Forecasting Tourism Demand in the Lower Northern Provinces of Thailand

Chalermpat Pongajarn[1], Natthawudh Konglumpun[2] and Kritsada Wattasaovaluk[3]

Received: December 19, 2023
Revised: June 21, 2024
Accepted: June 28, 2024

ABSTRACT

Tourism demand forecasting is critical for strategic planning and resource allocation, especially in emerging tourist destinations that often lack the necessary knowledge for effective resource management planning. The objective of this study is to determine the best appropriate predictive models for effectively projecting tourism demand in the Lower Northern Provinces of Thailand, including Phitsanulok, Uttaradit, Sukhothai, Tak, and Phetchabun. These provinces are emerging tourist destinations with underdeveloped tourism infrastructure compared to major cities. The study employs four different forecasting models: Holt-Winter's Exponential Smoothing, the Box-Jenkins approach (ARIMA), Artificial Neural Networks (ANN), and the Trigonometric Box-Cox ARMA Trend Seasonal Model (TBATS). The performance of these models is evaluated using Mean Absolute Percentage Error (MAPE) and Root Mean Square Error (RMSE), based on monthly tourist arrival data from January 2013 to December 2018. The findings indicate that the ARIMA model provides the highest accuracy for Phitsanulok, Uttaradit, Sukhothai, and Tak, while the ANN model performs best for Phetchabun. This study highlights the importance of selecting forecasting methods that align with the specific data pattern.

Keywords: Forecasting Models, Tourism Demand, Lower North Region, Artificial Neural Networks (ANN), TBATS Method.

[1] Lecturer, Faculty of Business, Economics and Communications, Naresuan University,
E-mail: chaleampatp@nu.ac.th

[2] Lecturer, Faculty of Business, Economics and Communications, Naresuan University,
E-mail: natthawudhk@nu.ac.th

[3] Lecturer, Faculty of Business, Economics and Communications, Naresuan University,
E-mail: kritsadaw@nu.ac.th

Background and Significance of the Research Problem

The tourism industry plays a vital role in economic development. The tourism industry is the most significant sector globally in terms of generating income and creating employment opportunities (Gokovali and Bahar, 2006). Thailand's gross value added of tourism industries (GVATI) in 2020 amounted to 995,006 million baht, which accounted for 5.65% of the country's GDP (Office of the Permanent Secretary of Ministry of Tourism and Sports, 2021). To ensure the resilience and re-balancing of Thai tourism, a crucial strategy is the equitable distribution of income and prosperity across all regions (Office of the Permanent Secretary of Ministry of Tourism and Sports, 2023). This approach involves attempts to promote tourism in less visited destinations or emerging tourist destinations. There is a significant disparity in the size of the tourism sector between the two categories of destinations. In 2020, the Lanna Civilization Tourism Development Zone, which includes popular tourist destinations like Chiang Mai, recorded a tourism expenditure of 67,771.72 million baht. In contrast, the World Cultural Heritage Tourism Development Zone, which encompasses the emerging tourist destinations in the lower northern region, only received 8,216.79 million baht in tourist spending (Office of the Permanent Secretary of Ministry of Tourism and Sports, 2021). Hence, the advancement of tourism in emerging tourist destinations would contribute to enhancing the stability of the nation's tourism revenue. Despite the acknowledged importance of promoting tourism development in emerging destinations, yet there is still a notable deficiency in research and crucial information, such as tourism demand, that is necessary for creating community tourism policies for planned development.

Forecasting tourism demand is essential for stakeholders in the tourism industry. It allows for effective strategic planning, resource allocation, and decision-making. Precise forecasting has practical implications that lead to several benefits, such as improved inventory management, targeted marketing strategies, sustainable development, cultural preservation, economic opportunities, and better decision-making (Song & Li, 2008). It aids policymakers and Destination Management Organizations (DMOs) in crafting tourism policies, allocating resources, managing seasonal fluctuations, and devising strategies for the sustainable growth of tourism (Frechtling, 2001). Forecasting enables local communities to anticipate a rise in visitors and strategically allocate resources for infrastructure development (Dwyer, 2022). The benefits mentioned above are obtained by utilizing sophisticated methodologies and pertinent data sources to produce accurate predictions.

Research on forecasting of tourism demand has garnered considerable attention due to its critical role in facilitating efficient planning and decision-making for many stakeholders in the tourism industry (Goh and Law, 2002; Cho, 2003; Song et al., 2010; Untong and Khampukka, 2009; and Amaiquema and Illesca, 2017). Multiple studies have applied various forecasting techniques, such as econometric models (Song et al., 2010), time series models (Chu, 2009), and artificial intelligence-based approaches like neural networks (Uysal and El Roubi, 1999; Law, 2000). The availability and quality of data, the forecasting horizon, and the specific characteristics of the tourism destination all contribute to the selection of an appropriate forecasting method (Song and Li, 2008). In addition, researchers have emphasized the significance of incorporating external factors, including economic indicators, meteorological conditions, and political events, to enhance the precision of tourism demand forecasts (Frechtling, 2001).

Due to the significance of tourism demand forecasting and the fact that the precision of that forecast is contingent on the particulars of the data. Hence, the purpose of this research is to investigate how accurate predictions can be generated for Thailand's emerging tourist destinations. This sector remains underdeveloped in comparison to major tourist destinations, and there is a dearth of pertinent studies that could furnish data to aid in the administration of tourism in this area. Furthermore, the availability of data regarding provincial tourism demand remains extremely limited. This is primarily due to the data storage being transferred from the Tourism Authority of Thailand to the Ministry of Tourism and Sports around 2007 and the 2020 outbreak of the COVID-19 pandemic. Precise prognostications require forecasts that are accurate despite the limited availability of data. This study aims to enhance the existing knowledge by identifying the most appropriate model for predicting tourism demand for emerging and major destinations in Thailand. The objective is to get accurate forecasting models that are tailored to the individual characteristics of the place.

Research Objective

The objective of this research is to evaluate and identify suitable forecasting models for predicting tourism demand in emerging tourist destinations within the lower northern region of Thailand. This area includes Phitsanulok, Uttaradit, Sukhothai, Tak, and Phetchabun provinces. Given that tourism development in these provinces is relatively underdeveloped compared to major tourist cities, and the availability of data is limited, the study aims to find models that can accurately forecast tourism demand despite these constraints.

Scope of Research

This study employed four different approaches, including Holt-Winter's Exponential Smoothing, Box-Jenkins, Artificial Neural Network (ANN), and the Trigonometric Box-Cox ARMA Trend Seasonal Model (TBATS), to develop models for predicting tourism demand. The four methods correspond to the primary forecasting techniques, where ARIMA represents the Time series models, the Holt Winter model represents the deterministic models, ANN represents Artificial Intelligence-Based Approaches, and TBAT represents additional forecasting methods.

The study utilized monthly data on the number of tourists who stayed overnight in six tourist destinations. The Ministry of Tourism and Sports, Thailand provided the data, which includes the monthly tourism demand from January 2013 to December 2018, including a total of 72 observations. The study examined the effectiveness of the forecasting models by comparing the Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE). The chosen destination for this study is emerging tourist destinations located within the World Cultural Heritage Tourism Development Zone in the lower northern area of Thailand, including Phitsanulok, Phetchabun, Uttaradit, Sukhothai, and Tak. In order to analyze the impact of the tourist city area's background on the forecast, this study specifically chose Chiang Mai Province, the only main tourist destination in the North, as a point of comparison.

Literature Review

The forecasting of tourism demand increasingly relies on financial models that utilize time series data to predict tourist numbers. This method is gaining popularity due to its ease of data collection, as it primarily involves a single variable, and its analytical flexibility. In their review of tourism demand forecasting models from 2000 to 2007 (Song and Li, 2008) identified 121 studies, noting that over half (59.50%) employed time series data (time series models). Commonly used time series models in these studies include ARIMA, SARIMA, MARIMA, and GARCH.

The Artificial Neural Network (ANN) method was increasingly used for forecasting tourism demand (Song and Li, 2008). ANN, a computational technique that mimics the human brain's learning process, was effectively applied (Law and Au, 1999) in predicting the influx of Japanese tourists to Hong Kong, demonstrating higher accuracy compared to multiple regression and other models like naïve, moving average, and single exponential smoothing. Similarly, Cho evaluated the precision of forecasting foreign tourist demand in Hong Kong using ANN, exponential smoothing, and univariate ARIMA, finding that the ANN model was superior, particularly with data exhibiting unclear patterns (Cho, 2003). Jain, Varshney, and Joshi (2001) also confirmed

ANN's accuracy in short-term weekly water demand forecasting against linear and non-linear multiple regression models, as well as auto regressive models. However, contrasting findings by Supriatna et al. (2019) indicated that the Holt Winters exponential smoothing model outperformed the ANN model in predicting monthly foreign tourist numbers to Indonesia.

The forecasting of tourism demand has recently advanced with the utilization of more sophisticated time series data models. One such model, the Trigonometric Box-Cox ARMA Trend Seasonal Model (TBATS), has been increasingly applied in this context. Hassani et al. (2017) employed various time series data models, including TBATS method, to forecast tourism demand in several European countries like Austria, Cyprus, Germany, Greece, the Netherlands, Portugal, Spain, Sweden, and the United Kingdom. They found that the TBATS method outperformed other methods such as Exponential Smoothing (ETS), Neural Networks (NN), Fractionalized ARIMA (ARFIMA), Moving Average (MA), and Weighted Moving Average (WMA) in terms of accuracy. This superiority was particularly noticeable when dealing with data exhibiting multiple seasonal effects. Furthermore, Naim, Mahara, and Idrisi (2018) demonstrated that the TBATS model was more accurate than the Exponential Smoothing State Space Model with Box-Cox transformation, ARMA errors, Trend, and Seasonal components (BATS) in predicting daily natural gas consumption in India.

Time series data models can be effectively applied in other scenario related tourist demand, such as in predicting the tourist recovery after natural disasters. Huang and Min (2002) employed the Seasonal Autoregressive Integrated Moving Average (SARIMA) model to analyze the impact of earthquakes on tourist arrivals in Taiwan. Similarly, Kuo et al. (2008) utilized the ARMAX model to assess the effects of SARS and bird flu outbreaks on tourist numbers in China, Hong Kong, Singapore, and Taiwan.

However, the effectiveness of forecasting model does not always improve with its complexity, as it largely depends on the nature of the data. Research has indicated that simpler models can sometimes outperform more elaborate ones. For instance, Hassani et al. (2017) discovered that MA and WMA models, with their lower RMSE values, were more accurate than NN (Neural Network) and ARFIMA in predicting tourism demand in European countries. Similarly, a study by Chan, Hui, and Yuen (1999) on forecasting monthly tourist numbers of Singapore, where the data was relatively stable, found that the Naïve II (constant growth) models yielded more accurate predictions than more complex models like ARIMA, exponential smoothing, and trend curve analysis. Tu (1996) explored the advantages of forecasting with artificial neural network (ANN) models, uncovering that: i) Developing the ANN models requires comparatively less statistical training than other methods. Ii) ANN are accomplished models for identifying complex non-linear relationships between independent and dependent variables. Iii) ANN models are capable of detecting all possible interactions between variables used in prediction. iv) ANN can be trained on a variety of data using various algorithms.

Peng, Song, and Crouch (2014) conducted a meta-analysis of 262 studies from 1961 to 2011, uncovering that numerous factors influence the accuracy of tourism demand forecasting model. These factors include the origin and destination countries, forecast period, frequency of data, variables representing demand, statistical values for accuracy measurement, and sample size. Consequently, no single model consistently yields the most accurate predictions in every scenario. They advised that time series data models are suitable for monthly data and can provide accurate short-term predictions (less than a year) but may not be as effective if expenditures are used to represent demand. This finding is consistent with the conclusions of Song and Li (2008) and Hassani et al. (2017) that there is no single best model suitable for all types of data.

The new methods and pioneering new data sources for forecasting tourism demand continue to develop. Recent research by Jassim et al. (2023) explores the use of diverse internet data sources, including website traffic metrics (e.g., visitor statistics, session durations), social media engagement (e.g., Facebook likes, Twitter retweets), and Google Trends search query data, to enhance the accuracy of tourism demand forecasting models. Recently, the deep learning model has become a tool for estimating tourism demand (Law et al., 2019). For instance, the Convolutional Block Attention Module (CBAM). (Sun et al., 2023)

For Thailand, tourism demand forecasting using time series data is relatively scarce. For instance, Untong and Khampukka (2010) explored tourism demand in Thailand using ARIMA and SARIMA models for the period 1985-2008. Similarly, Saothayanun et al. (2014) compared forecasts of foreign tourist arrivals to Thailand using the Box-Jenkins method and the Winter method, utilizing data from January 1997 to September 2012.

Research Methodology

This study obtained monthly tourist demand data from The Ministry of tourist and Sports, Thailand in order to anticipate tourism demand using Holt-Winter's Exponential Smoothing, Box-Jenkins, ANN, and TBATS methodologies. The time frame spans from January 2013 to December 2018, with a total of 72 recorded instances for each province. Table 1 provides statistical descriptions of various datasets to emphasize their properties.

Table 1 Descriptive Statistics of Tourist Numbers in Each Province

Province	Obs	Mean	Std. dev.	Min	Max
Phitsanulok	72	143,669.4	18,300.4	107,754	196,368
Phetchabun	72	140,364.8	42,842.2	72,726	269,490
Tak	72	124,650.2	25,195.1	78,761	179,874
Sukhothai	72	65,113.9	15,787.8	39,565	105,232
Uttaradit	72	43,145.5	15,126.4	16,641	78,964
Chiang Mai	72	578,650.4	119,586.9	312,985	863,826

Source: Author's Study

The process of developing a predictive model is primarily divided into two stages: model training and model forecasting. Therefore, the 72 observations in this inquiry were divided into two sections. The initial phase involved the training of forecasting models using monthly data from January 2013 to December 2017 (60 observations). The remaining part, which consisted of twelve observations from January 2018 to December 2018, was utilized for model testing. The models' efficacy was evaluated by determining the discrepancy between the forecasted and actual tourist numbers for the same period via the forecasting models' errors. In order to assess the veracity of the model, this investigation implemented two methods: Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE). The following are the details of the forecasting techniques that were implemented in this study:

1. The Holt Winter's Exponential Smoothing Method which is well-suited for short-to medium-term time series data that exhibit trends and seasonal patterns. The model is:

$$S_t = \alpha \frac{y_t}{I_{t-L}} + (1-\alpha)(S_{t-1} + b_{t-1})$$

$$b_t = \gamma(S_t - S_{t-1}) + (1-\gamma)b_{t-1}$$

$$I_t = \beta \frac{y_t}{S_t} + (1-\beta)I_{t-L}$$

$$F_{t+m} = (S_t + mb_t)I_{t-L+m}$$

Where y represents the actual data, S denotes the smoothed data, b is the trend component, I stands for the seasonal index, and F is the forecasted value.

Exponential smoothing models are advantageous because they are well understood, explainable, do not require stationary data, and are suitable for small datasets. They also have relatively few hyperparameters, lower time complexity, and can handle missing data. Additionally, they can be extended to handle nonlinear dependencies more easily than models like ARIMA. However, they cannot handle covariates or multiple seasonality, are designed only for univariate time series, and are sensitive to recent outliers. Despite these limitations, their simplicity and lower computational demands make them useful in certain scenarios.

2. The Autoregressive Integrated Moving Average model (ARIMA), employing the Box-Jenkins approach, is a forecasting technique for time series data. It integrates three key components: Moving Average model order q or $\mathrm{MA(q)}$, the Autoregressive model order p or $\mathrm{AR(p)}$, and the Integral d or $\mathrm{I(d)}$ for data stabilization by subtracting the current data from its previous values (differencing). This model is frequently referred to as $\mathrm{ARIMA(p, d, q)}$, representing the respective orders of its autoregressive, integrated, and moving average components.

The AR model or $\mathrm{AR(p)}$ is: $$\mathrm{y_t = \beta_0 + \beta_1 y_{t-1} + \cdots + \beta_p y_{t-p} + \mu_t}$$

The MA model or $\mathrm{MA(q)}$ is: $$\mathrm{y_t = C_0 + C_1 \mu_t + C_2 \mu_{t-1} + \cdots + C_q \mu_{t-q}}$$

The process of forecasting using the ARIMA model can be summarized in the following steps:

(1) Data Analysis for Parameter Identification: Initially, analyze the data to determine suitable values for p (order of the autoregressive component), q (order of the moving average component), and d (degree of differencing required to make the series stationary). Then estimating the parameters of the chosen model.

(2) Model Verification: After selecting a model in step (1), verify its appropriateness. This involves examining the statistical significance of the estimated coefficients and analyzing the residuals of the model to ensure they behave like white noise. If the model does not meet these criteria, return to step (1) and re-estimate the parameters.

(3) Forecasting with the chosen $\mathrm{ARIMA(p, d, q)}$

The primary advantage of ARIMA forecasting lies in its reliance solely on time series data for model construction. Kenny et al. (1998) highlighted several key benefits of this approach. First, the ARIMA model is particularly effective when used to forecast a large number of time series. Second, i) ARIMA circumvents some of the typical problems associated with time series models, such as the complexities of dealing with multiple interrelated variables. Third, it avoids the complexities inherent in forecasting in multivariate models with lagged variables which necessitate initial prediction of lagged variables before they can be applied to forecast the primary variables in the model.

However, ARIMA models have several disadvantages. They cannot natively handle multiple seasonality, making it difficult to model data with both daily and yearly trends. ARIMA models struggle with sudden mean shifts and are designed primarily for univariate time series, limiting their use for interconnected series. They cannot model nonlinear dependencies and are sensitive to outliers, requiring preprocessing for datasets with many outliers. Specifying ARIMA parameters is subjective, relying on the modeler's skill, and the models can be computationally intensive to train. Additionally, ARIMA models require the data to be stationary or made stationary through differencing, limiting their applicability in some scenarios.

3. Artificial Neural Networks (ANN): ANN is a forecasting approach that emulates human brain functioning to predict outcomes based on available data. This study used a feed-forward neural network with a single hidden layer and lagged input for forecasting a uni-variate timeseries. It comprises of three layers: the input layer (number of tourists), the hidden layer, and the output layer (the prediction of tourist number), as illustrated in Figure. 6. The network is trained for one-step forecasting, then multi-step forecasts are computed recursively. It means that nodes in each layer take in inputs from the layer before them, and they then use those inputs in the next layer. Each node receives a weighted linear sum as input, which is changed via non-linear functions and then passed on to the following layer.

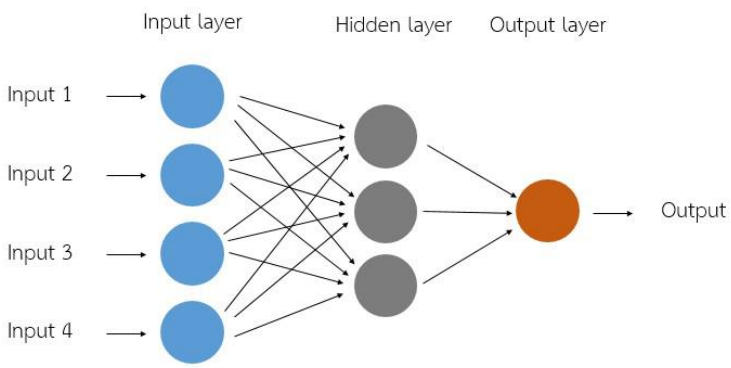

Figure 1 Artificial Neural Networks (ANN)
Source: Author’s Study

ANNs offer several advantages, including strong performance on unstructured data, the ability to detect complex features interaction without explicit specification, robustness to outliers, and the capacity to natively handle multiclass outcomes with a single mode. However, ANNs require substantial training data, extensive hyperparameter tuning, and are computationally intensive. They typically cannot handle missing data without preprocessing, and they lack interpretable coefficients, making them more challenging to interpret compared to models like linear or logistic regression. Despite these challenges, their performance advantages in complex data scenarios often outweigh the drawbacks.

4. TBATS method (Exponential Smoothing State Space Model with Box-Cox Transformation, ARMA Errors, Trend, and Seasonal Components) is a method for forecasting time series data influenced by multiple seasonal factors. The acronym TBATS method encapsulates its primary components: Trigonometric seasonality, Box-Cox transformation, ARMA errors, Trend, and Seasonal components. The structure of the TBATS model is outlined as follows:

$$y_t^{(\lambda)} = l_{t-1} + \phi b_{t-1} + \sum_{i=1}^{T} s_{t-m_i}^{(i)} + d_t$$

$$l_t = l_{t-1} + \phi b_{t-1} + \alpha d_t$$

$$b_t = \phi b_{t-1} + \beta d_t$$

$$d_t = \sum_{i=1}^{p} \varphi d_{t-1} + \sum_{i=1}^{q} \theta_i e_{t-1} + e_t$$

Where $y_t^{(\lambda)}$ represents the data a moment t (Box – Cox transformed), $s_t^{(i)}$ denotes the component of the i^{th} seasonal effect, l_t is local the level, b_t is the trend, d_t represents the $ARMA(p,q)$ of the error, and e_t is the Gaussian error.

The seasonal effect is:

$$s_t^{(i)} = \sum_{j=1}^{(k_j)} s_{j,t}^{(i)}$$

$$s_{j,t}^{(i)} = s_{j,t-1}^{(i)} \cos(\omega_i) + \gamma_1^{(i)} d_t$$

$$s_{j,t}^{*(i)} = -s_{j,t-1}^{(i)} \sin(\omega_i) + s_{j,t-1}^{*(i)} \cos(\omega_i) + \gamma_2^{(i)} d_t$$

$$\omega_i = \frac{2\pi_j}{m_i}$$

Where T represents the number of seasonal effects, m_i is the duration of the i^{th} seasonal effect, and k_i denotes the number of harmonics for the i^{th} seasonal effect. λ refers to the Box–Cox Transformation. α, β are the smoothing parameters, while ϕ is the trend value. ω_i and θ_i are coefficient of the $ARMA(p,q)$. $\gamma_1^{(i)}$ and $\gamma_2^{(i)}$ pertain to the adjustment and smoothing of the i^{th} seasonal effect. De Livera et al. (2011) identified several key benefits of the TBATS model. Firstly, it accommodates a robust parameter space to enhance forecasting potential. Secondly, the model is adaptable to both nested and non-nested data with multiple seasonality effects. Thirdly, it effectively manages typical non-linearity in time series data. Fourthly, it accommodates correlations between error values, specifically the autocorrelation of residuals. Finally, it provides a straightforward yet efficient estimation procedure.

However, they have notable disadvantages, such as the difficulty of incorporating covariates, limited understanding among practitioners, and complexity in explanation. TBATS models also require significant computational resources due to numerous hyperparameters and configurations, are only suitable for univariate time series, and are sensitive to outliers and mean shifts. Additionally, common implementations struggle with handling missing values natively, adding to their complexity.

Results

In Table 2 Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE) for all models were presented. The study found that all four predictive models had comparable predictive power. However, when compared, the ARIMA model is the most accurate compared to other models in forecasting the number of tourists in the four provinces of Phitsanulok, Uttaradit, Sukhothai and Tak with the error rates of 1.78%, 2.98%, 5.64%, and 2.67%, respectively. On the other hand, compared to other models, the ANN model estimates the number of tourists in Phetchabun and Chiang Mai more accurately with error rate of 3.54% and 3.30% respectively). Forecast model performance in predicting tourist numbers in Phitsanulok, Uttaradit, Sukhothai, Tak, Phetchabun, and Chiang Mai is shown in Table 2 and Figure2 - Figure6.

Table 2 Forecast Model Performance in Predicting Tourist Numbers in Phitsanulok, Uttaradit, Sukhothai, Tak and Phetchabun Province, January-December 2018

Province	Forecasting Performance	Holt Winter	ARIMA	ANN	TBATS
Phitsanulok	RMSE	4,890.70	3,516.80*	5,348.20	4,846.90
	MAPE	2.38	1.78*	2.55	2.61
Uttaradit	RMSE	6,671.70	1,853.80*	4,793.30	7,680.90
	MAPE	4.01	2.98*	3.04	4.92
Sukhothai	RMSE	6,437.40	4,739.60*	6,580.80	7,937.20
	MAPE	7.30	5.64*	7.56	8.30
Tak	RMSE	10,352.80	4,012.79*	5,641.80	8,506.80
	MAPE	6.90	2.67*	3.54	5.35
Phetchabun	RMSE	29,054.50	21,577.60	6614.10*	40,270.20
	MAPE	16.45	14.25	3.54*	21.41
Chiang Mai	RMSE	56,868.90	29,242.40	23,624.70*	61,795.10
	MAPE	8.09	4.17	3.30*	8.96

Source: Author’s Study

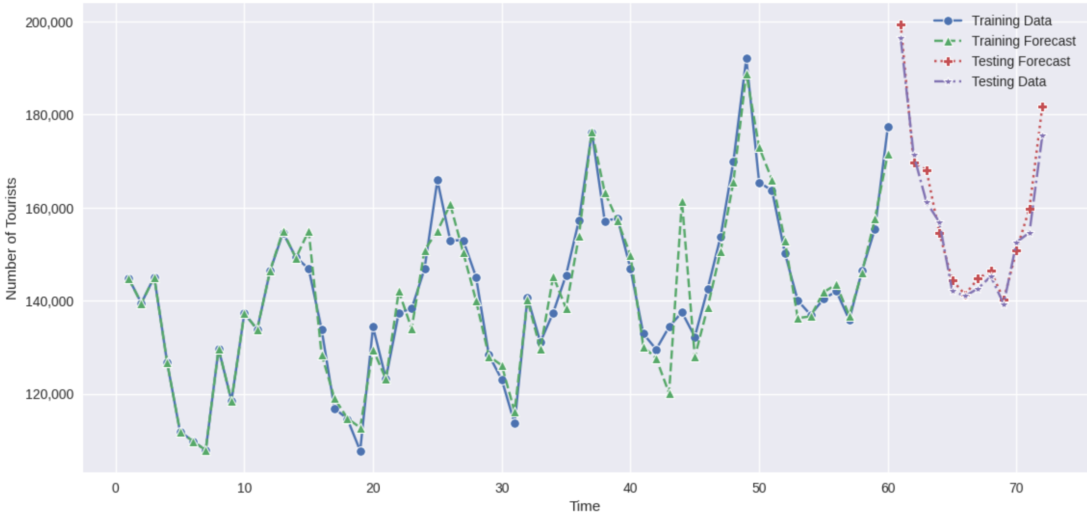

Figure 2 Forecasting Tourist Numbers in Phitsanulok by ARIMA Model

Source: Author's Study

Phitsanulok province experienced seasonality effect, with January having the greatest number of visitors (32,279.43 more tourists than typical). Then, the number of tourists progressively declined, and August was the month with the fewest tourists. ARIMA(0,1,0)(0,1,0) is the best reliable model for predicting Phitsanulok tourist number.

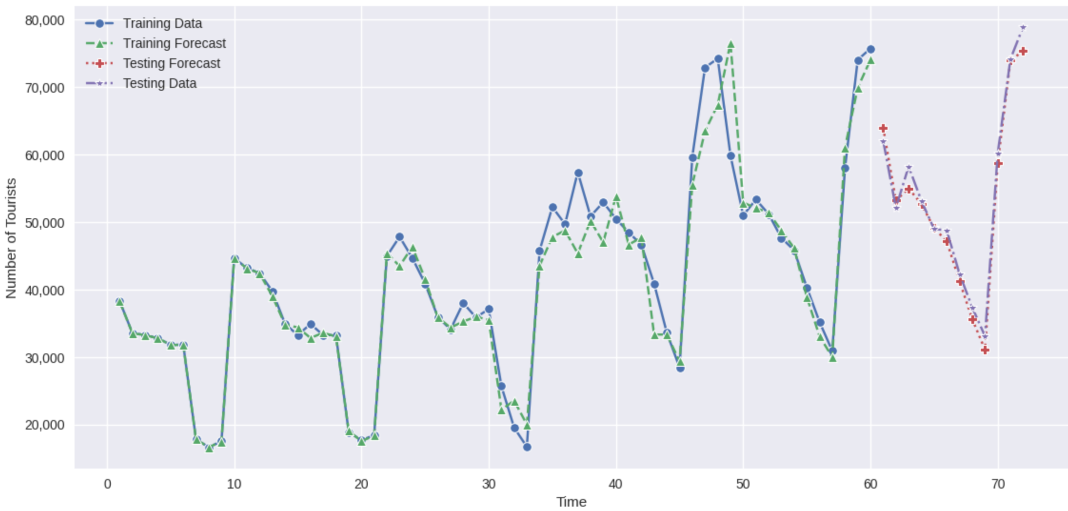

Figure 3 Forecasting Tourist Numbers in Uttaradit by ARIMA Model

Source: Author's Study

Uttaradit was a province where seasonality effect in tourism is significant. The number of tourists was very high in January and October, with more than usual, at 8,674.07 and 8,133.13 respectively. The most accurate model for forecasting tourists in Uttaradit province is ARIMA(1,0,0)(0,1,1).

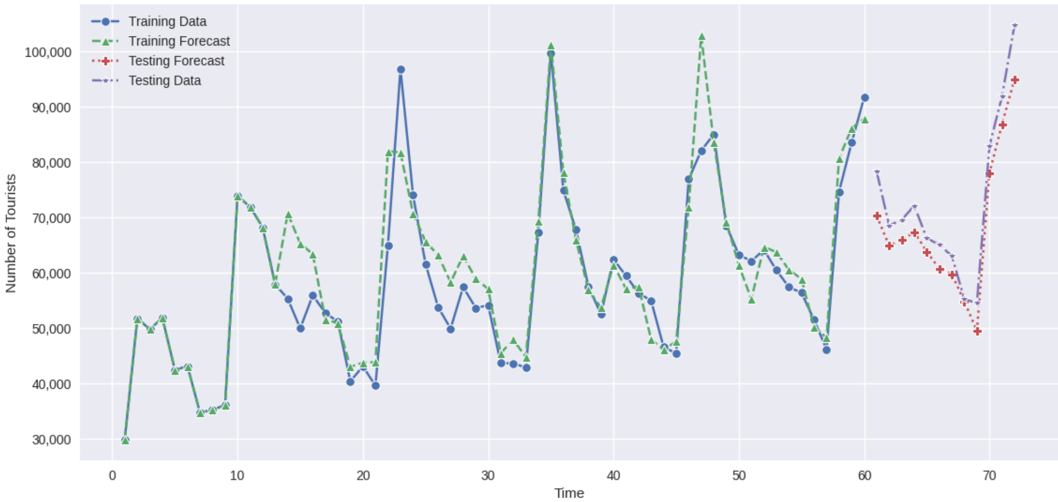

Figure 4 Forecasting Tourist Numbers in Sukhothai by ARIMA Model

Source: Author's Study

Sukhothai was a province where the number of tourists was exceptionally high in November and December, with 25,062.06 and 20,167.67 more tourists than usual, respectively, which was very different from the number of tourists in other months. When predicting Sukhothai tourists, ARIMA (1,1,1)(0,1,0) is the most accurate model.

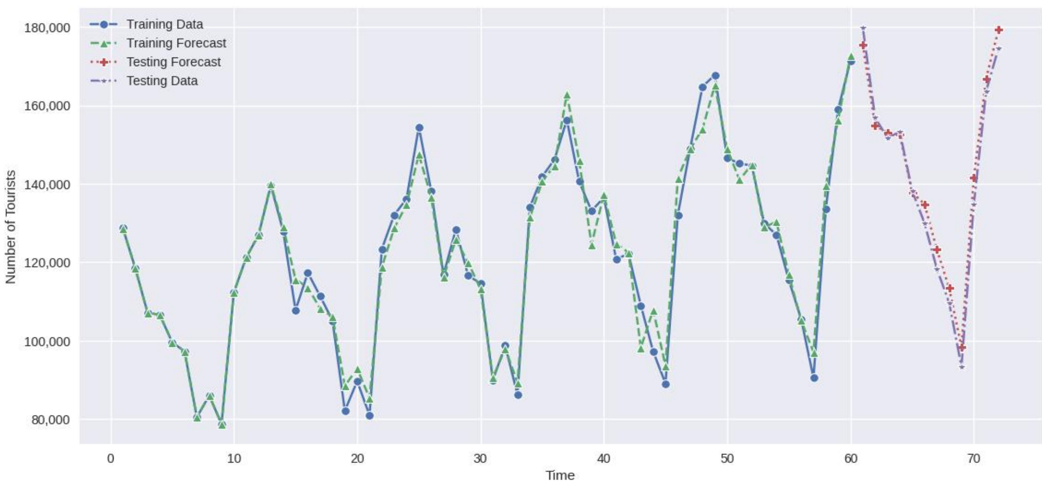

Figure 5 Forecasting Tourist Numbers in Tak by ARIMA Model

Source: Author's Study

Tak was a province experiences comparable seasonality effect every year. The highest month was January, when there were 33,640.57 more visitors than normal, and September was the month with the fewest visitors. Visitor numbers would progressively rise again in October. ARIMA(1,1,1)(0,1,0) is the best accurate model for predicting Tak visitors.

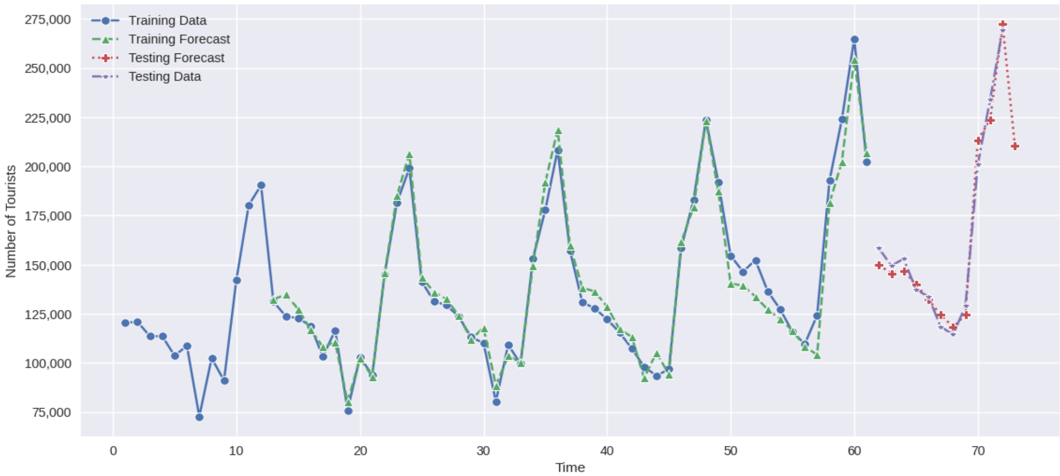

Figure 6 Forecasting Tourist Numbers in Phetchabun by ANN Model

Source: Author's Study

In Phetchabun province, the winter season, from October to January, is the high tourism season. December was the busiest month in Phetchabun, drawing 78,274.51 more visitors than average. The fitted model to forecasting tourist number in Phetchabun is NNAR *(*3,1,2)[12].

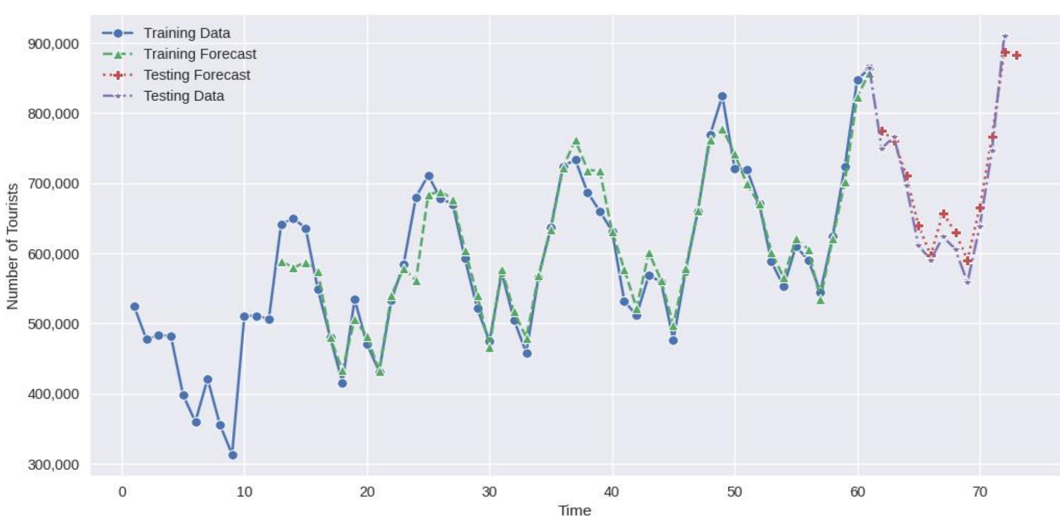

Figure 7 Forecasting Tourist Numbers in Chiang Mai by ANN model

Source: Author's Study

Chiang Mai is one of the most popular provinces for tourism that experiences distinct seasonal effects (in winter and early summer). Each year, the high tourist season occurs from October to April, peaking in January and February. The fitted model to forecasting tourist numbers in Chiang Mai province is NNAR(1,1,2)[12].

In this study, it was found that most fitted models were based on the ARIMA method, according to a comparison of prediction performance across several techniques. However, the outcome demonstrates that the ANN forecasting model is the only forecasting model that had no deviations exceeding 10% across the board. The results indicate that the ANN technique is a more flexible forecasting method than other methods. In other words, the error from the ANN analysis does not increase significantly in comparison to other methods, even if the data pattern is not very applicable. Furthermore, it was discovered that despite the scarcity of available data, the efficacy of forecasting remains satisfactory. Considering the forecasting performance of the models developed in this study, the majority of fitted models for predicting tourism demand in the lower northern region provinces of Thailand had a MAPE less than 5%, an acceptably accurate level (Swanson, 2015). In addition, the results in Table 2 and Figure 2-6 showed no empirical evidence of overfitting.

Discussion

Forecasting is necessary in all businesses to create relevant plans and make informed decisions. Hence, precise estimations of tourism demand are crucial for government agencies, tourism managers, and hospitality experts. The study's findings demonstrated that the ARIMA method's model had superior accuracy in predicting tourism demand in the emerging tourist destinations in the lower north region of Thailand, specifically in Phitsanulok, Uttaradit, Sukhothai, and Tak Provinces. Only Phetchabun Province and Chiang Mai, the main tourist destination in northern Thailand, have been proven to be more accurate by the ANN method's model.

Currently, the use of artificial neural networks (ANNs) for forecasting is becoming popular. However, the findings of this study align with the broader consensus in the field: there is no universally superior forecasting method. The efficacy of each model depends on the specific characteristics and fluctuations of the data (Song and Li, 2008; Peng, Song, and Crouch, 2014; and Hassani et al., 2017). The result in this study can be attributed to three primary factors: the quantity of data, the seasonal pattern and the quantity of tourists.

Due to the change in the governing body responsible for collecting tourism data and the impact of the COVID-19 epidemic, the availability of provincial-level tourist data in the same series has been restricted. Although the number of observations did not pose a problem in creating models that could accurately predict, the MAPE of the forecast is less than *5% (*Swanson, *2015*, p. 3*)*, ANN and TBAT require a large amount of training data to perform effectively. Notably, the Holt Winter model did not demonstrate greater accuracy in any province, potentially due to its deterministic nature, which sets it apart from other models in the study.

The second aspect pertains to the matter of seasonal trends. The ARIMA model's exceptional accuracy in four provinces may be attributed to the evident and constant seasonal patterns observed in their data. These findings align with the research undertaken by Song and Li *(*2008*)*, which indicated that 59.5% of the 121 studies completed between 2000 and 2007 utilized time series models such as ARIMA to anticipate tourism demand. The unique characteristics of demand-side data in tourism pose a challenge for accurate forecasting. While TBAT is effective in predicting multiple seasonal variations, ANN, which is suitable for data with nonlinear relationships, cannot be utilized to its full potential. However, in terms of efficiency in predicting tourist numbers across different provinces, ANN models generally were not the most efficient. However, the deviation in the ANN model's predictions was relatively minor when compared to other models. Particularly in provinces where tourists exhibit regular seasonal travel patterns, such as in Tak, Phetchabun, Uttaradit, and Phitsanulok Provinces, the Mean Absolute Percentage Error (MAPE) from ANN-based forecasts ranged between 2.545 and 3.542. This implies that if there is a circumstance where the seasonal pattern is unpredictable, particularly in the final year of forecasting, ANN will outperform other methods.

The final concern pertains to the quantity of tourists. The number of tourists in emerging tourist destinations is considerably smaller compared to that in the main tourist destinations. Hence, when tourism is encouraged by government initiatives or other irregular activities, it will significantly impact the fluctuation in tourist numbers during that specific timeframe. The study indicates that ARIMA may yield superior results in predicting such instances when used for short-term forecasting.

This research underscores the importance of selecting an appropriate forecasting model that aligns with the unique characteristics of the data and the specific context in which it will be applied. It emphasizes how the effectiveness of different forecasting methods can vary significantly based on factors such as data patterns, seasonal trends, and the volume of tourists. The findings illustrate that models like ARIMA, which are well-suited for data with clear and consistent seasonal patterns, can offer superior accuracy in certain regions.

Suggestions

The COVID-19 pandemic has brought unprecedented challenges to the tourism industry, necessitating the development of new forecasting models that can accurately predict tourism demand in the post-pandemic period. Recent studies have employed various forecasting techniques to predict post-COVID-19 tourist numbers. Mustafa et al. (2021) used ARIMA and Holt's Linear Models to forecast tourist arrivals based on GDP growth. Similarly, Ansarinasab and Saghaian (2023) analyzed the impact of different types of tourism on the spread of COVID-19 using quantile regression. These studies highlight how tourism dynamics have changed post-pandemic, with recovery varying across regions due to shifts in tourist behavior, local producers, and government policies. As a result, tourist data since 2020 has been anomalous compared to pre-pandemic data. Developing forecasting models to predict post-pandemic tourist numbers is crucial for both academic purposes and practical applications. The future research as well as the government plan for tourism development therefore should incorporate factors such as tourist's behavioral changes, regional recovery differences and data anomalies.

References

Amaiquema, J. P., and Illesca L. A. (2017). Forecasting International Tourism Demand For Ecuador. *International Journal of Research - Granthaalayah*, 5(8), 11-17.

Ansarinasab, M., and Saghaian, S. (2023). Outbound, Inbound and Domestic Tourism in the Post-COVID-19 Era in OECD Countries. *Sustainability*, 15(12), 9412: 1-29. Retrieved from https://doi.org/10.3390/su15129412

Chan, Y. M., Hui, T. K. and Yuen, E. (1999). Modeling the impact of sudden environmental changes on visitor arrival forecasts: The case of the Gulf War. *Journal of Travel Research*, 37(4), 391-394.

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

Chu, F.-L. (2009). Forecasting tourism demand with ARMA-based methods. *Tourism Management,* 30(5), 740–751. Retrieved from https://doi.org/10.1016/j.tourman.2008.10.016

De Livera, A. M., Hyndman, R. J., and Snyder, R. D. (2011). Forecasting time series with complex seasonal patterns using exponential smoothing. *Journal of the American Statistical Association*, 106(496), 1513-1527.

Dwyer, L. (2022). Tourism Economics and Policy. In D. Buhalis (Ed.), *Encyclopedia of Tourism Management and Marketing* (pp. 421–424). Edward Elgar Publishing. Retrieved from https://doi.org/10.4337/9781800377486.tourism.economics

Frechtling, D. (2001). *Forecasting Tourism Demand* (1st ed.). Routledge. Retrieved from https://doi.org/10.4324/ 9780080494968

Goh, C., and Law, R. (2002). Modeling and forecasting tourism demand for arrivals with stochastic nonstationary seasonality and intervention. *Tourism Management*, 23(5), 499-510.

Gokovali, U., and Bahar, O. (2006). Contribution of Tourism to Economic Growth: A Panel Data Approach. *Anatolia,* 17(2), 155–167. Retrieved from https://doi.org/10.1080/13032917.2006.9687184

Hassani, H., Silva, E. S., Antonakakis, N., Filis, G. and Gupta, R. (2017). Forecasting accuracy evaluation of tourist arrivals. *Annals of Tourism Research*, 63, 112-127.

Huang, J.-H., & Min, J. C. H. (2002). Earthquake devastation and recovery in tourism: the Taiwan case. *Tourism Management*, 23(2), 145–154. Retrieved from https://doi.org/https://doi.org/10.1016/S0261-5177(01)00051-6

Jain, A., Varshney, A. K. and Joshi, U. C. (2001). Short-term water demand forecast modelling at IIT Kanpur using artificial neural networks. *Water resources management,* 15(5), 299-321.

Jassim, R. S. Al, Jetly, K., Abushakra, A., & Mansori, S. Al. (2023). A Review of the Methods and Techniques Used in Tourism Demand Forecasting. *EAI Endorsed Transactions on Creative Technologies,* 9(4). Retrieved from https://doi.org/10.4108/eetct.v9i31.2986

Kenny, G., Meyler, A., and Quinn, T. (1998). *Forecasting Irish inflation using ARIMA models* (No. 3/RT/98). Central Bank of Ireland, 1-46.

Kuo, H. I., Chen, C. C., Tseng, W. C., Ju, L. F. and Huang, B. W. (2008). Assessing impacts of SARS and Avian Flu on international tourism demand to Asia. *Tourism management,* 29(5), 917-928.

Law, R. (2000). Back-propagation learning in improving the accuracy of neural network-based tourism demand forecasting. *Tourism Management*, 21(4), 331–340. Retrieved from https://doi.org/10.1016/S0261-5177(99)00067-9

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

Law, R., Li, G., Fong, D. K. C., & Han, X. (2019). Tourism demand forecasting: A deep learning approach. *Annals of Tourism Research*, 75, 410-423. Retrieved from https://doi.org/10.1016/j.annals.2019.01.014

Mustafa, H., Ahmed, F., Zainol, W. W., and Mat Enh, A. (2021). Forecasting the Impact of Gross Domestic Product (GDP) on International Tourist Arrivals to Langkawi, Malaysia: A PostCOVID-19 Future. *Sustainability*, 13(23), 13372: 1-16. Retrieved from https://doi.org/10.3390/su132313372

Naim, I., Mahara, T. and Idrisi, A. R. (2018). Effective short-term forecasting for daily time series with complex seasonal patterns. Procedia computer science, 132(10), 1832-1841.

Office of the Permanent Secretary of Ministry of Tourism and Sports. (2021). *Tourism Satellite Accounts in Thailand–2021.*

Office of the Permanent Secretary of Ministry of Tourism and Sports. (2023). *National Tourism Development Plan* No. 3 (2023 - 2027) Resilience, Sustainability, Inclusive Growth.

Peng, B., Song, H. and Crouch, G. I. (2014). A meta-analysis of international tourism demand forecasting and implications for practice. *Tourism management,* 45, 181-193. Retrieved from https://doi.org/10.1016/j.tourman.2014.04.005

Saothayanun, L., Taweesakulvatchara, S., Kanjanasakda, Y. and Somrang, B. (2014). A Forecasting Methods for the Number of International Tourists in Thailand: Box-Jenkins Method and Winter's Method. *Science and Technology Journal,* 22(1), 89-98.

Song, H., and Li, G. (2008). Tourism demand modelling and forecasting—A review of recent research. *Tourism Management*, 29(2), 203–220. Retrieved from https://doi.org/10.1016/j.tourman.2007.07.016

Song, H., Li, G., Witt, S. F., and Fei, B. (2010). Tourism Demand Modelling and Forecasting: How Should Demand Be Measured? *Tourism Econoomics*, 16(1), 63-81.

Sun, H., Yang, Y., Chen, Y., Liu, X., & Wang, J. (2023). Tourism demand forecasting of multi-attractions with spatiotemporal grid: a convolutional block attention module model. *Information Technology and Tourism,* 25, 205-233. Retrieved from https://doi.org/10.1007/s40558-023-00247-y

Supriatna, A., Hertini, E., Saputra, J., Subartini, B. and Robbani, A. A. (2019). The Forecasting of Foreign Tourists Arrival in Indonesia based on the Supply Chain Management: an Application of Artificial Neural Network and Holt Winters Approaches. International *Journal of Supply Chain Management,* 8(3), 156-163.

Swanson, D. A. (2015). On the Relationship among Values of the same Summary Measure of Error when used across Multiple Characteristics at the same point in time: An Examination of MALPE and MAPE. *Review of Economics and Finance*, 5(1), 1-14.

Tu, J. V. (1996). Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. *Journal of Clinical Epidemiology*, 49(11), 1225-1231. doi:10.1016/S0895-4356(96)00002-9

Untong, A. and Khampukka, P. (2009). Forecasting the Number of International tourists in Thailand by using the SARIMA Intervention Model. *Journal of Science and Technology, Ubon Ratchathani University*, 11(1), 196-218.

Untong, A. and Khampukka, P. (2010). Time Series Data Characteristics Checking for Forecasting Tourism Demand in Thailand*. Humanity and Social Science Journal, Ubon Ratchathani University,* 1(1), 60-85.

Uysal, M., and El Roubi, M. S. (1999). Artificial Neural Networks versus Multiple Regression in Tourism Demand Analysis. *Journal of Travel Research*, 38(2), 111–118. Retrieved from https://doi.org/10.1177/004728759903800203