Li-Chang Hsu and Chao-Hung Wang

Applied multivariate forecasting model to tourism industry

Abstract
Various forecast models can be adopted for predicting what types of tourism demand are vulnerable to wide fluctuations. This study employs the fuzzy grey model FGM(1,N) and back-propagation neural networks (BPNN) as the multivariate forecasting models. Various benchmark univariate forecasting models are also employed in this study including the naïve method, exponential smoothing model, Holt's method, and linear regression. We find that the multivariate forecasting models generates more accurate forecasts than univariate models in the tourism service industry. More specifically, the GM(1,N) model was applied to choose the critical influences on tourism demand. Then, FGM(1,N) model was applied to forecast tourism demand using officially published annual data which show tourists traveling from Taiwan to the United States and to Japan during the period 1990-2003. The results showed that the FGM (1,N) outperformed the benchmark statistical methods during the out-of-sample period. Moreover, when important determinants including service price, foreign exchange rate, population, and per capita income are ranked and selected, the GM(1,N) model was improved and achieved viable performance. Finally, in terms of deciding which model to use, the general finding that can be drawn from this study appears to be that in situations involving little sampling data the grey model is superior to other traditional forecasting models. Further discussion and managerial implication can be drawn from these findings.

Keywords: tourism service industry; multivariate forecasting model; service price

Introduction
The tourism industry, which comprises 11.7% of total global economic output and supplies 8% of the total number of jobs worldwide, is one of the largest-scale service industries compared with other competitive industries. Thus, the development of the tourism industry not only supplies numerous job opportunities to local residents but also generates considerable foreign exchange earnings for the economy. The globally accepted definition of the "tourism industry" is of a business with multiple aspects, including hotels, restaurants, transportation, entertainment, craft products, and so on.

Li-Chang Hsu, Department of Finance, Ling Tung University, Taichung, Taiwan, R.O.C.
E-mail: vtcntht87@mail.ltu.edu.tw.
Chao-Hung Wang, Department of Marketing and Logistics Management, Ling Tung University, Taichung, Taiwan, R.O.C.
E-mail: chw@mail.ltu.edu.tw.
The perishable nature of tourism demand means that it is impossible to satisfy fluctuating tourism demand with a stable supply of resources, equipment and employees. Therefore, forecasts are essential due to the importance of accurate and numerous practical applications, including financial statements, electronic power, international passenger demand, and water resources (Lee, Var & Blaine, 1996), and accurate forecasts would help managers to make better tactical strategic decisions. Furthermore, governments can use such forecasts to efficiently plan the tourism infrastructure.


The advantage of the econometrics model (including multiple regressions) is that scholars can identify and analyze variables from numerous affecting factors, and then model the relationship of dependant and independent variables. The drawbacks of the econometrics include high cost, time-consumption, and difficulties in gathering current data. Moreover, time series models including the naïve forecast, moving average, exponential smoothing, and the Box-Jenkins method involves a statistical analysis which employs only the historical data of the univariate being forecast, and this use of easily collected historical data is one of the main attraction features of the model.

However, the above mentioned methods need large amounts of data, normal distribution and stationary data trends. These basic assumptions limit their application validity of forecast modeling. Currently, many scholars are designing new forecasting methods to overcome the limitations of traditional methods. One such method is the artificial intelligence methods that do not require making as many assumptions as traditional statistical methods. An extensive literature exists on methods of forecasting tourism demand. The main aim of this current study is to investigate the feasibility of incorporating influencing multivariate of tourism demand into artificial intelligence, and to rank variables based on their importance.

Morley (1999) argued that modeling tourism demand has become a significant study area in tourism, since it aims to understand what contributes to this growth and occasional breaks in the growth, to assess the impacts of factors on tourism numbers and to produce forecasts. Thus, this investigation constructs five different forecasting methods, including fourteen forecasting models that are tested on the data; namely, the naïve method, exponential smoothing, Holt’s model, the regression model, artificial neural network and the grey forecasting model.

This investigation has the following objectives:
(1) To identify the influences on tourism demand, the grey quantitative relationship between the tourism demand and influential factor was established using multidimensional grey assessment theory.
(2) To formulate the grey forecasting procedure using the influences on tourism demand.
(3) To compare the performance of forecasting models using multivariate and univariate approaches and then select the best of these two models.
The first problem is the selection of a variable to be used as an indicator of tourism demand. In the literature, Gonzalez and Moral (1995) proposed regression forecasting models for international tourism demand using population, income, cost of living, foreign exchange rate, marketing expenditure, and relative prices as independent variables. Lim (1997) surveyed numerous empirical articles on international tourism demand and concluded that the most widely used explanatory variables were income, relative price, and transportation cost. Moreover, Law and Au (1999) and Law (2000) adapted the determinants designed by Gonzalez and Moral (1995). As noted above, the number of Taiwanese tourist arrivals can be calculated as:

\[ TD = f(Y, SP, POP, FER) \]  

where \( TD \) denotes the number of tourist visits from an origin country to a destination country, \( Y \) represents per capita income, \( SP \) is the relative price, \( POP \) denotes the population, and \( FER \) represents the currency exchange rate.

Substantial agreement exists regarding the important explanatory variables in the case of international tourism, such as income, population, foreign exchange and price. Income reflects the magnitude of the market because increased income is required to support increased buying power. According to Engel’s Law, spending on leisure and entertainment will increase as people’s incomes increase. Witt and Witt (1995) noted that origin country income or private consumption is generally included as an explanatory variable, and is commonly included in the demand function in per capita form. Crouch (1994) showed that income is the main explanatory variable. Consequently, income is considered the most important determinant of tourism demand (Song, Wong & Chon, 2003; Dritsakis, 2004). National disposable income (NDI), gross domestic product (GDP), gross national product (GNP), and gross national income (GNI), have also been used in many empirical studies (Song & Turner, 2004). This study investigation uses per capita GNP (US$) as a proxy variable for income level in Taiwan.

In terms of population, Witt and Witt (1995) hypothesized that the level of foreign tourism from a given origin is expected to depend upon the origin population, with demand increasing along with population growth. Sometimes population is a separate explanatory variable but, generally, the influence of population on the consuming market is accommodated by modifying the dependent variable to international tourism demand per capita (Law, 2000; Song & Turner, 2004).

Exchange rate is the price of foreign currency on international money markets. If the exchange rate rises, the buying power will decrease an international commodity markets. A higher exchange rate for the NT dollar to the US dollar thus decreases Taiwanese tourist demand for travel to foreign destinations. The exchange rate is also sometimes used separately to represent tourist living costs, possibly in addition to the exchange-rate-adjusted consumer price index. The justification of varying exchange rate is that consumers are aware of destination travel costs, and thus use the exchange rate as a proxy variable. However, the use of the exchange rates alone can be misleading because, even though the exchange rate for a destination may become more favorable, this could be counterbalanced by a relatively high inflation rate (Song & Witt, 2003; Song & Turner, 2004; Dritsakis, 2004).

Relative price is also an important explanatory variable. Service offering price is more difficult to measure than any other variable. This work cites service price as a proxy for
the relative price, as initially used by Law (2000), paid by Taiwanese visitors for services in the USA and Japan. The formula for calculating service price ($SP$) is denoted as:

$$SP = \frac{CPI_{k,0}}{CPI_{j,0}} \times \frac{CPI_{k,0}}{CPI_{j,0}}$$

(2)

where $SP$ denotes the service price, and $CPI_{k,0}$ represents the base period CPI of the destination country $k$. Moreover $CPI_{j,0}$ is the base period CPI of the original country $j$.

In this study, these four determinants for forecasting tourism demand from Taiwan to Japan and the USA - population, income, foreign exchange rate and service price – were applied. Secondary sources of data, issued by the Taiwan official government and the Taiwan Economic Journal Database, were used in this investigation (Table 1). Tables 2 contains relevant data for forecasting Taiwanese travel demand to the U.S.A. and Japan in the period 1990 to 2003. In each case, annual data from 1990 to 2003 are used with the estimation period 1990–2001 and the post-estimation period 2002–2003.

Table 1

<table>
<thead>
<tr>
<th>Variables</th>
<th>Sources of data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tourism demand</td>
<td>Yearly Statistics, Tourism Bureau Ministry of Transportation and Communications, R.O.C. (Taiwan)</td>
</tr>
<tr>
<td>Populations</td>
<td>Directorate General of Budget Accounting and Statistic</td>
</tr>
<tr>
<td>Per capita income</td>
<td>Executive Yuan, R.O.C. (Taiwan)</td>
</tr>
<tr>
<td>Exchange rate</td>
<td>Central Bank of China (Taiwan)</td>
</tr>
<tr>
<td>Consumer price index</td>
<td>Directorate General of Budget Accounting and Statistic Executive Yuan, R.O.C. Taiwan Economic Journal Database</td>
</tr>
</tbody>
</table>

Table 2

<table>
<thead>
<tr>
<th>Year</th>
<th>Taiwanese tourists to:</th>
<th>Service price</th>
<th>Foreign exchange rate</th>
<th>Taiwanese population</th>
<th>GNP (US$)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>USA</td>
<td>JAPAN</td>
<td>USA</td>
<td>JAPAN</td>
<td>TW$/</td>
</tr>
<tr>
<td>1990</td>
<td>239,325</td>
<td>591,495</td>
<td>1.021</td>
<td>0.829</td>
<td>26.890</td>
</tr>
<tr>
<td>1991</td>
<td>267,584</td>
<td>653,242</td>
<td>1.016</td>
<td>0.833</td>
<td>26.809</td>
</tr>
<tr>
<td>1992</td>
<td>286,966</td>
<td>748,112</td>
<td>1.030</td>
<td>0.855</td>
<td>25.163</td>
</tr>
<tr>
<td>1993</td>
<td>371,750</td>
<td>737,100</td>
<td>1.030</td>
<td>0.869</td>
<td>26.382</td>
</tr>
<tr>
<td>1994</td>
<td>453,924</td>
<td>676,944</td>
<td>1.045</td>
<td>0.898</td>
<td>26.455</td>
</tr>
<tr>
<td>1995</td>
<td>522,910</td>
<td>498,565</td>
<td>1.053</td>
<td>0.932</td>
<td>26.476</td>
</tr>
<tr>
<td>1996</td>
<td>579,488</td>
<td>600,146</td>
<td>1.055</td>
<td>0.960</td>
<td>27.458</td>
</tr>
<tr>
<td>1997</td>
<td>588,916</td>
<td>651,597</td>
<td>1.040</td>
<td>0.952</td>
<td>28.662</td>
</tr>
<tr>
<td>1998</td>
<td>577,178</td>
<td>674,089</td>
<td>1.042</td>
<td>0.962</td>
<td>33.445</td>
</tr>
<tr>
<td>1999</td>
<td>563,991</td>
<td>720,903</td>
<td>1.021</td>
<td>0.967</td>
<td>32.266</td>
</tr>
<tr>
<td>2000</td>
<td>651,134</td>
<td>811,388</td>
<td>1.000</td>
<td>0.985</td>
<td>31.225</td>
</tr>
<tr>
<td>2001</td>
<td>532,010</td>
<td>766,247</td>
<td>0.972</td>
<td>0.991</td>
<td>33.800</td>
</tr>
<tr>
<td>2002</td>
<td>532,180</td>
<td>797,460</td>
<td>0.955</td>
<td>1.000</td>
<td>34.575</td>
</tr>
<tr>
<td>2003</td>
<td>479,264</td>
<td>731,330</td>
<td>0.932</td>
<td>1.000</td>
<td>34.418</td>
</tr>
</tbody>
</table>
As already noted, there are many forecasting models proposed and used. In this section two multivariate models – grey forecasting model and back-propagation neural networks (BPNN) and univariate models selected for comparison in this study will be described in more details.

**MULTIVARIATE MODELS**

**Grey forecasting model**

Grey system theory was initially developed by Deng in early 1980s for considering systems involving uncertainties and information insufficiency. This had extensive applications, including systems analysis, data processing, modeling, forecasting, decision making and control. A forecasting model designed using grey theory is called a grey model (GM). Grey models are provided by grey differential equations, which are groups of abnormal differential equations with variable behavior parameters. Consequently, GM(n,m), where n denotes the order of the differential equation and m represents the number of type of observed data, is defined as a higher-order multivariable differential equation. The GM(1,1) is a single variable first-order grey model that is widely applied to various fields (Hsu & Wen, 1998; Hsu & Wang, 2002; Wang, 2004). However, development and formulation of GM(1,N) in the tourism industry have seldom been investigated.

In grey forecasting, the grey model exploits the accumulated generating operating (AGO) technique to outline system behavior. Assume an original series to be

\[
\begin{align*}
  x_1^{(0)} &= \left\{ x_1^{(0)}(1), x_1^{(0)}(2), \ldots, x_1^{(0)}(k) \right\} \\
  x_2^{(0)} &= \left\{ x_2^{(0)}(1), x_2^{(0)}(2), \ldots, x_2^{(0)}(k) \right\} \\
  &\vdots \\
  x_N^{(0)} &= \left\{ x_N^{(0)}(1), x_N^{(0)}(2), \ldots, x_N^{(0)}(k) \right\}, \quad k = 1, 2, \ldots, n.
\end{align*}
\]

where \( x_i^{(0)} \) is the predicted series, \( x_j^{(0)} \), \( j = 2, 3, \ldots, N \) is the associated series.

Then one-order AGO is defined as:

\[
\begin{align*}
  x_1^{(1)} &= \left\{ x_1^{(1)}(1), x_1^{(1)}(2), \ldots, x_1^{(1)}(k) \right\} \\
  x_2^{(1)} &= \left\{ x_2^{(1)}(1), x_2^{(1)}(2), \ldots, x_2^{(1)}(k) \right\} \\
  &\vdots \\
  x_N^{(1)} &= \left\{ x_N^{(1)}(1), x_N^{(1)}(2), \ldots, x_N^{(1)}(k) \right\}, \quad k = 1, 2, \ldots, n
\end{align*}
\]

where \( x_i^{(1)}(k) = \sum_{j=1}^{k} x_i^{(0)}(j), \quad i = 1, 2, \ldots, N \).

A grey differential equation with N variables is called GM(1,N) and can be expressed as follows:

\[
\frac{dx_i^{(1)}(k)}{dt} + a_i z_1^{(1)}(k) = \sum_{i=2}^{N} b_i x_i^{(1)}(k), \quad k = 1, 2, \ldots, n
\]

where the value of \( a_i \) indicates the coordinate degree between index \( x_i \) and various factors. Moreover, \( b_i \) is the ith influence coefficient, which means that \( x_i \) influences \( x_i \) (the behavior variable) and \( z_i^{(1)}(k) = 0.5x_i^{(1)}(k) + 0.5x_i^{(1)}(k-1), \quad k \geq 2 \).
The parameter of GM(1,N) is \( \hat{\mathbf{a}} = [a_1, b_2, \ldots, b_N]^T \), based on the least squared method, is denoted as follows:

\[
\hat{\mathbf{a}} = (B^T B)^{-1} B^T \mathbf{Y}_N, \tag{6}
\]

where,

\[
B = \begin{bmatrix}
-\hat{z}_1^{(1)}(2) & x_2^{(1)}(2) & \cdots & x_N^{(1)}(2) \\
-\hat{z}_1^{(1)}(3) & x_2^{(1)}(3) & \cdots & x_N^{(1)}(3) \\
\vdots & \vdots & \ddots & \vdots \\
-\hat{z}_1^{(1)}(n) & x_2^{(1)}(n) & \cdots & x_N^{(1)}(n)
\end{bmatrix}
\]

and \( \mathbf{Y}_N = \begin{bmatrix}
x_1^{(0)}(2) \\
x_1^{(0)}(3) \\
\vdots \\
x_1^{(0)}(n)
\end{bmatrix} \]

The optimal \( a_i \) and \( b_i \) are then determined and their correlation is illustrated in Eq. (6). The value of \( a_i \) indicates the developing coefficient for \( x_1^{(i)} \). These coefficients \( |b_i| \) display the degree of correlation between indices and factors.

For forecasting purpose, FGM(1,N) is derived from GM(1,N), the FGM(1,N) whitened differential equation is denoted by:

\[
\frac{dx_1^{(1)}(k)}{dt} + a_1 x_1^{(1)}(k) = \sum_{i=2}^{N} b_i x_i^{(1)}(k), \quad k = 1, 2, \ldots, n.
\]

Then the forecasting function of FGM(1,N) is defined as (Deng and Guo, 1996):

\[
x_1^{(0)}(k) = \sum_{i=2}^{N} \beta_i x_i^{(1)}(k) - \alpha x_1^{(1)}(k-1), \quad k = 2, 3, \ldots, n \tag{7}
\]

where \( \alpha = \frac{a}{1+0.5a} \), \( \beta_i = \frac{b_i}{1+0.5a} \), \( i = 2, 3, \ldots, N \).

When \( N = 1 \), Eq. (7) is transformed into a multivariate model GM(1,1), is used to forecast the time series with one variable. Then the forecasting function of GM(1,1) is rewritten as (Hsu, 2003; Wang, 2004):

\[
\hat{x}_1^{(0)}(k) = \begin{bmatrix} x^{(0)}(1) - \frac{b}{a} \end{bmatrix} (1 - e^{-a(k-1)}) e^{-a(k-1)}, \quad k = 2, 3, \ldots, n \tag{8}
\]

where \( a, b \) is a parameter, which is estimated using OLS.

**Back-propagation neural networks (BPNN)**

Artificial neural networks (ANNs) are a class of nonlinear models inspired by studies of the brain and nerve system. The advantage of ANNs over other conventional statistical models is that they can model complex, possibly nonlinear relationships without requiring any prior assumptions regarding the underlying data generating process. Back propagation neural networks (BPNN) are one sub-class of ANNs. The BPNN is a supervised feed-forward neural network that uses training data as a basis for identifying patterns representing input and output variables.
The BPNN has been widely applied in various fields, particularly for forecasting tourism demand (Law, 2000; Burger et al., 2001). BPNN is comprised of two processes: learning and recalling. During the learning stage, certain variables and referent output are calculated based on the processing data embedded in the learning sample. The amending weight and bias are then iterated using the gradient steepest descent method, and converging finally to an allowance output. To summarize, ANNs are expected to outperform traditional statistical methods in forecasting due to better recognition of the high-level features of processing the data, and using a small-sized training set.

UNIVARIATE MODELS

In addition to non-traditional procedures, various benchmark statistical approaches are employed in this study, namely the naïve method, exponential smoothing model, Holt's method, and linear regression. The four models are types commonly used in the tourism industry (Law & Au, 1999; Law, 2000).

The naïve method is very simple (Law & Au, 1999; Law, 2000), stating that the value of the period to be forecast equals the value of the last period for which information is available: \( F_t = A_{t-1} \). The simplest form of exponential smoothing is the single exponential smoothing model (Gaynor and Kirkpatrick, 1994; Law and Au, 1999), formally started as: \( F_t = F_{t-1} + \alpha(A_{t-1} - F_{t-1}) \), where \( \alpha = \) smoothing constant, \( F = \) forecast value, and \( A = \) actual value.

Another frequently used exponential smoothing model is Holt's method. The data series values are smoothed by weighting the current level and the previous smoothed adjusted estimate of the trend: \( \tilde{A}_t = \alpha A_{t-1} + (1-\alpha)(\tilde{A}_{t-1} - \tilde{B}_{t-1}) \). Moreover, the trend estimates are provided by the weighted differences of two smoothing series and the previous trend, \( t-1: \tilde{B}_t = \beta(\tilde{A}_{t-1} - \tilde{A}_t) + (1-\beta)\tilde{B}_{t-1} \). Moreover, the \( i \)-period-ahead forecast is given by: \( F_{t+i} = \tilde{A}_t + i\tilde{B}_{t+i} \). The values of both \( \alpha \) and \( \beta \) are allowed to vary to enable the identification of the optimum values of these parameters; that is, this study attempts to identify the values of \( \alpha \) and \( \beta \) that minimizes the mean square error.

The function of simple linear regression (LR) is: \( \hat{T_D}_t = a + bt \), where \( a, b \) are the parameters to be estimated, \( t \) denotes time, \( T_D \) is the sample value, and \( \hat{T_D}_t \) denotes the prediction value of period \( t \).

Multiple regression is widely used in tourism research for estimating tourism demand functions (Ong, 1995), which are required for planning police making, and also for budgeting purposes by tourism operators, investors and government bureaus. The main studies employing multiple regression analysis of linear models were performed by Morley (1997), Lim (1997) and Law (2000). Multiple linear regression models have the following general form: \( T_D_t = a + b_1Y_t + b_2SP_t + b_3POP_t + b_4FER_t \), where \( a \) = the intercept constant, and \( b_1, b_2, \ldots, b_4 = \) slope coefficient.
MAGNITUDE OF FORECASTING ERROR

The forecasting accuracy of the models are measured mainly using Theil inequality coefficient (TIC), mean absolute percentage error (MAPE) and root mean square error (RMSE). These measures have already been presented in detail by others, including Law (2000), Song et al. (2000), Kon and Turner (2005). The TIC is defined as

$$TIC = \frac{\sqrt{\frac{1}{n} \sum_{t=1}^{n} (\hat{F}_t - A_t)^2}}{\left(\frac{1}{n} \sum_{t=1}^{n} \hat{F}_t^2 + \frac{1}{n} \sum_{t=1}^{n} A_t^2\right)}.$$  

The Theil inequality coefficient (TIC) lie between [0,1] where 0 is represents a perfect fit, 1 is represents a poor prediction. Mean absolute percentage error is defined as

$$MAPE = \left\{\frac{\sum_{t=1}^{n} (A_t - \hat{F}_t) / A_t}{n}\right\} \times 100\% ;$$

and root mean square error is defined as

$$RMSE = \sqrt{\frac{\sum_{t=1}^{n} (A_t - \hat{F}_t)^2}{n}},$$

where $\hat{F}_t$ denotes the predicted number of visitors, and $A_t$ represents the actual number of visitors. The model yielded plausible prediction values when $MAPE$ was low, and vice versa. Lewis (1982) summarized the criteria of $MAPE$ in Table 3.

<table>
<thead>
<tr>
<th>MAPE</th>
<th>Forecasting power</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;10%</td>
<td>Highly accurate forecasting</td>
</tr>
<tr>
<td>10%–20%</td>
<td>Good forecasting</td>
</tr>
<tr>
<td>20%–50%</td>
<td>Reasonable forecasting</td>
</tr>
<tr>
<td>&gt;50%</td>
<td>Weak and inaccurate forecasting</td>
</tr>
</tbody>
</table>

Source: Lewis (1982)

Empirical results

To test these forecasting models data from Taiwan will be used. Taiwanese economy has grown significantly during the past decade, leading to increasing per capita income, improved living standards, and increased leisure and tourism consumption. Since 1981, the government of Taiwan permitted its residents to travel abroad. Gradually, Taiwanese tourists turned into major tourist sources of other countries. According to the publications of the World Travel and Tourism Council (WTTC), the rate of tourist arrivals from Taiwan has grown faster than that from other countries in Asia. Apart from Hong Kong, the USA and Japan are the two main destination countries. In addition, the significant growth of 0.203 million inbound tourists annually from Taiwan to Japan far exceeded the growth rate for other destination countries. According to official Taiwanese data published in 2003, the United States and Japan are the main markets for Taiwanese tourists in the area of America and Asia, respectively.

Fourteen forecasting models were employed to forecast Taiwanese tourist arrival in USA and Japan based on data from the period of 1990-2001 (i.e., in-sample data). The forecast accuracy is inspected using the out-of-sample data for the 2002-03 periods.
The difference between the actual and the forecast for 2002-03 were used to evaluate
the accuracy of the forecasting model.

EMPIRICAL RESULTS OF THE GREY FORECASTING MODEL
GM(1,5) attempts to identify the relevant variables and estimate the relationship
between the independent and dependent variables in terms of parameters. All of the
models considered in this section operate within an Excel spreadsheet environment. To
build the GM(1,5) model of U.S.A., we first obtain original series as follows.

\[
x_1^{(0)} = \{239325, 267584, \cdots, 532010\}
\]
\[
x_2^{(0)} = \{20401, 20606, \cdots, 22406\}
\]
\[
x_3^{(0)} = \{1.021, 1.016, \cdots, 0.972\}.
\]

From Eq. (4), one-order AGO sequence is defined as follows:

\[
x_1^{(1)} = \{239325, 506909, \cdots, 563517\}
\]
\[
x_2^{(1)} = \{20401, 41007, \cdots, 257312\}
\]
\[
x_3^{(1)} = \{1.021, 2.037, \cdots, 12.325\}.
\]

Additionally, matrix \( B \) and constant vector \( Y_N \) are accumulated as follows:

\[
B = \begin{bmatrix}
-373117 & 41007 & \cdots & 2.037 \\
-650392 & 61810 & \cdots & 3.067 \\
\vdots & \vdots & & \vdots \\
-5369171 & 257312 & \cdots & 12.325
\end{bmatrix},
Y_N = \begin{bmatrix}
267584 \\
286966 \\
\vdots \\
532010
\end{bmatrix}.
\]

From Eq. (6), the parameters are estimated using the OLS method, yielding the following matrix.

\[
\hat{a}_{U.S.A} = [a_1, b_2, \cdots, b_5]^T = \begin{bmatrix}
0.695, -58.33, 46.12, 10373.73, 741148.47
\end{bmatrix}^T \quad (9)
\]

The relationship between the TD sequence and the influence sequence (Pop, Y, FER, SP) can be found by comparing the value of \( b_2, \cdots, b_5 \). Following the same procedures
as the abovementioned steps, we can calculate the parameters of Japan. The parameters
(\( \hat{a} \)) are estimated using the OLS method, yielding the following matrix.

\[
\hat{a}_{JP} = [a_1, b_2, \cdots, b_5]^T = \begin{bmatrix}
0.15, -17.05, -0.07, 8037.16, 891054.68
\end{bmatrix}^T \quad (10)
\]

Eq.(9) yielded the same ranking result as that of Eq.(10), that is, \( |b_5| > |b_4| > |b_2| > |b_3| \). The sequential function demonstrates that the relative importance of factors is ranked
based on service price \( (x_1) \) > exchange rate \( (x_2) \) > populations \( (x_3) \) > per capita
income \( (x_5)\). Clearly, the service price has more impact than any of the other factors
considered. Regarding forecasting, the parameter matrix of $\hat{a}_{U.S.A}$ and $\hat{a}_{JP}$ are transformed into a $\alpha$ value and a matrix of $\beta$. The fitted models used for export forecasting of tourist arrivals from Taiwan are as follows.

\[
\hat{x}_{UL}^{(0)}(k) = -43.29x_1^{(0)}(k) + 34.23x_2^{(0)}(k) + 7698.9x_3^{(0)}(k) + 550046.5x_4^{(0)}(k) - 0.516x_5^{(0)}(k - 1) \quad (11)
\]

\[
\hat{x}_{UL}^{(0)}(k) = -15.856x_1^{(0)}(k) - 0.06 x_2^{(0)}(k) + 7474x_3^{(0)}(k) + 828618.17 x_4^{(0)}(k) - 0.144x_5^{(0)}(k - 1) \quad (12)
\]

Equations (11) and (12) represent the FGM(1,5) forecasting models for the U.S.A. and Japan, respectively.

Song and Turner (2004) argued that it is very difficult to obtain a single model that consistently outperforms all other models in all situations. Hence, according to the GM(1,5) ranking results, we built FGM(1,5) for $TD = f(Pop, Y, FER, SP)$, FGM(1,4) for $TD = f(Pop, FER, SP)$, FGM(1,3) for $TD = f(FER, SP)$, and FGM(1,2) for $TD = f(SP)$ models for greater forecasting accuracy. The out-of-sample forecast performance is listed in Tables 4 and 5, respectively.

**EMPIRICAL RESULTS OF BPNN**

All neural networks were estimated (trained) and forecasted using PCNeuron (Yeh, 2001). The input layer comprises Y, Pop, FER, and SP, while the output layer is tourism demand (TD). The data include eight periods of discrete data for training and two periods of discrete data (1999-2001) for testing.

The value of the learning rate $\eta$ and the momentum factor $\alpha$ will significantly affect the efficiency and the convergence of neural network learning algorithm. In addition to this, the number of neurons in the input layer also affects the performance of the neural network. To build a forecasting model, the parameter values were investigated by running the neural networks with different hidden nodes (from 3 to 10), number of epochs (from 300 to 2000 by 200), learning rates (from 0.1 to 2 by 0.2), and moment rates (from 0.1 to 0.8). A stopping criterion was established to determine when to terminate the training process to achieve optimum performance (minimize $\text{RMS}_E$) to avoid overtraining. The best forecasting result and errors are listed in Tables 4 and 5, respectively.

**EMPIRICAL RESULTS OF BENCHMARK STATISTICAL METHODS**

Historically, the exponential smoothing model uses a smoothing average of past values, in which the weight reduces geometrically over time to suppress short term fluctuations in the data. The smoothing constant ($\alpha$) ranges between 0.1 and 1, and was selected to be 0.98 with the smallest mean square error. Holt’s method is a two parameter exponential model. At the end of time period $t$, Holt’s method generates an estimate of the base level and the per period of the series. The two smoothing constants $\alpha$ and $\beta$ are set to 0.9 and 0 for the USA example and 1 and 0 for the Japan example.

This study used Excel’s data analysis tool for the linear regression analysis. The fitted equations are $TD_{USA}^{f^a} = 240067.2 + 35312.43 t$ and $TD_{JP}^{f^p} = 612524.9 + 9981.147 t$. The above mentioned empirical results are shown in Table 4 and 5.
**COMPARING OF THE OUT-OF-SAMPLE FORECASTING RESULTS**

Considering all of the empirical results for out-of-sample (see Table 4 and 5), based on the univariate model results, for case of travel to the USA case, naïve model had the smallest MAPE (5.54%), RMSE (37,417.46) and TIC (0.0360). For the case of Japan, the LR and GM(1,1) model had low TIC value almost equal to zero. The GM(1,1) model had the smallest MAPE (4.85%), had excellent forecasting power (Table 3).

---

**Table 4**

**EMPIRICAL RESULTS OF FORECASTING TAIWANESE DEMAND FOR TRAVEL TO THE U.S.A.**

<table>
<thead>
<tr>
<th>Model</th>
<th>2002</th>
<th>2003</th>
<th>RMSE*</th>
<th>TIC*</th>
<th>MAPE*</th>
<th>Forecasting Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual value</td>
<td>532,180.00</td>
<td>479,264.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Naïve</td>
<td>532,010.00</td>
<td>532,180.00</td>
<td>37,417.46</td>
<td>0.0360</td>
<td>5.54</td>
<td>Excellent</td>
</tr>
<tr>
<td>Exponential smoothing (0.98)</td>
<td>534,357.73</td>
<td>532,223.55</td>
<td>37,479.70</td>
<td>0.0361</td>
<td>5.73</td>
<td>Excellent</td>
</tr>
<tr>
<td>Holt's (0.9,0)</td>
<td>543,065.17</td>
<td>533,268.52</td>
<td>38,954.94</td>
<td>0.0373</td>
<td>6.66</td>
<td>Excellent</td>
</tr>
<tr>
<td>LR</td>
<td>699,128.80</td>
<td>734,441.20</td>
<td>215,623.87</td>
<td>0.1762</td>
<td>42.31</td>
<td>Reasonable</td>
</tr>
<tr>
<td>MR (1,4)</td>
<td>537,546.83</td>
<td>524,194.93</td>
<td>31,996.80</td>
<td>0.0308</td>
<td>5.19</td>
<td>Excellent</td>
</tr>
<tr>
<td>BPNN</td>
<td>434,490.00</td>
<td>376,230.00</td>
<td>100,397.56</td>
<td>0.1100</td>
<td>19.93</td>
<td>Good</td>
</tr>
<tr>
<td>GM (1,1)</td>
<td>698,929.43</td>
<td>742,985.27</td>
<td>220,628.97</td>
<td>0.1797</td>
<td>43.18</td>
<td>Reasonable</td>
</tr>
<tr>
<td>FGM (1,5)</td>
<td>528,190.15</td>
<td>502,843.86</td>
<td>16,910.48</td>
<td>0.0165</td>
<td>2.84</td>
<td>Excellent</td>
</tr>
</tbody>
</table>

* RMSE is the root mean square error, TIC is the Theil inequality coefficient and MAPE is the mean absolute percentage error.

**Table 5**

**EMPIRICAL RESULTS OF FORECASTING TAIWANESE DEMAND FOR TRAVEL TO JAPAN**

<table>
<thead>
<tr>
<th>Model</th>
<th>2002</th>
<th>2003</th>
<th>RMSE*</th>
<th>TIC*</th>
<th>MAPE*</th>
<th>Forecasting Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual value</td>
<td>797,460.00</td>
<td>731,330.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Naïve</td>
<td>765,243.00</td>
<td>797,460.00</td>
<td>52,014.96</td>
<td>0.0336</td>
<td>6.54</td>
<td>Excellent</td>
</tr>
<tr>
<td>Exponential Smoothing (0.98)</td>
<td>766,133.25</td>
<td>796,833.46</td>
<td>51,342.32</td>
<td>0.0332</td>
<td>6.44</td>
<td>Excellent</td>
</tr>
<tr>
<td>Holt's (1,0)</td>
<td>765,247.00</td>
<td>797,460.00</td>
<td>52,013.72</td>
<td>0.0336</td>
<td>6.54</td>
<td>Excellent</td>
</tr>
<tr>
<td>LR</td>
<td>742,279.81</td>
<td>752,260.96</td>
<td>41,731.05</td>
<td>0.0276</td>
<td>4.89</td>
<td>Excellent</td>
</tr>
<tr>
<td>MR (1,3)</td>
<td>727,208.49</td>
<td>713,638.42</td>
<td>51,226.30</td>
<td>0.0345</td>
<td>5.61</td>
<td>Excellent</td>
</tr>
<tr>
<td>BPNN</td>
<td>639,570.00</td>
<td>639,150.00</td>
<td>129,279.55</td>
<td>0.0920</td>
<td>16.20</td>
<td>Good</td>
</tr>
<tr>
<td>GM (1,1)</td>
<td>739,520.43</td>
<td>749,099.90</td>
<td>42,853.02</td>
<td>0.0284</td>
<td>4.85</td>
<td>Excellent</td>
</tr>
<tr>
<td>FGM (1,2)</td>
<td>800,247.07</td>
<td>771,110.85</td>
<td>28,198.26</td>
<td>0.0182</td>
<td>2.90</td>
<td>Excellent</td>
</tr>
</tbody>
</table>

* RMSE is the root mean square error, TIC is the Theil inequality coefficient and MAPE is the mean absolute percentage error.
Base on the multivariate model results, for the U.S.A. case, FGM(1,5) had the smallest MAPE (2.84%), RMSE (16910.48) and TIC (0.0165), for the Japan case, FMG(1,2) had the smallest MAPE (2.90%), RMSE (28198.26) and TIC (0.0182), that is to say, had highly accurate forecasting power (Table 3). Moreover, all of the Theil inequality coefficient value almost equal to zero. These results indicate that the forecasting performance of the model is relatively good. From the point of view of Witt and Witt (1992) or Law (2000), the smallest MAPE had "highly" accurate forecasting. We may proceed from this to the conclusion that FGM(1,N) model may apply to the case of tourism demand forecasting.

Conclusions

This study presents a novel means of forecasting tourism demand, called FGM(1,N), which is based on grey theory. Due to the lack of large quantities of sample data, the grey model is very useful for establishing an accurate forecasting model. Consequently, this study assesses the impact of independent variables on tourism demand, and then forecasts future trends using the FGM(1,N) model.

The following conclusions are reached based on the empirical results. First, the FGM(1,N) outperformed the benchmark statistical methods during the out-of-sample period, while the training and learning results were poor for the BPNNs that lacked large amounts of data. Second, when important independent variables, including service price, foreign exchange rate, population, and per capita income are ranked and selected, the GM(1,N) model was improved and achieved viable performance. For the case of Taiwanese tourist arrival in America and Japan, this study shows that service price (SP) is the most important factor that affects the decisions of tourist to travel. The exchange rate (FER) is another important variable of demand determinant. Third, in terms of deciding which model to use, the general finding that can be drawn from this study appears to be that in situations involving little sampling data the grey model is superior to other traditional forecasting models.

The findings of this work have important managerial and practical implications. Tourism service providers should assess the costs and benefits of each model before choosing an appropriate forecasting tool. The ten models in this work differ in terms of the necessary amounts of historical data, personnel background, and software requirements. Since the BPNN model performs poorly overall due to the lack of large amounts of historical data to training and learning, thus BPNN requires special software and artificial intelligence expertise. The exponential models and multiple regression can be applied using existing software and do not require complicated statistical skills, but the sample data must satisfy the basic assumptions, such as normal distribution and homogeneous variants. Therefore, tourism managers must carefully assess the trade-off between forecasting accuracy and the above limitations of available methods.

References


Received: 10/30/2007
Accepted: 04/29/2008