A Hybrid Approach on Tourism Demand Forecasting

To cite this article: M. E Nor et al 2018 J. Phys.: Conf. Ser. 995 012034

View the article online for updates and enhancements.
A Hybrid Approach on Tourism Demand Forecasting

M. E Nor, A. I M Nurul and M. S Rusiman

Department of Statistics and Mathematics, University of Tun Hussein Onn Malaysia, 86400 Johor Malaysia

Abstract. Tourism has become one of the important industries that contributes to the country’s economy. Tourism demand forecasting gives valuable information to policy makers, decision makers and organizations related to tourism industry in order to make crucial decision and planning. However, it is challenging to produce an accurate forecast since economic data such as the tourism data is affected by social, economic and environmental factors. In this study, an equally-weighted hybrid method, which is a combination of Box-Jenkins and Artificial Neural Networks, was applied to forecast Malaysia’s tourism demand. The forecasting performance was assessed by taking the each individual method as a benchmark. The results showed that this hybrid approach outperformed the other two models.

1. Introduction

Tourism is one of the important sectors that influences the country’s economy. This is due to the considerable intake of money for businesses in terms of goods, services and the opportunities for employment in service industries such as hospitality administration and transportation [1]. There are three basic forms of tourism; and those are domestic tourism, inbound tourism and outbound tourism. Domestic tourism refers to activities of tourists within their country of residence and outside of their homes. Inbound tourism refers to the activities of tourists from outside of the country of residence. Meanwhile, outbound tourism refers to the activities of resident tourists outside of their country of residence. The increasing number of tourist arrivals by year [2] proved that Malaysia is a country of great potential as a tourism country. This is because it has its own charm since Malaysia has a tropical climate which is warm and humid throughout the year. The United Nations World Tourism Organisation (UNWTO) conference held in Kuala Lumpur in 2007, the Malaysia’s capital, commemorated the success of the country’s tourism performance; mainly in terms of tourist arrivals and receipts [3]. This statement is supported by the increase in estimation of tourist arrivals from 2000 to 2015; i.e. from 10.2 million to 25.7 million of tourists with RM 17.3 billion to RM 69.1 billion in revenue, respectively [2].

The increase in the number of tourist arrivals and revenue over the years had motivated the government to be more concerned and to actively promoting tourism sector which currently stands as Malaysia’s top second foreign exchange earner after manufacturing [4]. As a systematic and progressive developing country, the plans for not spending too much or too little money in providing services or goods to upcoming tourists is necessary. Therefore, forecasting tourism demand is essential to a country since it affects the country’s planning and decision making. Various methods can be used to forecast the
number of incoming tourists, but the accuracy of each method would be questioned later. Many researchers have used Box-Jenkins (BJ) and artificial neural network (ANN) in forecasting time series data [5], [6], [7] and [8]. Hence, a linear model, namely the Box-Jenkins (BJ), and nonlinear model, the artificial neural networks (ANN) have been chosen for this study.

2. Time series forecasting models

Box-Jenkins forecasting method belongs to the family of algebraic models known as ARIMA model, which has the ability to forecast based on a given stationary time series [9]. ARIMA method proposed by Box and Jenkins in 1970 has an autoregressive component. It uses past statistical data of the variables and also has a treatment of the random or stochastic factors through the uses of moving averages. The combination of both components enables the ARIMA method to incorporate seasonal analysis and examine interrelations deeper [7]. All stationary time series are modelled by ARIMA processes, as long as the appropriate order of p, the number of autoregressive (AR) terms, and q, the number of moving average (MA) terms, are appropriately specified [10]. ARIMA allows one to choose from a wide class of models rather than being restricted to one particular model [11]. ARIMA also could outperform large and complex econometric models in variety of situations [12, 13].

Meanwhile, ANN is a mathematical model used in many business applications for pattern recognition, forecasting, prediction and classification. This is due to its ability to “learn” from the data, its nonparametric nature and the ability to generalize [14]. The second method used to forecast tourism demand in this study is by using artificial neural network. This method is a non-linear model that mimics human brain’s function, with the complex system of biological neurons model. ANNs demonstrate the capacity of improving time series forecasting through the analysis of additional information, reducing its size and lessening its complexity [15]. ANN is also capable of applying learning process on sample data besides solving complicated and nonlinear forecast on tourist arrivals [8]. As other methods faced difficulties, such as time-consuming and expensive, this intelligent ANN method handles the challenges well as it does not learn from past data. Other than that, [16] and [17]’s research on forecasting tourist arrivals in Hong Kong and South Africa had also concluded that neural network performs better than the other time series forecasting methods.

Since tourism industry’s data often fluctuate and are complex in nature, it is difficult for linear model such ARIMA to capture the non-stationary property and moving tendency; thus, causing inaccurate forecasting [18]. However, a study cannot easily assumed that the time series data follow nonlinear pattern and does a forecasting using nonlinear model such as ANN without further analysis [19]. Thus, many hybrid methods are being proposed and applied in the literature, which combined two or more individual models; hence, providing an additional advantage over single method. The novel forecasting model based on empirical mode decomposition (EMD) and neural network had outperformed the single back-propagation neural network (BPN) model without EMD pre-processing and the traditional ARIMA [20]. Meanwhile [21] had used hybrid method of between vector error correction model (VECM) and multi-output support vector regression (MSVR) in capturing the linear and non-linear patterns exhibited in agricultural commodity future prices; and the results were promising. Meanwhile, a study on tourism services in Australia was improved by using a hybrid semantic enhanced approach by combining the new Ontology-based Semantic Similarity (IOBSS) measure and the standard item-based Collaborative Filtering approach [22]. A hybrid evolutionary system composed by a simple exponential smoothing filter, ARIMA and autoregressive (AR) linear
model as well as support vector regression (SVR) model had shown promising results in the forecasting domain [23].

Commonly, if hybrid models are being used, the accuracy of forecasting will be improved. However, if a hybrid uses many decompositions and models, the accuracy will degrade after some limitations; and the model will no longer be successful. As such, hybrid models should contain a limited number of individual models as to retain the easiness of the model as well as to achieve accuracy [24]. In this paper, a Hybrid Equally-Weighted BJ-ANN forecasting model was applied. This method was invented when Franklin, its pioneer, realized that people and organizations commonly make decisions by combining information from various inputs [25]. The effectiveness of Equally-Weighted forecast or simple forecasting combination method was tested by [26] and [27]; and it was proven in improving the forecasting performance as the number of combined single methods increases. Graefe [25] had used equal-weighted forecast in providing new evidence for U.S. presidential election forecasting; a field dominated by multiple regression analysis’s application. In his results, equal-weighted appeared to perform well when predicting new data with the addition of all relevant variables.

3. Materials and forecasting models

Details about BJ and ANN described as in previous chapter. In this section, a more complete discussion of these modelling presented.

3.1. Study area and data

The data set is number of tourist arrivals in Malaysia that covers from 1998 to 2016. These monthly data were obtained from Department of Malaysia Tourism. The 18-year (1998-2015) of tourist arrival data were used to attain the best model fit for all models. Meanwhile, the remaining year’s (2016) records were set aside for model validation and comparisons for forecasting purposes.

3.2. Box-Jenkins (BJ)

An ARIMA model contains autoregressive (AR), moving average model (MA) and is extended as it includes differencing. The ARIMA model is ARIMA \((p,d,q)\); where \(p\) is the order of the AR, \(d\) is the number of times differencing being carried out and \(q\) is the order of the MA [5]. The general ARIMA model, which allocates seasonality is written as equation 1.

\[ X_t = \varnothing_0 y_{t-1} + \varnothing_2 y_{t-2} + \ldots + \varnothing_p y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \ldots - \theta_q \varepsilon_{t-q} \]

or define as:

\[ \varnothing(B)\Phi(p)(B^d) = \theta(B)\Theta(q)(B^s)\alpha_t \]

where \( X_t \) and \( \varepsilon_t \) are the actual value and random error at time period \( t \), respectively; \( \varnothing_i \) \((i=1,2,\ldots,p)\) and \( \theta_j \) \((j=0,1,2,\ldots,q)\) are model parameters, \( p \) and \( q \) are integers and often referred to as orders of the model. Random errors \( \varepsilon_t \) assumed independently and identically distributed with a mean of zero and a constant variance of \( \sigma^2 \). The process \{\( Y_t \)\} said as autoregressive integrated moving average process, ARIMA \((p,d,q)\), if \( X_t = \nabla^d Y_t \) is an ARMA \((p,q)\) process.
3.3. Artificial Neural Networks (ANN)

Neural networks include several models, such as MLPs, which are useful for statistical applications [28]. The major steps in designing the data forecasting model are choosing variables, data collection, data pre-processing, dividing the data set into smaller sets (training, test and verification), determining network’s topology (number of hidden layers, number of neurons in each layer, number of neurons in output layer and the transformation function), calculating the forecast value, and lastly calculating the forecast error. The number of input nodes, p and the number of hidden nodes, q are often selected by doing some experiments. In this paper, p and q are enumerated from the BJ method. Once a network structure (p,q) specified, the network will be ready for training [5]. The Levenberg-Marquardt (LM) training algorithm is used to find the weights that minimize some overall error measures such as the sum of squared errors (SSE) or mean squared errors (MSE). LM reduces the number of epoch needed, reduces fitting error and downsizes the training set [29].

For a time series forecasting’s problem, a training pattern consists of a fixed number of lagged observations of the series. The lag(s) is(are) determined based on the autoregressive order in ARIMA model [30], [31] and [32]. Each input connects to all neurons, and the neurons are connected to the output. Arrows indicate that arrow’s source is function’s argument computed at the destination of the arrow and each arrow has an estimated corresponding weight or parameter. There are constants or bias connected to each neuron and output which denoted as 0. To illustrate this, an example of NNs architecture with two inputs is presented in Figure 1.

![Neural Networks architecture with two inputs and two neurons](image)

**Figure 1.** Neural Networks architecture with two inputs and two neurons
Each neuron is a processing unit that uses a logistic function to calculate the linear combination of inputs. At each \( i \)-th neuron, the linear combination of \( k \) inputs is calculated as follows [33]:

\[
y_i = \frac{1}{1 + e^{-x_i}}
\]

where

\[
y = b_i + \sum_{j=1}^{k} w_{i,j} x_j.
\]  \( (2) \)

Here \( b_i \) are bias, \( x_j \) are independent variables or inputs and \( w_{i,j} \) are weights from input \( x_j \) to \( i \)-th neuron in the hidden layer. The \( y \) is the total of all the input elements and weight that enters the \( i \)-th neuron. The number of neurons is denoted by \( h \) and the number of neurons used is one up to five neurons, \( h = 1, 2, \ldots, 5 \). Finally, the predicted values obtained by using the linear combination of input that is given by equation 3.

\[
\hat{y}_{i+h} = \sum_{i=1}^{k} \gamma_i n_i + \gamma_0,
\]  \( (3) \)

where \( \gamma_0 \) are bias for output and \( \gamma_i \) are weights from \( n_i \) to output. After software has forecast the tourism and residual data by using artificial neural network method, the data need to be transformed back to its original scale for first differencing of non-seasonal and seasonal data. The formula is given as follows:

\[
F(t) = Fd(t) + Y_{t-1} + Y_{t-12} - Y_{t-13}
\]  \( (4) \)

### 3.4. Hybrid Equally-Weighted BJ-ANN model

A specific value of weight assigned to each method. In Equally-Weight hybrid forecast method, the combination weights are dispensing equally to each of the forecasts and total weight would add up to one. The formula of forecast combination and weight assigned to each forecast [34] were given by:

\[
f_c = \sum_{i=1}^{n} w_i f_i
\]

\[
w_i = \frac{1}{n}
\]  \( (5) \)
where $f_i$ was the $i$th single forecast, $f_c$ was the combined forecast generated by the $n$ single forecast $f_i$ and $w_i$ was the combination weight assigned to $f_i$. The value of $n$ is two since there are two methods to be hybrid. Thus half of forecast value from ARIMA model would be summed up with forecast value from ANN. Figure 2 show simple process of Equally-Weighted hybrid forecast between ARIMA and ANN models.

$$f_c = \sum_{i=1}^{n} w_i f_i$$

**Figure 2. Flowchart of Equally-Weighted hybrid ARIMA-ANN**

3.5. Forecast Accuracy Measurement

In order to measure performances of forecasting time series methods between autoregressive integrated moving average (ARIMA), artificial neural network (ANN) and hybrid of ARIMA-ANN method; error measurement root mean square error (RMSE), mean absolute deviation (MAD) and mean absolute percentage error (MAPE) were estimated. MSE is the average of the squared individual errors that measures dispersion of forecast errors. The smaller the MSE value, the more stable the model is [35]. However, MAPE is a better measurement than MSE because it does not accentuate large errors. The final accuracy measurement is the MAD. It is the average of the absolute value of the error, regardless whether the error is being overestimated or underestimated. Among these three accuracy measurements, MAPE provided the most accurate and fair comparison of forecasting methods [35].

4. Result and findings

Study used three methods to forecast monthly tourism demand in 2016; and those are Box-Jenkins (BJ), Artificial Neural Networks (ANN) and Hybrid Equally-Weighted of BJ and ANN. The study required data of tourist arrivals from January 1998 until December 2016; plotted as in Figure 3. Data from January 1998 until December 2015 are in-sample data, $y_1$, used for model fitting. Meanwhile data from January 2016 until December 2016 are out-sample data, $y_2$, used for model validation. Small deviation of out-sample data with forecast value shall determine which model shows better performance.
4.1. Box-Jenkins (BJ)

The ACF and PACF plots indicated that the mean is not stationary. Non-seasonal and seasonal differencing had been conducted in order to ensure the mean is stationary. Several tentative models were identified and their in-sample MSE values are shown in Table 1.

**Table 1.** The mean squared error value for all tentative models

<table>
<thead>
<tr>
<th>Model Parameter</th>
<th>Mean Squared Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>SARIMA (1,1,1)(0,1,1)_{12}</td>
<td>17236963394</td>
</tr>
<tr>
<td>SARIMA (2,1,4)(0,1,1)_{12}</td>
<td>16425550460</td>
</tr>
<tr>
<td>SARIMA (1,1,1)(0,1,2)_{12}</td>
<td>16724164692</td>
</tr>
</tbody>
</table>

SARIMA with the smallest MSE value was chosen and it will be used to determine the input lag in ANN model. The model is expanded as follows:
\[ \varnothing_r(B)\Phi_r(B)^{\prime}v^{\prime}\gamma y_t = \theta_q(B)\Theta_q(B)^{\prime} \alpha, \]
\[ \varnothing_s(B)v^{\prime}\gamma y_t = \theta_s(B)\Theta_s(B)^{\prime} \alpha, \]
\[ (I - \varnothing_j(B) - \Phi_j(B)^{\prime})(I - B - B^{12} + B^{13}) y_{t} = \theta_j(B)\Theta_j(B)^{\prime} \alpha, \]
\[ (1 - B^{12} - B + B^{13} - \varnothing_j(B) + \varnothing_j(B)^{\prime} + \varnothing_j(B)^{\prime} + \varnothing_j(B) - \varnothing_j(B)^{14} + \varnothing_j(B) - \varnothing_j(B)^{15}) y_{t} = \theta_j(B)\Theta_j(B)^{\prime} \alpha, \]
\[ y_t = (I + \varnothing_j(B))y_{t-1} + (\varnothing_j - \varnothing_j(B))y_{t-2} + \varnothing_jy_{t-3} - y_{t-12} + (I + \varnothing_j) y_{t-13} \]
\[ + (\varnothing_j - \varnothing_j(B))y_{t-14} \Theta_jy_{t-15} = \theta_j(B)\Theta_j(B)^{\prime} \alpha, \]
\[ y_t = (I + \varnothing_j(B))y_{t-1} + (\varnothing_j - \varnothing_j(B))y_{t-2} + \varnothing_jy_{t-3} + y_{t-12} - (I + \varnothing_j) y_{t-13} \]
\[ - (\varnothing_j - \varnothing_j(B))y_{t-14} + \varnothing_jy_{t-15} + \theta_j(B)\Theta_j(B)^{\prime} \alpha, \]

(6)

From (6), it indicated that the appropriate lags for ANN model to forecast the residual are lags 1, 2, 3, 12, 13, 14 and 15.

4.2. Artificial neural networks (ANN)

In general, the first step in learning steps of neural network is completed with a defined fixed number of inputs, hidden nodes and outputs. Study had used feed-forward network algorithm with seven inputs, one hidden layer and one output.

4.3. Hybrid Equally-Weighted BJ-ANN

Study combined the forecast values from SARIMA method and ANN method since this study used two methods only. The error measurements of RMSE, MAD and MAPE are stated in Table 2. The forecast values are plotted as in Figure 4.
Figure 4. Actual and forecast values on tourism demand on 2016.

Table 2. Error measurement of all models

<table>
<thead>
<tr>
<th></th>
<th>RMSE</th>
<th>MAD</th>
<th>MAPE</th>
</tr>
</thead>
<tbody>
<tr>
<td>BJ</td>
<td>219671.8090</td>
<td>166474.5833</td>
<td>7.5476</td>
</tr>
<tr>
<td>ANN</td>
<td>156775.4521</td>
<td>138574.0833</td>
<td>6.1814</td>
</tr>
<tr>
<td>Hybrid Equal-Weighted BJ-ANN</td>
<td>126815.4396</td>
<td>106850.0833</td>
<td>4.6903</td>
</tr>
</tbody>
</table>

In view of model performance results, Hybrid Equally-Weighted BJ-ANN ranked the first, followed by Artificial Neural Networks and Box-Jenkins. Hybrid Equally-Weighted BJ-ANN is able to reduce error by at least 35.82% from the BJ model and at least 19.11% from the ANN model. Thus, study concludes that Hybrid Equally-Weighted BJ-ANN is most accurate model to better forecast Malaysia tourism demand.

5. Conclusion

In this study, hybrid model was applied to overcome main limitations, take advantage on superior performance yet unique strength of models BJ and ANN, besides to improve on the accuracy of the forecasting performance. The model used was from individual models of BJ and ANN; and also their hybrid model Hybrid Equally-Weighted BJ-ANN. The data covered the period from 1998 to 2016 of monthly tourist arrivals in Malaysia. By using individual models as the benchmark, the results showed that the hybrid model performs more efficiently in forecasting tourism demand.
This model has the ability to produce a wide range of error reduction percentage compared to individual models. Thus, hybrid method would be well-thought-out to reduce the risk of forecasting failure. This is because this hybrid method is able to explore the advantage or the strength of BJ and ANN in determining different patterns separately [36]. Besides that, ARIMA alone may be inadequate for complex patterned problem, while ANN model can reveal the correlation of nonlinear pattern well [37]. In a nutshell, hybrid model between BJ-ANN proved that it is competent in handling nonlinearity and uncertainty problems within data. This method is flexible in solving both linear and non-linear models by combining the forecast from both methods with the aim of enhancing forecasting performance [5].

6. References

Acknowledgments

We would like to express our thanks to University Tun Hussein Onn Malaysia (UTHM) for providing facilities and financial supports. This research was carried out under UTHM Short Term Grant, Vot U369.