Forecasting International Tourism Demand Using the Recurrent Neural Network Model with Genetic Algorithms and ARIMAX Model in Tourism Supply Chains

Yi-Hui Liang

Abstract—Forecasting is the basis of planning, and the key to tourism supply chain. The international tourists visiting Taiwan from Mainland China, Japan, and South Korea is the major international tourist source markets for Taiwan. Recurrent neural network model is a fairly new and promising neural network technology. Genetic algorithms can be optimized the neural network structure. The ARIMAX model is recently utilized time series model for tourism demand forecasting. Accordingly, this work presents recurrent neural network model to forecast numbers of international tourists to Taiwan from Mainland China, Japan, and South Korea to help the Taiwanese tourism industry. This work also compares the forecast accuracy of the ARIMAX model with that of the ARIMAX model. Empirical results demonstrate that the recurrent neural network model with genetic algorithms is better than the ARIMAX model for Mainland China and South Korea, and worse than for Japan. The findings of this study can contribute to management and policy-decision issues related to the tourism industry for Taiwan in tourism supply chains.

Index Terms—Recurrent neural network, genetic algorithm, arimax, tourism demand, forecasting.

I. INTRODUCTION

Supply chain management (SCM) in the manufacturing industry has attracted extensively research focus while SCM studies in the tourism industry are very limited. Tourism products and services can be perceived as a value-added chain of different components, constructing a tourism supply chain. Stakeholders in the tourism supply chain interact with each other to resolve their different business goals with different operating scopes [1].

Tourism is one of the major industries in Taiwan, contributes significantly to the economy of Taiwan. The tourism sector can raise the national income and employment level of a country. Forecasting is the basis of planning for tourism supply chain. Accordingly, forecasting international tourists is fundamental for effective formulation and implementation of tourism supply chain planning by governments and enterprises, and can carry significant economic benefits for Taiwan's tourism industry. Nevertheless, the perishable nature of tourism goods makes forecasting an important subject for future success [2].

Baggio and Sainaghi (2016) [3] also observed that tourism demand is characterized by high volatility and rapid changes in trends.


Conversely, the ARMA (autoregressive moving average) of past observations of international tourist arrivals to Taiwan is normally analyzed by Box-Jenkins models. However, other factors also influence the demand of international tourists [8]. The ARIMAX (autoregressive integrated moving average cause effect) model is an extension of the ARIMA (autoregressive integrated moving average) with explanatory variables (X), and is a casual time-series regression models that can consider the influence forces of endogenous and exogenous variables, and impulse effects [9]. Tsut et al. (2014) [9] observed that accurate and reliable tourist demand forecasting must address the impacts of domestic and foreign endogenous and exogenous variables, and these can be analyzed by the ARIMAX models. In this study, select relative consumer price index and relative exchange rate as explanatory variables.

Mainland China, Japan and South Korea are the top three major international tourist source markets for Taiwan. Nowadays, the Taiwanese tourism industry encounters great difficulties in a varying global tourist market. The government of Taiwan must concentrate its attention on international tourists from Mainland China, Japan and South Korea.

Accordingly, this work presents the recurrent neural network model with genetic algorithms to forecast numbers
of international tourists to Taiwan from Mainland China, Japan, and South Korea to help the Taiwanese tourism industry. This work also compares the forecast accuracy of the recurrent neural network model with genetic algorithms with that of the ARIMAX model for forecasting the tourists from Mainland China, Japan, and South Korea. The findings of this study can contribute to management and policy-decision issues related to the tourism industry for Taiwan.

II. 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) [10] found that the SARIMA models outperformed other eight models for forecasting the tourism demand of Hong Kong. Chu (2008) [11] employed the ARFIMA models to forecast the tourism demand for Singapore. Chu (2009) [12] used ARFIMA, ARAR and SARIMA models, to forecast the tourism demand of nine Asian Pacific countries. Chang and Liao (2010) [13] employed the SARIMA model for forecasting the tourism demand of Taiwan.


Among economic factors, 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) [22] suggested that the consumer price index can be viewed as an alternative for a tourism price index because non-fare tourism prices are highly correlated with the consumer price index. Theoretically, the relative consumer price index is more suitable 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) [23] indicated that tourism prices at a destination are expected to affect tourism demand based on the relative consumer price index between the destination and origin. However, Martins et al. (2017) [24] concluded that the relative consumer price index is only correlated to low and middle income countries for tourism demand.

Among economic factors, 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 other countries 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) [25] suggested that relative exchange rates would affect international tourist arrivals to Turkey. Pham et al. (2017) [26] shared this perspective. However, Sheldon (1993) [27] suggested that the exchange rate will influence tourism demand, not the relative exchange rate.

Baggio and Sainaghi (2016) [3] 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) [4] also mentioned that every forecasting method has its own advantages for a particular problem, but none is superior in various occasions. Recently, Dogru et al. (2017) [8] 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.

III. RESEARCH METHOD

This study employs the neural network model and with genetic algorithms to forecast future the international tourists from Mainland China, Japan, and South Korea. This study selects the recurrent neural network model among numerous neural network models. Recurrent neural networks have been an interesting and important part of neural network research during the 1990’s. They have already been applied to a wide variety of problems involving time sequences of events and ordered data such as characters in words. A recurrent neural network is a class of artificial neural network where connections between units form a directed cycle. This allows it to demonstrate dynamic temporal behavior. For a typical recurrent neural network model, the calculating processes during forward propagation are as following.

\[ s_j(t) = f(\sum_{i=1}^l x_i(t)v_{ji} + \sum_h s_h(t-1)u_{jh} + b_j), \]

where \( s_j(t) \) represent the input of hidden layers at time \( t \), \( x_i(t)v_{ji} \) represent the input of input layer at time \( t \), \( s_h(t-1)u_{jh} \) represent the input of hidden layers at time \( t - 1 \).

Gradient descent is utilized to change each weight in
proportion to the derivative of the error with respect to that weight to minimize total error during the recurrent neural network training stage.

The most common global optimization method for training the recurrent neural network model is genetic algorithms. Genetic Algorithm (GA) is an adaptive method and often to solve the optimization problem. GA comprises of evaluation, selection, crossover, and mutation stages.

This study adopts GA to optimize the recurrent neural network model structure. The chromosomes represent the number of neurons in the hidden layer, and the step size, additive, multiplicative and smoothing parameters for DBD learning rule of the neural network architecture in binary as strings of 0s and 1s. Randomly initialize the chromosomes population. Training cycles are nested within evolutionary cycles in populations. Through cycles, select the root of MSE as the fitness function to assess the fitness of networks to progress with evolution of network optimization [28].

The analytical steps of the recurrent neural network model with GA are expressed as following:

Step 1: The recurrent neural network model with GA
Step 2: Forecast the values.

Next, attempts to construct the ARIMAX model to analyze and forecast the international tourism demand. This study incorporates relative consumer price index and relative exchange rate variables into the ARIMAX model. The ARIMAX model is expressed as following.

\[ y_t = \mu + \sum_{i=1}^{2} \Theta_i(B)B_i x_t + \varepsilon_t \]

where \( \Phi \) and \( \Theta \) are seasonal AR (autoregressive) and MA (moving average) parts.

The analytical steps of the ARIMAX model are expressed as following:

Step 1: Check whether the series is stationary.
Step 2: Set up the economic variables.
Step 3: Built up the model.
Step 4: Check whether the residual series is stationary.
Step 5: Check whether the residual series is serial correlated.
Step 6: Forecast the values.

The forecasting performance of the different models is assessed using the mean absolute deviation (MAD), the root mean squared error (RMSE) and the mean absolute percent error (MAPE) measures. For above three measures, MAPE is better because it can standardize the error terms to ameliorate comparisons among variables with distinct scales.

IV. RESULTS

This study collected the tourism demand of the number of tourists visiting Taiwan from South Korea, Japan and Mainland China from the open statistical data of Taiwan’s Tourism Bureau, Ministry of Transportation and Communication (MOTC). This study collects the data of tourism demand in the period from January 2001 to December 2014 as the train set and employed the period from January 2015 to July 2016 as the test set. The above three series are shown in Fig. 1, 2 and 3.

For South Korea, this study finds that the explanation power of the ARIMAX model is 89.9617% using the period from January 2001 to July 2016, the international tourism demand series from South Korea is positively influenced by relative consumer price index and relative exchange price, and exists the volatility effect and a seasonal cycle of twelve
months. This study uses the period from January 2001 to December 2014 as the train set to construct the recurrent neural network model with GA and the ARIMAX model, and uses the period from January 2015 to July 2016 as the test set to compare the forecasting results. The results illustrate in Table I. From the results, the recurrent neural network model with GA is superior to the ARIMAX model. In terms of MAPE, MAPE is 33.94% lower in the recurrent neural network model with GA than in the ARIMAX model, 63.80% lower than in the single exponential smoothing approach, 2.04% lower than in the single exponential smoothing approach, 13.28% lower than in the Holt-winter multiplicative seasonal exponential smoothing approach, 1.57% lower than in the Holt-winter multiplicative seasonal exponential smoothing approach, and 3.54% lower than in the Holt-winter multiplicative seasonal exponential smoothing approach.

### Table I: The Results for the International Tourism Demand Series from South Korea

<table>
<thead>
<tr>
<th>Model</th>
<th>MAD</th>
<th>MAPE</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARIMAX</td>
<td>14232.72143</td>
<td>0.228131231</td>
<td>15969.08237</td>
</tr>
<tr>
<td>Recurrent networks with GA</td>
<td>9109.219367</td>
<td>0.150692858</td>
<td>11387.46194</td>
</tr>
</tbody>
</table>

For Japan, this study finds that the explanation power of the ARIMAX model is 57.0536% using the period from January 2001 to July 2016, the international tourism demand series from Japan is positively influenced by relative consumer price index and relative exchange price, and exists the volatility effect and a seasonal cycle of twelve months. This study uses the period from January 2001 to December 2014 as the train set to construct the recurrent neural network model with GA and the ARIMAX model, and uses the period from January 2015 to July 2016 as the test set to compare the forecasting results. The results illustrate in Table II. From the results, the ARIMAX model is superior to the recurrent neural network model with GA. In terms of MAPE, MAPE is 9.24% lower in the ARIMAX model than in the recurrent neural network model with GA, 5.62% lower than in the regression approach, 25.63% lower than in the single exponential smoothing approach, 31.22% lower than in the single exponential smoothing approach, 26.36% lower than in the Holt-winter exponential smoothing approach, 24.73% lower than in the Holt-winter multiplicative seasonal exponential smoothing approach, and 24.44% lower than in the Holt-winter multiplicative seasonal exponential smoothing approach.

### Table II: The Results for the International Tourism Demand Series from Japan

<table>
<thead>
<tr>
<th>Model</th>
<th>MAD</th>
<th>MAPE</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARIMAX</td>
<td>17348.03036</td>
<td>0.114809888</td>
<td>24257.72830</td>
</tr>
<tr>
<td>Recurrent networks with GA</td>
<td>17789.18010</td>
<td>0.126496778</td>
<td>22089.02815</td>
</tr>
</tbody>
</table>

For Mainland China, this study finds that the explanation power of the ARIMAX model is 93.0910% using the period from July 2008 to July 2016, the international tourism demand series from Mainland China is positively influenced by relative consumer price index and relative exchange price, and exists the volatility effect and a seasonal cycle of twelve months. This study uses the period from July 2008 to December 2014 as the train set to construct the recurrent neural network model with GA and the ARIMAX model, and uses the period from January 2015 to July 2016 as the test set to compare the forecasting results. The results illustrate in Table III. From the results, the recurrent neural network model with GA is superior to the ARIMAX model. In terms of MAPE, MAPE is 4.39% lower in the recurrent neural network model with GA than in the ARIMAX model, 13.92% lower than in the regression approach, 5.20% lower than in the single exponential smoothing approach, 2.04% lower than in the single exponential smoothing approach, 5.83% lower than in the Holt-winter multiplicative seasonal exponential smoothing approach, 3.88% lower than in the Holt-winter additive seasonal exponential smoothing approach, and 4.17% lower than in the Holt-winter multiplicative seasonal exponential smoothing approach.

### Table III: The Results for the International Tourism Demand Series from Mainland China

<table>
<thead>
<tr>
<th>Model</th>
<th>MAD</th>
<th>MAPE</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARIMAX</td>
<td>36099.41341</td>
<td>0.106238668</td>
<td>48044.58751</td>
</tr>
<tr>
<td>Recurrent networks with GA</td>
<td>33927.7358</td>
<td>0.101572756</td>
<td>42662.94421</td>
</tr>
</tbody>
</table>

V. Conclusions

Peng et al. (2014) [4] suggested that suitable forecasting approaches need to be selected for various forecasting occasions. The model is adopted to estimate the three major source markets of Mainland China, Japan, and South Korea, which together made up 59.48% of the total foreign arrivals to Taiwan in July 2016, on the basis of MOTC Taiwan, ROC.

The Taiwanese government has implement a series of plans to open up the inbound tourism market and boost the tourism industry development, and hope to bring more tourists. Forecasting the number of international tourists is both a demanding and important task for Taiwan.

This work explores and forecasts the numbers of international tourists to Taiwan from Mainland China, Japan, and South Korea using two models, of which the recurrent neural network model with GA is better than the ARIMAX model for Mainland China and South Korea, and worse than Japan.

The ARIMA model is usually considered as a benchmark to compare the forecasting accuracy with the proposed method by many researchers for tourism demand. The ARIMAX model bases the ARIMA model and adds the explanatory variables. In theory, the ARIMAX model has better explanation power than the ARIMA model. Therefore, this study proposes the recurrent neural network model with GA to compare with the ARIMAX model in forecasting tourism demand.

In this study, the neural network model with GA can generate forecasts without additional assumptions and limitations because of its nonparametric properties. However, the ARIMAX model can offer richer information, such as trends, seasonality, economic variables relating to consumer price indexes and exchange rates, autoregressive terms and moving average terms.

The recurrent neural network model with GA and the ARIMAX model are both more new approaches for tourism demand forecasting. The recurrent neural network model with GA belongs to artificial intelligence approach, and the ARIMAX model belongs to time series models and the causal
econometric models. From this study, we can learn that the recurrent neural network model with GA is not superior to the ARIMAX model for all situations. The results is consistent with the results of Peng et al. (2014). Different countries must have different setting for international tourism forecasting.

The results by the ARIMAX model and the neural network model with GA can provide a basis for planning and implementing policy, in particular in resource allocation, by investors and managers in tourism.

This work has five main limitations. First, this study only discusses Mainland China, Japan, and South Korea, yet findings from other countries can also help in the development of international tourism. Second, Baggio and Sainaghi (2016) [3] observed that tourist destinations are complex networks of components, where nodes are organizations and people, and links are various types of business or personal relationships, owing to nonlinear characteristics. This work does not deeply explore the interrelationships among Mainland China, Japan, and South Korea. Third, this work adopts monthly tourism demand data, does not utilizes other frequency tourism data. Fourth, this work only considers relative consumer price indexes and relative exchange rates, and thus ignores other variables, such as GDP or CPI, the effect of consumer preferences and behavior, destination promotion, and the effect of specific events or one-off impacts including the global financial crisis and SARS disease. Fifth, owing to the project reason, this study selected the data from January 2001 to July 2016. Other studies can select more data for further research.

Future studies could further explore other AI and/or advanced approaches, the relationship structure among South Korea, Japan, China and other countries, and other factors influencing tourism demand. Otherwise, future studies also can examine the frequency issues by transforming the frequency to weekly or yearly.

CONFLICT OF INTEREST

Please declare whether or not the submitted work was carried out with a conflict of interest. If yes, please state any personal, professional or financial relationships that could potentially be construed as a conflict of interest. If no, please add “The authors declare no conflict of interest”.

REFERENCES


Copyright © 2022 by the authors. This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited (CC BY 4.0).