimization, in predicting tourist arrivals during uncertain conditions like pandemics. Moreover, while it is widely acknowledged that the pandemic brings about economic damage, this study is crucial in that it scientifically proves the extent of that impact.

Literature Review

One of the most significant effects of Covid-19 on the tourism industry is a significant decrease in demand. Travel restrictions, lockdowns, and the fear of contracting the virus have all contributed to the drop in demand. These remarkable developments appear to have piqued the interest of tourism researchers, as studies (Chen et al., 2020; Qiu et al. 2020; Buckley & Westaway, 2020; Arabadzhyan et al., 2021; Mach & Ponting, 2021; Park et al., 2021; Yousaf, 2021; Müller & Wittmer, 2022; Yang & Smith, 2022; Hunter, 2022; Brune et al., 2023; Hüsser & Ohnmacht, 2023; Yu et al., 2023; Milone et al., 2023) and solution proposals on the subject have been published one after the other. Sigala (2020), for instance, explores the influences of Covid-19 on the tourism industry in her study and provides suggestions for advancing and resetting the industry. The author highlights the need for innovation and adaptation in the industry, including the use of digital technologies and sustainable tourism practices. In his paper, Vargas (2020) reflects on the need to adapt tourism governance models to meet the changing requirements of the industry in response to the Covid-19 crisis. Using a survey of tourism enterprises, Wang et al. (2021) investigate the impact of Covid-19 on the Chinese tourism industry. According to the authors, the pandemic has had a significant negative impact on the industry, resulting in decreased demand, cancellations, and financial losses. Yiwei et al. (2022) aim to investigate the spillover effect of industrial sectors by emphasizing the tourism sector. Regarding the Chinese and US industrial sectors in the first quarter of 2020, both countries' stocks exhibited high volatility. Throughout the pandemic period, the tourism industry suffered as well. Agarwal et al. (2022) look at how artificial intelligence (AI) has the potential to transform the tourism industry in the post-Covid-19 era. The authors discuss how AI can be used to improve customer experiences, improve sustainability, and enable personalized tourism services. In conclusion, the studies discussed here emphasize the significant impacts of Covid-19 on the tourism industry, as well as the need for innovation and adaptation in response to the crisis. Some potential strategies for transforming the industry in the post-pandemic period include the use of digital technologies, sustainable tourism practices, and artificial intelligence.

This paper also employs artificial intelligence-based methods to understand the impact of the Covid-19 pandemic. The study also presents a novel approach by excluding the year affected by the pandemic from the training and testing dataset and applying the K* algorithm to the dataset, both with and without the pandemic year. The use of the K* algorithm in this research is intriguing because it is a type of unsupervised machine learning technique that has been

used in a variety of fields, including social network analysis and bioinformatics. However, few studies have investigated its use in tourism forecasting. It will contribute to the existing literature by demonstrating the potential of machine learning and artificial intelligence techniques for this task. Hence, the total number of tourist arrivals is estimated as what would have happened if the pandemic had not occurred along with machine learning-based techniques to predict tourist arrivals and determine the direct economic losses. These techniques have the potential to improve the accuracy and reliability of forecasts and can be used to analyse large and complex datasets. It will be of interest to researchers, policymakers, and practitioners in the fields of tourism and forecasting.

Materials and Methods

Based on the literature review, material and method analysis; this study relies on estimating tourist arrivals during the Covid-19 pandemic using a combination of mathematical modelling namely exponential growth function and machine learning algorithm specifically the K* algorithm. The main reason is to using the K* algorithm as compared to kNN and LWL algorithms is that there are limited studies in the literature on the K* algorithm application in the tourism setting. On the other hand, instance-based learning algorithms generally involve the calculation of the uncertainty conditions or disorders of any system. Furthermore, these learning methods could capture complex relationships in the dataset and unexpected factors such as travel restrictions and pandemic conditions.

Initially, historical data ranging from 1950 to 2020 is collected from the Turkish Statistical Institute database since these years do not contain uncertain conditions namely the Covid-19 pandemic. Thus, the dataset covers pre-pandemic years. The number of tourist arrivals based on the pre-pandemic years is analysed with respect to the exponential growth function optimized by the Levenberg-Marquardt algorithm. After obtaining the closed form mathematical expressions of the number of tourist arrivals are calculated during pandemic periods. To validate this numerical analysis, the K* algorithm is utilized for the estimation of the number of tourists. Sensitivity and robustness are checked by comparing the numerical analysis and estimation results along with the calculated and predicted values for pandemic periods. In other words, cross-validation is evaluated in a binary way including exponential growth function and estimation of the proposed algorithm.

In conclusion, this study adds a great deal to the body of literature by using both machine learning and mathematical modelling, carrying out thorough testing, and offering a thorough examination of visitor arrivals over a long-time span. Understanding the nuances of various algorithms' applicability in the context of tourism forecasting is enhanced by comparing them.

In recent years, the field of tourist arrival forecasting has seen a growing interest, and various methods have been proposed and tested in the literature. In general, commonly used techniques with their pros and cons for forecasting procedure is given in detail in Table 1.

Method	Pros	Cons
Time Series Analysis	• Simple and easy to understand • Good for forecasting trends and patterns • Works well with univariate data	• Limited in its ability to handle external factors that may impact revenue • May not work well with irregular or non-stationary data
Regression Analysis	• Can handle multiple independent variables • Can account for external factors that may impact revenue • Can provide insights into the relationship between independent and dependent variables	• May not work well with non-linear relationships • May not work well with time-series data that have a strong autocorrelation
ARIMA Models	• Good for handling time-series data • Can handle non-stationary data • Can handle trend and seasonality in data	• Limited in its ability to handle external factors that may impact revenue • Can be complex to implement
Exponential Smoothing	• Simple to implement • Good for handling time-series data • Can handle trend and seasonality in data	• Limited in its ability to handle external factors that may impact revenue • May not work well with irregular or non-stationary data
Artificial Neural Networks	• Good for handling complex data relationships • Can handle multiple inputs and outputs • Can handle non-linear relationships	• Can be complex to implement • May be prone to overfitting if not properly validated

Table 1. A Comparison Table of Some Commonly Used Forecast Methods for International Tourist Arrivals

It's important to bear in mind that the pros and cons of each method will vary depending on the specific situation and the characteristics of the data. Additionally, it's often useful to use multiple methods and compare their results to get a better overall understanding of the situation. The K-Star (K*) Algorithm and the Levenberg-Marquardt Algorithm are two different methods used in the field of machine learning for forecasting international tourist arrivals. The K* Algorithm is a clustering algorithm that can identify patterns and structures in data. It works by dividing the data into groups or clusters based on similarity and then using these clusters to make predictions. The K* Algorithm is suitable for data with clear patterns and structures and is easy to interpret. However, it may not be suitable for data with complex relationships or non-linear patterns. The Levenberg-Marquardt Algorithm is an optimization algorithm that can minimize the error between the predicted and actual values of a model. It works by adjusting the parameters of a model to minimize the error between the predicted and actual values. The Levenberg-Marquardt Algorithm is suitable for data with complex relationships and is good for optimizing the parameters of models. However, it may not be suitable for data with clear patterns and structures and can be complex and difficult to interpret. In terms of international tourist arrivals, the choice of method will depend on the specific characteristics of the data. If the data has clear patterns and structures, the K* Algorithm may be a good choice, while if the data has complex relationships, the Levenberg-Marquardt Algorithm may be a better choice. However, it is important to note that the accuracy of the predictions will also depend on the quality of the data and the specific parameters used in each method.

A comparison table of the K* Algorithm and the Levenberg-Marquardt Algorithm in terms of international tourist arrivals may include the following aspects:

Feature	K-Star (K*) Algorithm	Levenberg-Marquardt Algorithm
Suitability	Suitable for data with clear patterns and structures	Suitable for data with complex relationships
Interpretability	Easy to interpret	Complex and difficult to interpret
Accuracy	Depends on the quality of the data and the specific parameters used	Depends on the quality of the data and the specific parameters used

Table 2. Comparison of the K* and Levenberg-Marquardt Algorithms

It is important to note that this comparison table is not exhaustive, and there may be other factors that are important to consider when choosing a method for forecasting international tourist arrivals. In this paper, two algorithms K* and Levenberg-Marquardt, are utilized to apply the estimating the international tourist arrivals. However, this paper differs from other studies due to the fact that an optimization algorithm is used for the coefficient determination of the exponential function. Then, this paper suggests the efficacy of both the K* algorithm and the Levenberg-Marquardt algorithm application to determine if the covid-19 does not occur.

Exponential Functions and Optimization Algorithm

Exponential functions are mathematical functions that can be used to model a wide range of phenomena, including growth and decay. In the context of tourist arrivals, exponential functions can be used to model the rate at which tourist numbers change over time. One common type of exponential function is the exponential growth function, which is given by the equation in Table 3.

y = a * e^(kt)

Table 3. The Exponential Growth Function

Where y is the number of tourist arrivals, a is the initial number of tourists, k is the growth rate, and t is the time. The growth rate k can be positive or negative, depending on whether the number of tourists is increasing or decreasing. Another common type of exponential function is the exponential decay function, which is given by the equation in Table 4.

y = a * e^(-kt)

Table 4. The Exponential Decay Function

This function models the rate at which the number of tourists is decreasing over time. Exponential functions can be useful for modelling tourist arrivals because they can capture the non-linear relationship between the number of tourists and time. They can also be used to estimate the number of tourists in the future by extrapolating from historical data. It's worth noting that these functions are the idealized representation of reality, in practice the data may not exactly follow this pattern, and the estimation may need to be adjusted by other factors.

The Levenberg-Marquardt algorithm is an optimization method that can be used to estimate the parameters of an exponential function for the purpose of modelling tourist arrivals. The exponential function that is commonly used for modelling tourist arrivals is the exponential growth function, which is given by the equation in Table 5.

y = a * e^(kt)

Table 5. The Exponential Growth Function for Estimating Tourist Arrivals

Where y is the number of tourist arrivals, a is the initial number of tourists, k is the growth rate, and t is the time. The goal is to estimate the parameters a and k from the available data. The Levenberg-Marquardt algorithm is an iterative method that finds the best estimates of the parameters by minimizing the sum of the squares of the residuals, which are the differences between the observed data and the predicted values. The algorithm starts with an initial estimate of the parameters and iteratively updates them by adjusting them in the direction of the negative gradient of the cost function. One of the advantages of the Levenberg-Marquardt algorithm is that it is relatively fast and efficient, and can find a good solution even when the data is noisy or has outliers. The Levenberg-Marquardt algorithm can be applied to estimate the parameters of an exponential function for the purpose of modelling tourist arrivals by following these steps in Table 6.

Levenberg-Marquardt Optimization Algorithm Principle: Estimation of the tourist arrivals
• Collect data on the number of tourist arrivals over time. • Select an initial estimate of the parameters a and k. • Use the Levenberg-Marquardt algorithm to iteratively update the estimates of the parameters by minimizing the sum of the squares of the residuals. • Once the algorithm has converged, the final estimates of the parameters are the best estimates of the number of tourist arrivals.

Table 6. Levenberg-Marquardt Optimization Steps

It's worth noting that this method assumes that the data follows the exponential function, in practice the data may not exactly follow this pattern, and the estimation may need to be adjusted by other factors. Also, the Levenberg-Marquardt algorithm is a local optimization method, it's important to check that the solution found is not a local minimum.

Machine Learning Method

Estimating the number of tourist arrivals during the Covid-19 pandemic using machine learning algorithms would involve using historical data and other relevant factors to make predictions about future tourist arrivals. Some possible steps for using machine learning to estimate tourist arrivals during the pandemic might be included as given in the proposed steps. Collecting and pre-processing data, such as the number of confirmed Covid-19 cases, travel restrictions, and economic indicators, as well as historical data on tourist arrivals. In this paper, historical data is retrieved from the Turkish Statistical Institute database including the years from 2000 to 2020 (Turkish Statistical Institute, 2022) and the previous historical data is obtained from the study of Ünlüönen and Kılıçlar (2004), which includes the years from 1950 to 2000. After gathering the dataset, it was trained via machine learning model K-star (K*), by using the total tourist arrivals with and without pandemic cases.

K* algorithm is improved by Clearly and Trigg (1995), which is a type of instance-based learning algorithm. In general, the working principle of the instance-based learning algorithm is to compare the instances via the metric distances including Manhattan, Euclidean, Chebyshev etc. However, the K* instance-based algorithm depends on the entropy calculation based on the probability theory. There are limited studies based on the K* algorithm in the literature as compared to other instance-based learning algorithms namely k-nearest neighbour(kNN) and locally weighted learning (LWL). Thus, the effectiveness of the K* algorithm in the tourism sector is analysed and interpreted by comparing kNN and LWL. The model was tested using a holdout or cross-validation set to evaluate its performance. The model was used to make predictions about future tourist arrivals, considering current conditions such as travel restrictions and Covid-19 case numbers. The model was regularly updated with new data and retrained

as the situation evolved.

Exponential functions and machine learning algorithms are both methods that can be used for estimating the number of tourist arrivals, but they have some key differences. Exponential functions typically require a relatively small amount of data and can be used to model simple relationships between variables. Machine learning algorithms, on the other hand, typically require large amounts of data and are capable of modelling complex relationships (Chen et al., 2022). Flexibility Exponential functions are fixed mathematical equations that can model a limited set of relationships. Machine learning algorithms, on the other hand, can be trained to model a wide range of relationships and can adapt to changing conditions (Ahmed et al., 2022). Exponential functions can be used to make predictions about future tourist arrivals by extrapolating from historical data. Machine learning algorithms, however, can be used to make predictions based on a wide range of factors and can consider changing conditions. Exponential functions may not generalize well to new situations and may not capture the complexity of real-world data (Sutthimat et al., 2022). Machine learning algorithms can generalize well and can be adapted to new situations, provided that they are trained on large and diverse datasets. Exponential functions are simple and easy to understand, the parameters of the function can be interpreted, and the results are easy to interpret. Machine learning algorithms are more complex and may be difficult to interpret (Yanzhou et al., 2022). Overall, both exponential functions and machine learning algorithms can be used for estimating the number of tourist arrivals, but machine learning algorithms are more powerful and flexible, and can be used to make more accurate predictions. However, exponential functions are simpler and easier to understand. The choice of method will depend on the specific use case and the data available.

Results

The Levenberg-Marquardt optimization algorithm is applied to determine the optimal coefficients of the exponential function for the estimation of tourist arrivals. Initially, the total tourist arrivals are given along with the years, which range from 1950 to 2020 as given in Figure 1. When the number of tourists from 1950 to 2020 is examined, there has been a sharp decrease in 2020 due to the Covid-19 pandemic, and this decrease does not fit the exponential function. Hence, the tourist arrivals in 2020 are not included to train and test the dataset effectively.

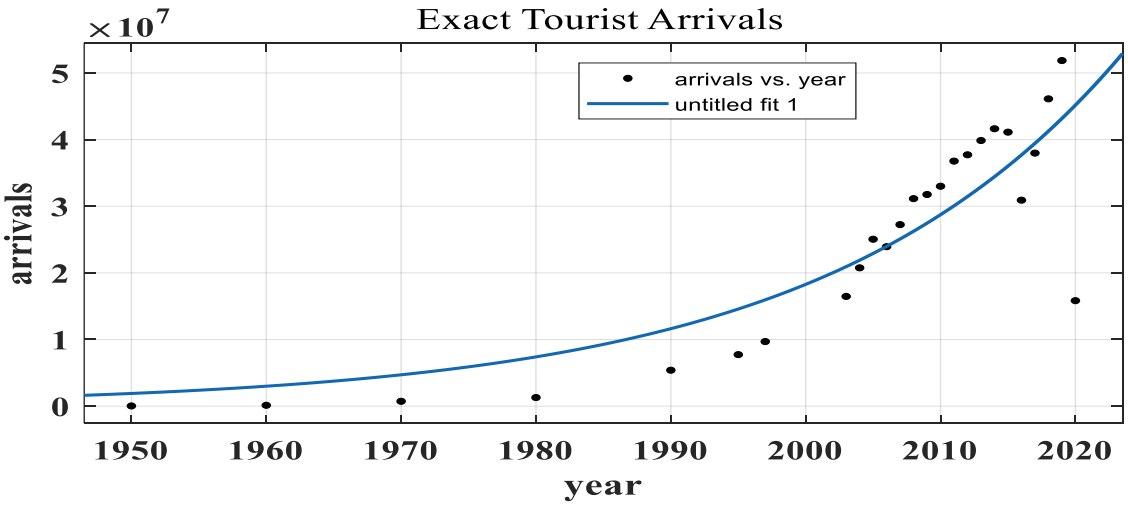


Figure 1. The Exact Tourist Arrivals from 1950 to 2020

Based on the dataset, exponential function coefficients of tourist arrivals, are optimized as a=9.1e-44 and b=0.05787, respectively. During the optimization procedure, a 95% confidence boundary is applied. The optimized

exponential function is obtained as follows in Table 7.

$$y= (9.1x10^{-44})\ e^{0.05787t}$$

Table 7. The Optimized Exponential Function

The detailed code of the algorithm is also depicted in Table 8.

```
function [fitresult, gof] = createFit(Year, Arrivals)
%CREATEFIT (YEAR, ARRIVALS)
% Create a fit.
% Data for 'Fitted Data' fit:
%      X Input: Year
%      Y Output: Arrivals
% Output:
%      fit result: a fit object representing the fit.
%      gof: structure with goodness-of fit info.

%% Fit: 'Fitted Data'.
[xData, yData] = prepareCurveData (Year, Arrivals);

% Set up fittype and options.
ft = fittype (‘exp1’);
opts = fitoptions(‘Method', 'NonlinearLeastSquares' );
opts.Algorithm = 'Levenberg-Marquardt';
opts.Display = 'Off';
opts.Normalize = 'on';
opts.StartPoint = [10450313.0866315 1.87239738169156];

% Fit model to data.
[fitresult, gof] = fit (xData, yData, ft, opts);

% Plot fit with data.
figure (‘Name', 'Fitted Data’);
h = plot (fitresult, xData, yData, 'predobs’);
legend (h, 'Arrivals vs. Year', 'Fitted Data', 'Lower bounds (Fitted Data)', 'Upper bounds (Fitted Data)', 'Location', 'NorthEast',
'Interpreter', 'none’);
% Label axes
xlabel(‘Year', 'Interpreter', 'none' );
ylabel(‘Arrivals', 'Interpreter', 'none' );
grid on
```

Table 8. Levenberg-Marquardt Algorithm Based Optimization

After applying K* algorithm to the dataset with and without Covid-19 cases, the correlation coefficients and predicted tourist arrivals are obtained. The correlation coefficient with and without Covid-19 cases is calculated as 0.87 and 0.97, respectively. This reveals that uncertain conditions such as pandemic outbreaks cause a sharp decrease in tourist arrivals and in the effectiveness of the algorithm as well. Table 9 depicts the estimated and actual tourist numbers for 2020 based on the K* machine learning algorithm.

t	$y= (9.1x10^{-44})\ e^{0.05787t}$	Exact
2020	53325773	15826266

Table 9. Estimated and Exact Tourist Numbers for 2020

Conclusion

A pandemic outbreak, such as Covid-19, can significantly impact tourism. This can include a decrease in international travellers as well as changes to travel plans and destinations (Vargas, 2020). One significant impact of

the pandemic on tourism has been the restriction of international travel as countries implemented quarantine measures and closed their borders to reduce the virus's spread. This has resulted in a significant decrease in the number of tourists visiting destinations, resulting in a decrease in revenue for the tourism industry. In addition to travel restrictions, the pandemic has resulted in a decrease in consumer confidence and disposable income, both of which have impacted the tourism industry. Many individuals and families have lost their jobs or had their income reduced, making it more difficult for them to travel. According to Morakabati (2020), political crises can also have an impact on tourism by causing a lack of stability and safety in a destination, making it less appealing to visitors. Tourists may be less likely to visit a destination if there is a civil war or an increase in terrorism, for example. This can reduce revenue for the tourism industry and have a negative impact on the local economy. Besides, extraordinary statistics caused by unforeseen events such as wars, pandemics, and political crises can make it difficult to estimate tourism demand (Flectcher and Morakabati, 2008). The impact of uncertain conditions on algorithms has been a topic of interest in the literature. In the field of machine learning, many studies have explored the effect of changes in data distribution or characteristics on the performance of models. For example, Wyatt et al. (2022) investigated the impact of dataset shifts on the performance of deep learning models for image classification. They found that the models performed poorly when tested on data that differed significantly from the training data.

Initially, the Levenberg-Marquardt optimization algorithm was employed to determine the best exponential function coefficients for estimating tourist arrivals from 1950 to 2020. However, due to the dramatic decrease in tourist numbers during the pandemic, 2020 did not fit the exponential function and was thus excluded from the training and testing datasets. Regarding the specific approach used in this study, the Levenberg-Marquardt optimization method has been widely used in fitting exponential functions to data in a range of fields. (e.g., neuroscience, physics, economics). However, because the exponential function is relatively new to tourism theory, it should be emphasized that it is not always used for modelling tourism data. Nonetheless, research has proposed that other functions or models, such as ARIMA or neural networks, be used for tourism forecasting. (Chang & Liao, 2010; Palmer et al., 2006). The K* algorithm was utilized, then, to estimate tourist arrivals on the dataset, both with and without Covid-19 cases. The application of the K* algorithm, both with and without the year affected by Covid-19, provides insight into the influence of uncertain conditions on machine learning algorithms and the importance of incorporating such conditions in the training and testing dataset. The correlation coefficient was found to be higher without the Covid year. Therefore, this research tried to estimate that year by ignoring the number of tourists arriving in the year of covid cases.

According to the K* machine learning algorithm, if the pandemic had not occurred, more than 53 million tourists would have visited Türkiye in 2020, rather than the actual 15 million. According to the United Nations World Tourism Organization data, since a tourist spends an average of 765 dollars in Türkiye, this difference of 38 million means a loss of 29 billion dollars to the country's economy (approx. 3% of Türkiye's GDP). Although everyone accepts the damage caused by the pandemic to the national economies, empirically proving how much this damage is makes this study unique. The findings of the research emphasize the negative impact that uncertain conditions, such as pandemics, can have on the efficacy of algorithms as well as tourism revenue. Within the tourism literature, there have been studies that have looked at the impact of external factors on the accuracy of forecasting models. Cho (2001), for instance, examined the effect of economic fluctuations on the forecasting accuracy of tourist arrivals in Hong Kong. It is found that the accuracy of the forecasting models decreased during periods of economic instability.

The study results suggest that excluding years with drastic decreases in tourist numbers from the training and testing dataset could lead to more accurate estimations. The results also demonstrate the potential of machine learning algorithms, such as the K* algorithm, in providing accurate estimations for the tourism industry, despite the challenges posed by uncertain conditions like pandemics. The use of the K* algorithm in this research is also intriguing because it is a type of unsupervised machine learning technique that has been used in a variety of fields, including social network analysis and bioinformatics. However, few studies have investigated its use in tourism estimation, and more research may be required to compare its performance to other more established methods.

The theoretical framework of this study is based on probability theory, which provides a powerful tool for modelling and analysing uncertain and random phenomena. Specifically, the study employs probability theory to estimate the accuracy of algorithms and quantify the losses incurred by a country's tourism industry under uncertain conditions, such as the Covid-19 pandemic. The contribution of this research article to the international body of knowledge is twofold: firstly, it advances theoretical understanding of the use of probability theory in modelling real-world problems; and secondly, it offers a methodological approach for estimating tourist arrivals that accounts for the impact of extreme events. The study shows how Covid-19 affects the accuracy of the algorithms used to estimate tourist arrivals. The authors acknowledged the drastic decrease in tourist numbers due to the pandemic by excluding 2020 from the training and testing datasets, resulting in a deviation from the exponential function. This method ensures that the model is not biased toward pre-pandemic data and that estimations are more accurate. Furthermore, the use of the K* machine learning algorithm to estimate tourist arrivals with and without Covid-19 cases in the study sheds light on the extent of the pandemic's impact on the tourism industry. The fact that the correlation coefficient was higher in the absence of the Covid year suggests that the pandemic disrupted the relationship between the variables in the model. The estimated loss to the Turkish economy of 29 billion dollars as a result of the pandemic emphasizes the importance of understanding the effects of uncertain conditions on tourism forecasting and planning.

The findings of this research are important for tourism policymakers, academics, and industry practitioners because they shed light on the potential effect of future pandemics or other uncertain conditions on tourism demand. As Vargas (2018) has noted, while tourism has demonstrated a high degree of resilience in rebounding from adverse events in a relatively short period, the industry's managers and destinations must not underestimate the effects of such events. This study's research approach and methodology can be replicated in other countries to estimate the impact of the pandemic on their tourism sector. Furthermore, the study offers an excellent opportunity for future research in tourism forecasting and planning. Future research can build on this work by investigating the effect of other uncertain conditions on tourism demand and developing more accurate forecasting models that take these factors into consideration. Additionally, further research can explore the impact of the pandemic on different tourism sectors, such as accommodation, food and beverage, and transportation, to provide a more comprehensive understanding of its impact on the tourism industry.

In conclusion, this research article contributes significantly to the international body of knowledge in tourism theory by emphasizing the importance of considering uncertain conditions when developing tourism forecasting models and providing a methodological approach for estimating tourism industry losses due to the Covid-19 pandemic. Its findings have practical implications for tourism policymakers, academics, and industry practitioners, and provide a foundation for future research in this field.

Deceleration

All authors of the article contributed equally to the article process. The authors have no conflicts of interest to declare.

REFERENCES

Agarwal, P., Swami, S., & Malhotra, S. (2022). Artificial Intelligence Adoption in the Post COVID-19 New-Normal and Role of Smart Technologies in Transforming Business: a Review. *Journal of Science and Technology Policy Management*. DOI: http://dx.doi.org/10.1108/JSTPM-08-2021-0122.

Ahmed, A., Yafouz, A., Birima, A., Kisi, O., Huang, Y., Mohsen S., Sefelnasr, A., & El-Shafie, A. (2022). Water level prediction using various machine learning algorithms: a case study of Durian Tunggal river, Malaysia. *Engineering Applications of Computational Fluid Mechanics*, 16 (1), 422-440. DOI: 10.1080/19942060.2021.2019128.

Arabadzhyan, A., Figini, P., & Zirulia, L. (2021). Hotels, prices and risk premium in exceptional times: The case of Milan hotels during the first COVID-19 outbreak. *Annals of Tourism Research Empirical Insights*, 2, 100023. DOI: https://doi.org/10.1016/j.annale.2021.100023.

Bi, J.-W., Han, T.-Y., & Li, H. (2022). International tourism demand forecasting with machine learning models: The power of the number of lagged inputs. *Tourism Economics*, 28 (3), 621–645. DOI: https://doi.org/10.1177/1354816620976954.

Brune, S., Knollenberg, W., & Vila, O. (2023). Agritourism resilience during the COVID-19 crisis. *Annals of Tourism Research*, 99, 1035238. DOI: https://doi.org/10.1016/j.annals.2023.103538.

Buckley, R., & Westaway, D. (2020). Mental health rescue effects of women's outdoor tourism: A role in COVID-19 recovery. *Annals of Tourism Research*, 85, 103041. DOI: https://doi.org/10.1016/j.annals.2020.103041.

Chang, Y.W., & Liao, M.Y. (2010) A Seasonal ARIMA Model of Tourism Forecasting: The Case of Taiwan. *Asia Pacific Journal of Tourism Research*, 15:2, 215-221, DOI: 10.1080/10941661003630001.

Chen, M., Demir, E., Gomez, C., & Zaremba, A. (2020). The impact of policy responses to COVID-19 on U.S. travel and leisure companies. *Annals of Tourism Research Empirical Insights*, 1, 100003. DOI: http://dx.doi.org/10.1016/j.annale.2020.100003.

Chen, S., Wang, X., Zhang, H., Wang, J., & Peng, J. (2021). Customer purchase forecasting for online tourism: A data-driven method with multiplex behavior data. *Tourism Management*, Volume 87, 104357, ISSN 0261-5177. DOI: https://doi.org/10.1016/j.tourman.2021.104357.

Chen, X., Wu, H., Lichti, D., Han, X., Ban, Y. Li, P., & Deng, H. (2022). Extraction of indoor objects based on the exponential function density clustering model. *Information Sciences*, Volume 607, Pages 1111-1135, ISSN 0020-0255. DOI: https://doi.org/10.1016/j.ins.2022.06.032.

Cho, V. (2001). Tourism Forecasting and its Relationship with Leading Economic Indicators. *Journal of Hospitality & Tourism Research*, 25(4), 399–420. DOI: https://doi.org/10.1177/109634800102500404.

Cleary, J., & L. Trigg (1995). K*: An Instance-based Learner Using an Entropic Distance Measure, in *12th International Conference on Machine Learning*. p. 108-114.

Fletcher, J., & Morakabati, Y. (2008). Tourism activity, terrorism and political instability within the commonwealth: the cases of Fiji and Kenya. *International Journal of Tourism Research*, 10 (6), 537-556. DOI: https://doi.org/10.1002/jtr.699.

Fu, Y. Downey, A., Yuan, L., Zhang, T., Pratt, A., & Balogun, Y. (2022). Machine learning algorithms for defect detection in metal laser-based additive manufacturing: A review. *Journal of Manufacturing Processes*, Volume 75, 2022, Pages 693-710, ISSN 1526-6125. DOI: https://doi.org/10.1016/j.jmapro.2021.12.061.

Heidari, A., Navimipour, N., Unal, M., & Toumaj, S. (2022). Machine learning applications for COVID-19 outbreak management. *Neural Computing & Applications*, 34, 15313–15348. DOI: https://doi.org/10.1007/s00521-022-07424-w.

Hunter, W. (2022). Semiotic fieldwork on chaordic tourism destination image management in Seoul during COVID-19. *Tourism Management*, 93, 104565. DOI: https://doi.org/10.1016/j.tourman.2022.104565.

Hüsser, A.P., & Ohnmacht, T., A. (2023). A comparative study of eight COVID-19 protective measures and their impact on swiss tourists' travel intentions. *Tourism Management*, 97, 104734. DOI: https://doi.org/10.1016/j.tourman.2023.104734.

Joanne Y., & Roman E. (2021). Color and engagement in touristic Instagram pictures: A machine learning approach. *Annals of Tourism Research*, Volume 89, 103204, ISSN 0160-7383. DOI: https://doi.org/10.1016/j.annals.2021.103204.

Li, X., Li, H., Pan, B., & Law, R. (2021). Machine Learning in Internet Search Query Selection for Tourism Forecasting. *Journal of Travel Research*, 60 (6), 1213–1231. DOI: https://doi.org/10.1177/0047287520934871.

Mach, L., & Ponting, J. (2021). Establishing a pre-COVID-19 baseline for surf tourism: Trip expenditure and attitudes, behaviors and willingness to pay for sustainability. *Annals of Tourism Research Empirical Insights*, 2, 100011. DOI: http://dx.doi.org/10.1016/j.annale.2021.100011.

Milone, F., Gunter, U., & Zekan, B. (2023). The pricing of European airbnb listings during the pandemic: A difference-in-differences approach employing COVID-19 response strategies as a continuous treatment. *Tourism Management*, 97, 104738. DOI: https://doi.org/10.1016/j.tourman.2023.104738.

Morakabati, Y. (2020). A question of confidence. Is tourism as vulnerable to civil unrest as we think? A comparative analysis of the impact of Arab Spring on total reserves and tourism receipts. *International Journal of Tourism Research*, 22 (2), 252-265. DOI: https://doi.org/10.1002/jtr.2333.

Müller, A., & Wittmer, A. (2023). The choice between business travel and video conferencing after COVID-19 – Insights from a choice experiment among frequent travelers. *Tourism Management*, 96, 104688. DOI: https://doi.org/10.1016/j.tourman.2022.104688.

Oh, M. & Kim, S. (2022). Role of Emotions in Fine Dining Restaurant Online Reviews: The Applications of Semantic Network Analysis and a Machine Learning Algorithm. International *Journal of Hospitality & Tourism*

Administration, 23 (5), 875-903. DOI: 10.1080/15256480.2021.1881938.

Qiu, R., Park, J., Li, S., & Song, H. (2020). Social costs of tourism during the COVID-19 pandemic. *Annals of Tourism Research*, 84, 102994. DOI: https://doi.org/10.1016/j.annals.2020.102994.

Palmer, A., Montano, J.J., & Sese, A. (2006). Designing an artificial neural network for forecasting tourism time series. *Tourism Management*, 27 (5): 781-790. DOI: https://doi.org/10.1016/j.tourman.2005.05.006.

Park, I., Kim, J., Kim, S., Lee, J., & Giroux, M. (2021). Impact of the COVID-19 pandemic on travelers' preference for crowded versus non-crowded options. *Tourism Management*, 87, 104398. DOI: https://doi.org/10.1016/j.tourman.2021.104398.

Sangkaew, N. & Zhu, H. (2022). Understanding Tourists' Experiences at Local Markets in Phuket: An Analysis of TripAdvisor Reviews. *Journal of Quality Assurance in Hospitality & Tourism*, 23 (1), 89-114. DOI: 10.1080/1528008X.2020.1848747.

Sigala, M. (2020). Tourism and COVID-19: Impacts and implications for advancing and resetting industry and research. *Journal of Business Research*, 117, 312-321. DOI: https://doi.org/10.1016/j.jbusres.2020.06.015.

Sutthimat, P., Rujivan, S., Mekchay, K., & Rakwongwan, U. (2022). Analytical formula for conditional expectations of path-dependent product of polynomial and exponential functions of extended Cox–Ingersoll–Ross process. *Research in the Mathematical Sciences*, 9 (10). DOI: https://doi.org/10.1007/s40687-021-00309-9.

Turkish Statistical Institute. Statistical Tables, Tourism Income and number of visitors. Available online: https://data.tuik.gov.tr (accessed on 30 December 2022).

Ünlüönen K, & Kiliçlar A. (2004). Eighty Years of Turkish Tourism with Economical Reflections. *Journal of Gazi University Faculty of Commerce and Tourism Education*, 0(1), 131 - 156.

Vargas-Sánchez, A. (2018). Crisis Situations in Tourist Destinations: How Can They Be Managed? Enlightening Tourism. *A Pathmaking Journal*, 8(1), 47-69. DOI: https://doi.org/10.33776/et.v8i1.3439.

Vargas, A. (2020). Covid-19 crisis: a new model of tourism governance for a new time. *Worldwide Hospitality and Tourism Themes*, 12 (6), 691-699. DOI: https://doi.org/10.1108/WHATT-07-2020-0066.

Wang, C., Meng, X., Siriwardana, M., & Pham, T. (2022). The impact of COVID-19 on the Chinese tourism industry. *Tourism Economics*, 28(1), 131–152. DOI: https://doi.org/10.1177/13548166211041209.

William L., Park, S., Pan, B., & Newman, P. (2019). Forecasting campground demand in US national parks. *Annals of Tourism Research*, Volume 75, Pages 424-438, ISSN 0160-7383. DOI: https://doi.org/10.1016/j.annals.2019.01.013.

World Health Organization. (2020). WHO Director-General's opening remarks at the media briefing on COVID-19-11 March 2020.

World Health Organization. (2020). *Global Surveillance for Human Infection with Novel Coronavirus (2019-Ncov): Interim Guidance*, 21 January 2020 (No. WHO/2019-nCoV/Surveillance Guidance/2020.1). World Health Organization.

World Tourism Organization. (2021). Tourism Highlights 2021 Edition. Retrieved from

https://www.unwto.org/tourism-highlights-2021-edition.

Wyatt, M., Radford, B., Callow, N., Bennamoun, M. & Hickey, S. (2022). Using ensemble methods to improve the robustness of deep learning for image classification in marine environments. *Methods in Ecology and Evolution*, 13, 1317– 1328. DOI: https://doi.org/10.1111/2041-210X.13841.

Yang, E., & Smith, J. (2023). The spatial and temporal resilience of the tourism and outdoor recreation industries in the United States throughout the COVID-19 pandemic. *Tourism Management*, 95, 104661. DOI: https://doi.org/10.1016/j.tourman.2022.104661.

Yang, H., Wang, L., Xu, Y., & Liu, X. (2023). CovidViT: a novel neural network with self-attention mechanism to detect Covid-19 through X-ray images. *International Journal of Machine Learning and Cybernetics*. DOI: https://doi.org/10.1007/s13042-022-01676-7.

Yang, Y., Tian, X., Ng, W., Wang, R., & Kwong, S. (2022). Generative face inpainting hashing for occluded face retrieval. *International Journal of Machine Learning and Cybernetics*. DOI: https://doi.org/10.1007/s13042-022-01723-3.

Yiwei, W., Najaf, K., Frederico, G., & Atayah, O. (2022). Influence of COVID-19 pandemic on the tourism sector: evidence from China and United States stocks. *Current Issues in Tourism*, 23 (23), 3783-3798. DOI: https://doi.org/10.1080/13683500.2021.1972944.

Yousaf, S. (2021). Travel burnout: Exploring the return journeys of pilgrim-tourists amidst the COVID-19 pandemic. *Tourism Management*, 84, 104285. DOI: https://doi.org/10.1016/j.tourman.2021.104285.

Yu, L., Zhao P., Tang, J., & Pang, L. (2023). Changes in tourist mobility after COVID-19 outbreaks. *Annals of Tourism Research*, 98, 103522. DOI: https://doi.org/10.1016/j.annals.2022.103522.

Zhu N., Zhang D., Wang W., Li X., Yang B., Song J., & Tan, W. (2020). A novel coronavirus from patients with pneumonia in China, 2019. *New England Journal of Medicine*. DOI: 10.1056/NEJMoa2001017.