Forecasting of visitors arrived in Taiwan for tourism supply chain demand using big data

YI-HUI LIANG

Department of Information Management, I-Shou University No.1, Sec. 1, Syuecheng Rd., Dashu District, Kaohsiung City 84001, TAIWAN

Abstract - The fast development of Information and Communication Technology, generate, collect and operate a large amount of data, which is termed big data. The search queries in web search engines can be retrieved by visitors to obtain useful infor-mation for the selected next visiting destinations. Google Trends on Google search engine can evaluate and compare how many times users are searching for specific terms or topics. Otherwise, economic factors, covering income, the rela-tive prices, and relative exchange rate usually influence the international tourist demand. However, there are different conclusions in different settings. Accord-ingly, this work presents the ARIMAX model for modelling and forecasting numbers of international tourists visiting Taiwan from Japan for different purposes and provides an analysis of the effects of big data and economic factors. The results can contribute to the decision makers of the tourism industry in Tai-wan.

Key-Words: Big data, Forecasting, Taiwan, ARIMAX, Tourism demand.

Received: April 14, 2021. Revised: July 2, 2021. Accepted: July 23, 2021. Published: July 28, 2021.

1. Introduction

The development of tourism is inextricably connected to the undertaking of tourism enterprises and tourism goods supplied by them. The specificity of product develop-ment in the tourism industry needs the engagement and collaboration of many part-nerships among enterprises in the tourism supply chain.

Forecasting is done on the basis of planning. Accordingly, forecasting tourism sup-ply chain demand is fundamental for effective formulation and implementation of planning by government and enterprises for tourism supply chain and can lead to significant economic benefits [1].

Big data can be generated, collected and processed through Information and Communication Technology (ICT). Big data can be analyzed for wisdom that conduct to better decisions and strategic business advances. The search queries in google of a visitor can illustrate many things as well as next visiting destination. Google Trends is a Google web that evaluates the frequencies of search queries across diverse regions and languages over time. For, example, Yu et al. (2019) [2] use Google Trends for oil consumption forecasting.

Otherwise, economical factors, including income, the consumer price index, and exchange rate usually influence the international tourist demand [3][4][5].

The ARIMAX (autregressive integrated moving average cause effect) model is an augmentation of the ARIMA (autregressive integrated moving average), and is a causal time-series regression model that can consider the influences of both endoge-nous and exogenous variables, as well as impulse effects [6]. Tsui et al. (2014) [7] observed that accurate, reliable tourist demand forecasting must address the impacts of both domestic and foreign endogenous and exogenous variables, which can be analyzed using the ARIMAX models.

In summary, this work on Taiwan's industries focuses primarily on the ARIMAX model to model and forecast numbers of international tourists visiting Taiwan for different purposes and analyze of the effects of Google Trends and economic fac-tors. The empirical data comprises the monthly arrival data of international tourists visiting Taiwan from Tourism Bureau, Ministry of Transportation and Communication (MOTC). The proposed models can consider Google Trends, income, the relative consumer price index, and relative exchange rates. The findings of this study can contribute to management and policydecision problems for tourism supply chain in Taiwan.

2. Literature Overview

Most forecasting tourism demand researches used quantitative methods, including time series models, artificial intelligence approaches, big data analysis, and the causal econometric models. Time series models, such as numerous exponential smoothing models and the ARIMA, SARIMA, ARFIMA and ARAR models, can be used in tourism forecasting. Goh and Law (2002) [8] found that the SARIMA models outperformed other eight models for forecasting the tourism demand of Hong Kong. Chu (2008) [9] employed the ARFIMA models to forecast the tourism demand for Singapore. Chu (2009) [10] used ARFIMA, ARAR and SARIMA models, to forecast the tourism demand of nine Asian Pacific countries. Chang and Liao (2010) [11] employed the SARIMA model for forecasting the tourism demand of Taiwan.

Neural networks are the most used among artificial intelligence approaches. Neural networks have numerous neural networks approaches. Law and Au (1999) [12] used a supervised feed-forward neural networks for forecasting the tourism demand. Recently, Peng et al. (2014) [13] proposed two forecasting models using radial basis function and support vector regression (SVR) neural networks for the UK inbound tourism quarterly arrival data. Zhang et al. (2017) [14] proposed an approach that hybridizes SVR neural networks with bat algorithm (BA) for forecasting tourist volume.

Big data analysis mainly utilizes search engine and social media data for forecast-ing tourism demand. Li et al. (2018) [15] indicated that the tourism related big data can be divide into three types: generated by user data, device data, as well as transaction data. Lately, it become a hot topic for forecasting tourism demand (mainly hotel occupancy and tourist volume). Pan and Yang (2017) [16] utilized big data for forecast-ing destination weekly hotel occupancy. Liu et al. (2018) [17] used big data for forecast-ing tourist volume.

Econometric models base on the relationships between the tourism demand and its impacting factors. For, instance, Cho (2001) [18] adjusted the ARIMA model using the economic indicators for forecasting the tourism of Hong Kong. Song et al. (2010) [19] explored different measures for tourism demand forecasting.

For income, Sheldon (1993) [3] concluded that there are three major determinant var-iables for tourism demand modelling: real per capita income of tourist producing countries, exchange rate, and relative price. Recently, Martins et al. (2017) [5] indicated that the influence of macroeconomic determinants on world tourism demand cover exchange rate, the relative domestic prices, and the world GDP per capital. For consumer price index, the consumer price index is designed to track the price level of a consumer basket of goods and services payed by a typical consumer. Morley (1994) [20] suggested that the consumer price index can been viewed as an alternative for a tourism price index because non-fare tourism prices are highly correlated with the consumer price index. In theory, the relative consumer price index is more suita-ble than the consumer price index from the view of tourists, and it can be used to measure consumer price changes in the original countries compared to the destination country. Recently, Wu et al. (2017) [21] indicated that tourism prices at a destination are expected to affect tourism demand based on the relative consumer price index be-tween the destination and origin. However, Martins et al. (2017) [5] concluded that the relative consumer price index is relative to low and middle income countries for tourism demand.

For exchange rate, the exchange rate measures the amount of one currency traded for one unit of another currency. The exchange rate can be used to calculate the dollar cost of a suit from Japan relative to what it would cost in Taiwan. In theory, for the same reason as the relative consumer price index, the relative exchange rate is more suitable than the exchange rate from the perspective of international tourists. Uysal and Crompton (1984) [22] suggested that real per capita income, relative exchange rates, transportation costs, and promotional expenses are the critical factors affecting international tourist arrivals to Turkey. Pham et al. (2017) [23] shared this perspective. However, Sheldon (1993) [3] suggested that the exchange rate will influence tourism demand, not the relative exchange rate.

Baggio and Sainaghi (2016) [24] suggested that developing different methods to obtain the more accurate tourism demand forecasts is a future direction for tourism demand researches. However, Peng et al. (2014) [13] also mentioned that every forecasting meth-od has its own advantages for a particular problem, but none is superior in various occasions.

Recently, Dogru et al. (2017) [4] found that it is necessary to analyze country-specific coefficients in order to accurately account for determinants of tourism demand for different countries. Consequently, there is a requirement to further explore the effects of various factors contributing to international tourist demand.

3. Research Method

Japan has long been the major international tourist source market for Taiwan. This study select visitors arrived in Taiwan from Japan (the residence is Japan) for different purposes to validate the proposed model.

For Google Trends, this study utilizes the web search data which the keyword and the topic is Taiwan, and the categories is business & industrial and travel as the big data indicator.

For income, this study employs GNI per capital data of Japan. For the consumer price index and exchange rate, for Taiwan as a destination, it is assumed in this study that the relative consumer price index and relative exchange rate influence international tourists visiting Taiwan since international tourists attempt to compensate for traveling expenditures by imposing the relative consumer price index and relative exchange rate between the country of origin and Taiwan. Use the consumer price index of Japan versus Taiwan as the relative consumer price index, using the exchange rate of Japan and Taiwan versus America, and taking the exchange rate of Japan versus Taiwan as the relative exchange rate.

$$y_{t} = \mu + \sum_{i=1}^{2} \frac{\Theta_{i}(B)}{\Phi_{i}(B)} B^{l_{i}} x_{it} + \varepsilon_{t}$$
(1)
$$\varepsilon_{t} = \frac{\Theta(B)}{\Phi(B)} a_{t}$$
(2)

Where Φ and Θ are the seasonal AR (autoregressive) and MA (moving average) parts, respectively.

4. Results

The visitor data based on the number of tourists visiting Taiwan from Japan for business and pleasure purpose and was obtained from the open statistical data of Taiwan's Tourism Bureau, MOTC. Data for the monthly tourism demand in the period from January 2004 to Sep 2019 were obtained and divided into a training set and a test set. The twelve-years data from January 2004 to December 2015 were used as the training set, and the 3-years-and-nine-months data from January 2016 to Sep 2019 were used as the test set. The number of tourists visiting Taiwan from Japan for business purpose is shown in figures 1. The number of tourists visiting Taiwan for pleasure purpose is shown in figures 2.



TIGUIE 1 THE NUMBER OF TOURISTS VISITING TAIWAN FROM JAPAN FOR BUSINESS PURPOSE



THE NUMBER OF TOURISTS VISITING TAIWAN FROM JAPAN FOR PLEASURE PURPOSE

The analytical steps for using the ARIMAX model are expressed as follows:

First, the number of tourists for business purpose and the number of tourists for pleasure purpose were transformed by using the original data logarithms. The Phillips and Perron test was adopted to do a unit root test. The transformed time series were stationary.

Next, aimed at Google Trends and the economic variables, including income, the relative consumer price index and relative exchange rate, Google Trends, income, the relative consumer price index, and relative exchange rates were transformed by using the original data logarithms, the unit root test was conducted. The transformed series after the original data logarithms for Google Trends were stationary. The first difference for income, the relative consumer price index, and relative exchange rates was taken to make the transformed time series stationary.

Third, the ARIMAX model was used to analyze and forecast the international tourists from Japan. The AR and MA structure was first estimated, and the explainable variables were added into the model. The results are shown in Table 1. For the tourists for business purpose, the R-squared was 0.705147; the AIC value was -2.049717, and the SIC value was -1.842525. The seasonality had a seasonal cycle of 12 months, and Big data and Relative exchange rate influences were at the 0.05 significance level. For the tourists for pleasure purpose, the R-squared was 0.870288; the AIC value was -1.212114, and the SIC value was -1.004922. The seasonality had a seasonal cycle of 12 months, and Big data were at the 0.05 significance level.

Additional quantitative approach covering backpropagation neural network with genetic algorithm as a benchmark were adopted, and a comparison was made with the benchmark model utilizing Neural Networks the software. NeuroSolutions, for the analysis. The results are illustrated in Table 2. The superiority of the ARIMAX model over the benchmark model was confirmed with regard to predicting the number of the international tourists traveling from Japan to Taiwan for different purposes.

TABLE I

The results for the international tourism demand series from Japan					
	Ln(the number of	Ln(the number of			
	tourists for	tourists for			
	pleasure purpose)	business purpose)			
Constant	9.891464(2.8292	9.824422(15.414			
	48***)	60***)			
	ARMA structure	ARMA structure			
AR(1)	0.985009	0.959731			
	(48.07597***)	(28.71533***)			
SAR(1)	0.971055	0.999955			
	(34.96860***)	(1448.452***)			
MA(1)	-0.598661 (-	-0.585266 (-			
	8.361225***)	7.182346***)			
SMA(1)	-0.686525 (-	-0.991328 (-			
	6.019533***)	28.21628***)			
	Explainable	Explainable			
	Variable	Variable			
Big data		0.087539(2.5356			
Ln(Google		36**)			
Trends(the					
category is					
business &					
industrial))					
Big data	0.249334(2.2384				
Ln(Google	42**)				
Trends(the					
category is travel))					
Income Δ	0.322496	0.163178			
Ln(income)	(0.580669)	(0.932270)			

Relative consumer	0.010003	0.001673			
price index Δ	(0.169863)	(0.025832)			
Ln(relative					
consumer price					
index)					
Relative exchange	-0.313108 (-	-0.577470 (-			
rate Δ Ln(relative	0.711178)	2.092688**)			
exchange rate)					
R-squared	0.870288	0.705147			
Adjusted R-	0.861510	0.685195			
squared					
AIC	-1.212114	-2.049717			
SIC	-1.004922	-1.842525			
	the number of	the number of			
	tourists for	tourists for			
	pleasure purpose	business purpose			
RMSE	12142.67	2025.040			
MAD	9263.941	1586.667			
MAPE	0.07659405	0.08180387			
Note: t statistics are revealed in parentheses respectively * ** and ***					

Note: t statistics are revealed in parentheses., respectively. * , **, and *** indicate significance at the 0.1, 0.05, and 0.01 levels, respectively.

TABLE II
he regults for ADIMAV command to Neurol network

The results for ARIMAX compared to Neural networks					
The number of tourists for business purpose					
	RMSE	MAD	MAPE		
ARIMAX	2025.040	1586.667	0.08180387		
Neural	2670.624	2538.009	0.09534500		
networks					
The number of tourists for pleasure purpose					
	RMSE	MAD	MAPE		
ARIMAX	12142.67	9263.941	0.07659405		
Neural	23986.40	21735.727	0.12584906		
networks					

5. Conclusions and Implications

The Taiwanese government has implemented a series of plans to open up the in-bound tourism market and boost the tourism industry between Taiwan and Japan, in the hopes of bringing more tourists from other countries. International tourists from Japan accounted for 15.28% of visitors to Taiwan in July 2016. This work explores and forecasts the number of international tourists traveling to Taiwan from Japan using the ARIMAX model. The ARIMAX model is a rich source of information, such as trends, seasonality, economic variables relating to income, consumer price index and exchange rates, autoregressive terms, and moving average terms. The ARIMAX model can provide a basis for planning and implementing policy, in particular policy related to resource allocations, by investors and managers in tourism.

The determinants of tourism demand differ on the basis of the country of origin. It is necessary to secure the specific characteristics of the tourism market since such determinants can differ for various groups, and the same strategy or policy may not be applicable to all markets. The analytical results indicate that tourists from Japan are influenced by trends, seasonality, autoregressive terms, and moving average terms, and differences in consumer price index (Ln(relative consumer price index)). Taiwan's government can understand the characteristics of their international tourists from Japan, it will develop marketing strategies or policies for the tourism industry in Taiwan, such as promotion activities.

This work has three main limitations. First, this study only discusses Japan, yet findings from other countries could potentially also help in the development of inter-national tourism. Second, Baggio and Sainaghi (2016) observed that tourist destinations are complex networks of components, where nodes are organizations and the people, and links are various types of business or personal relationships, owing to their nonlinear characteristics. This work does not deeply explore the interrelationships between Japan and other countries. Third, this work adopts monthly tourism demand data and does not utilize tourism data for other frequencies. Fourth, this work only considers incomes, consumer price indexes and exchange rates and thus ignores other variables, such as IPI, the effect of consumer preferences and behavior, destination promotions, and the effect of specific events or oneoff impacts, including the global financial crisis and SARS disease. Future studies could further explore other and/or advanced approaches, AI the relationships among Japan and other countries, and other factors influencing tourism demand. Otherwise, future studies could also examine frequency issues by transforming the frequency to weekly or yearly data.

Acknowledgements

Thanks to the support of I-Shou University (ISU), Taiwan R.O.C. no. ISU-106-IUC-06.

References

- C. Wu and Q. Zhao, "An uncooperative ordering policy with timevarying price and learning curve for time-varying demand under trade credit," European Journal of Industrial Engineering, vol. 11, no. 3, pp. 380–402, 2017.
- [2] L. Yu, Y. Zhao, L. Tang and Z. Yang, "Online Big Data-driven Oil Consumption Forecasting with Google Trends," International Journal of Forecasting, vol. 35, no. 1, pp. 213–223, 2019.

- [3] P. J. Sheldon, "Forecasting Tourism: Expenditures versus Arrivals," Journal of Travel Research, vol. 32, no. 1, pp. 13–20, 1993.
- [4] T. Dogru, E. Sirakaya-Turk and G. I. Crouch, "Remodeling International Tourism Demand: Old Theory and New Evidence," Tourism Management, vol. 60, pp. 47–55, 2017.
- [5] L. F. Martins, Y. Gan and A. Ferreira-Lopes, "An Empirical Analysis of the Influence of Macroeconomic Determinants on World Tourism Demand," Tourism Management, vol. 61, pp. 248–260, 2017.
- [6] A. C. Harvey, "The Econometric Analysis of Time Series" MIT press, MA, 1990.
- [7] W. H. K. Tsui, H. Ozer Balli, A. Gilbey and H. Gow, "Forecasting of Hong Kong Airport's Passenger Throughput," Tourism Management, vol. 42, pp. 62–76, 2014.
- [8] C. Goh and R. Law, "Modeling and Forecasting Tourism Demand for Arrivals with Stochastic Nonstationary Seasonality and Intervention," Tourism Management, vol. 23, pp. 499–510, 2002.
- [9] F. L. Chu, "A Fractionally Integrated Autoregressive Moving Average Approach to Forecasting Tourism Demand," Tourism Management, vol. 29, pp. 79–88, 2008.
- [10]F. L. Chu, "Forecasting Tourism Demand with ARMA-based Methods," Tourism Management, vol. 30, pp. 740–751, 2009.
- [11]Y. W. Chang and M. Y. Liao, "A Seasonal ARIMA Model of Tourism Forecasting: The case of Taiwan," Asia Pacific journal of Tourism research, vol. 15, no. 2, pp. 215–221, 2010.
- [12] R. Law and N. Au, "A Neural Network Model to Forecast Japanese Demand for Travel to Hong Kong," Tourism Management, vol. 20, pp. 89–97, 2002.
- [13]B. Peng, H. Song and G. I. Crouch, "A Mata-analysis of International Tourism Demand Forecasting and Implications for Practice," Tourism Management, vol. 45, pp. 181–193, 2014.
- [14]B. Zhang, X. Huang, N. Li and R. Law, "A Novel Hybrid Model for Tourist Volume Forecasting Incorporating Search Engine Data,"Asia Pacific Journal of Tourism Research, vol. 22, no. 3, pp. 245–254, 2017.
- [15] J. Li, L. Xu, L. Tang, S. Wang and L. Li, "Big Data in Tourism Research: A Literature Review," Tourism Management, vol. 68, pp. 301–323, 2018.
- [16]B. Pan and Y. Yang, "Forecasting Destination Weekly Hotel Occupancy with Big Data," Journal of Travel Research, vol. 56, no. 7, pp. 957–970, 2017.
- [17] Y. Y. Liu, F. M. Tseng and Y. H. Tseng, "Big Data Analytics for Forecasting Tourism Destination Arrivals with the Applied Vector Autoregression Model," Technological Forecasting and Social Change, vol. 130, pp. 123–134, 2018.
- [18] V. Cho, "Tourism Forecasting and its Relationship with Leading Economic Indicators," Journal of Hospitality & Tourism Research, vol. 25, no. 4, pp. 399–420, 2010.
- [19] H. Song, G. Li, S. F. Witt and B. Fei, "Tourism demand modelling and forecasting: how should demand be measured?," Tourism Economics, vol. 16, no. 1, pp. 63–81, 2010.
- [20]C. L. Morley, "The Use of CPI for Tourism Prices in Demand Modelling," Tourism Management, vol. 15, pp. 342–346, 1994.
- [21]D. C. Wu, H. Song, S. Shen, "New developments in tourism and hotel demand modeling and forecasting," International Journal of Contemporary Hospitality Management, vol. 29, no. 1, pp. 507–529, 2017.
- [22]M. Uysal and J. L. Crompton, "Determinants of Demand for International Tourist Flows to Turkey," Tourism Management, vol. 5, pp. 288–297, 1984.
- [23] T. D. Pham, S. Nghiem and L. Dwyer, "The Determinants of Chinese Visitors to Australia: A Dynamic Demand Analysis," Tourism Management, vol. 63, pp. 268–276, 2017.
- [24] R. Baggio and R. Sainaghi, "Mapping Time Series into Networks as a Tool to Assess the Complex Dynamics of Tourism Systems," Tourism Management, vol. 54, pp. 23–33, 2016.

Creative Commons Attribution License 4.0 (Attribution 4.0 International, CC BY 4.0)

This article is published under the terms of the Creative Commons Attribution License 4.0 <u>https://creativecommons.org/licenses/by/4.0/deed.en_US</u>