Available online at www.sciencedirect.com
SCIENCE DIRECT
PERGAMON
Tourism Management 25 (2004) 367–374
TOURISM MANAGEMENT
www.elsevier.com/locate/tourman

Predicting tourism demand using fuzzy time series and hybrid grey theory

Chao-Hung Wang*

International Trade Department, Ling Tung College, 1 Ling Tung Rd, Nantun 408, Taichung, Taiwan, ROC

Received 12 November 2002; accepted 5 May 2003

Abstract

Forecasting tourism demand in a capacity constrained service industry has been a major theme in this field. This study presents two models that can be used to predict tourism demand. Both two models are based on artificial intelligent (AI). Neural network theory was first applied to tourism demand forecasting in 2000 and empirically tested using the raw data from Hong Kong. This work provides empirical evidence using grey theory and fuzzy time series, which do not need large sample and long past time series. These AI models are estimated for tourist arrivals to Taiwan from Hong Kong, United States and Germany during the period of 1989–2000. GM(1,1) model achieves an accurate forecast when the sample data show a stable increase trend. Nevertheless, the Markov modification model can efficiently improved the GM(1,1) model when the sample data show significant fluctuations.

Keywords: Tourist demand; Forecast; Fuzzy time series; GM(1,1)

1. Introduction

A service industry with a fixed level of capacity encounters large fluctuations in demand for its product. Under conditions of fluctuation, an uncertainty manager may make wrong decisions and always turn customers away in peak periods. In low periods, facilities are idle and employees are standing around. This study explores the nature of the demand of tourism that is a major challenge for many tourism marketers and sightseeing organization, which include public non-profit governments sectors and private profit-companies.

According to data from the Tourism Bureau of Republic of China (ROC), a trend of gradually increasing tourist number is obvious in Taiwan within last decade. In 2002, there were some 2.4 million tourists arriving in Taiwan. The major tourist originating countries are the United States, Hong Kong, and Southeast Asian countries.

Forecasting methods include uni-variant (using past data), multi-variant (using the relationship among many variants) and qualitative analysis (using researcher's judgment). Time series used to forecast future trends include exponential smoothing, ARIMA and trend analysis. Multi-variant prediction methods include multi regression model, econometrics and state space. Delphi marketing research, situational analysis and historical analogue belong to qualitative methodologies.

The above-mentioned forecasting methods were developed to forecast trends over different time horizon. There are large differences in time length being considered when using different forecast methods. Basically, uni-variant methods in short-term forecasts usually generate higher accuracy than those of multi-variants. Therefore, one of the major objectives of this paper is to develop a short-term forecast method to enable the government sector to make appropriate decisions.

The topic of tourism demand forecasting is not an immaturity and a significant literature exists, but most of it concentrates on traditional time series. For example, Lim and McAleer (2002) use Box-Jenkins' Autoregressive Integrated Moving Average (ARIMA) model to forecast tourist arrivals to Australia from Hong Kong, Malaysia and Singapore. Goh and Law (2002) present the use of the time series SARIMA and MARIMA with interventions in forecasting tourism demand using ten arrival series for Hong Kong. Neural

*Tel.: +886-4-2389-5624x808; fax: +886-4-2386-4342.
E-mail address: chw@webmail.ltc.edu.tw (C.-H. Wang).

0261-5177/$ - see front matter
doi:10.1016/S0261-5177(03)00132-8

networks were first applied to forecast tourism demand by Law and Au (1999) to forecast Japanese demand for travel to Honk Kong. Experimental results showed that neural network model forecasts outperformed multiple regression, moving average, and exponent smoothing. Law (2000) extended the applicability of neural networks in tourism demand forecasting by incorporating the back-propagation learning process into non-linearly separable tourism demand data. Empirical results indicated that utilizing a back-propagation neural network outperformed regression models, time-series model and feed-forward neural networks in terms of forecasting accuracy.

Service industries are currently undergoing dramatic changes compared to manufacturing industries. Many factors underlie the ongoing transformation of service management. Among factors obstructing, forecasting is the ambiguous and constrained nature of available information in the short-term. Therefore, the major objective of this work is to present methods of artificial intelligence that overcome the disadvantages of conventional forecast methods. Artificial intelligence means that a computer-based system has capabilities of problem solving, storing memory and understanding human language. The research tools of artificial intelligent (AI) include fuzzy theory, grey theory, neural network model, Genetic Algorithms and expert systems.

One constraint is the necessity of large sample for the ARIMA and neural network to produce low forecast errors, but unfortunately it is difficult to gather such data in dramatic changing environments in a short period of time.

2. Research methodology

After reviewing the literature, the only AI methodology applied to tourism demand forecast was that of neural networks. This study presents grey theory and fuzzy time series, which are suitable for short-term time series data given a shortage of sample data.

2.1. Fuzzy time series

Fuzzy theory, originally explored by Zadeh in 1965, describes linguistic fuzzy information using mathematical modelling. Because the existing statistical time series methods could not effectively analyze time series with small amounts of data, fuzzy time series methods were developed. Song and Chissom (1993a, b) (hereafter S&C) proposed a first-order time-invariant model and a time-variant model of fuzzy time series in 1993. They fuzzified the enrollment at the University of Alabama in 1993 in the first application of fuzzy time series to forecasting. Then, in 1994, they proposed a new fuzzy time series and compared three different defuzzification models. The empirical result showed that the best prediction results are obtained when the neural network method is applied to defuzzify the data (Song & Chissom, 1994).

Chen (1996) considered that the neural network method is too complicated to apply; he therefore presented arithmetic operations instead of the logic max–min composition. The arithmetic operations have a robust specification and are superior to those applied in S&C's model.

Hwang, Chen, and Lee (1998) defined a fuzzy set for each year, established the fuzzy relationship, and finally forecast the enrollment in the University of Alabama using the relation matrix. Empirical analysis revealed that the average error rate of Hwang's model was smaller than those of Chen and S&C.

The following steps construct Hwang's fuzzy time series (Hwang et al., 1998):

Step 1: Calculate the variations using the historical data.

Step 2: Separate the universe of discourse U into several even-length intervals. In this step, the universe must first be defined: it includes the minimum number of units ($D_{\min}$) and the maximum number of units ($D_{\max}$), according to known historical data. Based on $D_{\min}$ and $D_{\max}$, the universe U is defined as $[D_{\min}-D_1, D_{\max}+D_2]$. D_1 and D_2 are two proper positive numbers. Then, U is divided into intervals with equal length.

Step 3: Define the fuzzy time series $F(t)$. The fuzzy time series is expressed as follows. $F(t) = I_{C1}/u_1 + I_{C2}/u_2 + \cdots + I_{Cm}/u_m$, where I_{Ci} is the memberships and $0 \leqslant I_{Ci} \leqslant 1$. Thus, the fuzzy sets ($A_i$) are expressed as: $A_i = \{I_{C1}/u_1, I_{C2}/u_2, \ldots, I_{Cm}/u_m\}$.

Step 4: Fuzzify the variations of historical data. This step determines a fuzzy set equivalent to each set of data. If the variation falls within u_i, then the degree of each historical datum belongs to each A_i is determined.

Step 5: Calculate the relation matrix $R(t)$. Two variables, which the operation matrix $O^w(t)$ and criterion matrix $C(t)$, govern the relation matrix $R(t)$, where w ($w = 2,3,\ldots,n$) is the window base and t is year. The operation matrix is expressed as follows:

$$O^w(T) = \begin{bmatrix} F(t-2) \\ F(t-3) \\ \vdots \\ F(t-w-1) \end{bmatrix} = \begin{bmatrix} O_{11} & O_{12} & \cdots & O_{1m} \\ O_{21} & O_{22} & \cdots & O_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ O_{w1} & O_{w2} & \cdots & O_{wm} \end{bmatrix}.$$

The criterion matrix is expressed as follows:

$C(t) = F(t-1) = [C_1, C_2, \ldots, C_m]$ where C_1 represents "decreases"; C_2 represents "increases a little", and C_m represents "increases too much". The relation for changing the degree of period t is thus obtained. The relation matrix $R(t)$ is expressed as follows:

$$R_{ij}(t) = O_{ij}^{w}(t) \times C_j(t), \quad 1 \leqslant i \leqslant w, \quad 1 \leqslant j \leqslant m,$$

$$R(t) = \begin{bmatrix} O_{11} \times C_1 & O_{12} \times C_2 & \cdots & O_{1m} \times C_m \\ O_{21} \times C_1 & O_{22} \times C_2 & \cdots & O_{2m} \times C_m \\ \vdots & \vdots & \vdots & \vdots \\ O_{w1} \times C_1 & O_{w2} \times C_2 & \cdots & O_{wm} \times C_m \end{bmatrix} = \begin{bmatrix} R_{11} & R_{12} & \cdots & R_{1m} \\ R_{21} & R_{22} & \cdots & R_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ R_{w1} & R_{w2} & \cdots & R_{wm} \end{bmatrix},$$

$$F(t) = [\mathrm{Max}(R_{11}, R_{21}, \ldots, R_{w1})\ \mathrm{Max}(R_{12}, R_{22}, \ldots, R_{w2}) \ldots \times \mathrm{Max}(R_{1m}, R_{2m}, \ldots, R_{wm})] = [r_1, r_2, \ldots, r_m].$$

Step 6: Defuzzify the fuzzified predicted variations in Step 5. The principles of defuzzification are as follows:

(1) If the membership of an output has only one maximum u_i, then select the midpoint of the interval that corresponds to the maximum forecast value.
(2) If the membership of an output has one or more consecutive maximum, then select the midpoint of the corresponding conjunct interval as the forecast.
(3) If the membership of an output is zero, then no maximum exists. Thus, the predicted degree of change is zero.

Step 7: Calculate the outputs. The actual value of change for the preceding year is added to the forecast degree of change, yielding the forecast value for this year.

2.2. Grey forecasting model

Grey theory, originally developed by Deng (1982), focuses on model uncertainty and information insufficiency in analyzing and understanding systems via research on conditional analysis, prediction and decision making. In the field of information research, deep or light colours represent information that is clear or ambiguous, respectively. Meanwhile, black indicates that the researchers have absolutely no knowledge of system structure, parameters, and characteristics; while white represents that the information is completely clear. Colours between black and white indicate systems that are not clear, such as social, economic, or weather systems.

The grey forecasting model adopts the essential part of the grey system theory and it has been successfully used in finance, integrated circuit industry and the market for air travel (Hsu & Wang, 2002; Hsu, 2003; Hsu & Wen, 1998). The grey forecasting model uses the operations of accumulated generation to build differential equations. Intrinsically speaking, it has the characteristics of requiring less data. The GM(1,1), can be denoted by the function as follows (Hsu, 2001):

Step 1: Assume an original series to be $x^{(0)}$, $x^{(0)} = x^{(0)}(1), x^{(0)}(2), x^{(0)}(3), \ldots, x^{(0)}(n)$.

Step 2: A new sequence $x^{(1)}$ is generated by the accumulated generating operation (AGO).

$$x^{(1)} = (x^{(1)}(1), x^{(1)}(2), x^{(1)}(3), \ldots, x^{(1)}(n)),$$

$$\text{where } x^{(1)}(k) = \sum_{i=1}^{k} x^{(0)}(i).$$

Step 3: Estabishing a first-order differential equation. $(\mathrm{d}x^{(1)}/\mathrm{d}t) + az = u$, where $z^{(1)}(k) = \alpha x^{(1)}(k) + (1-\alpha)x^{(1)}(k+1)$, $k = 1, 2, \ldots, n-1$. α denotes a horizontal adjument coefficient, and $0 < \alpha < 1$. The selecting criterion of α value is to yield the smallest forecasting error rate (Wen, Huang, & Wen, 2000).

Step 4: From Step 3, we have

$$\hat{x}^{(1)}(k+1) = \left(x^{(0)}(1) - \frac{u}{a}\right)\mathrm{e}^{-ak} + \frac{u}{a},$$

$$\text{where } \hat{\theta} = \begin{bmatrix} a \\ u \end{bmatrix} = (B^{\mathrm{T}}B)^{-1} B^{\mathrm{T}}\, Y,$$

$$B = \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \ldots & \ldots \\ -z^{(1)}(n) & 1 \end{bmatrix},$$

$$Y = (x^{(0)}(2), x^{(0)}(3), \ldots, x^{(0)}(n))^{\mathrm{T}}.$$

Step 5: Inverse accumulated generation operation (IAGO). Because the grey forecasting model is formulated using the data of AGO rather than original data, IAGO can be used to reverse the forecasting value. Namely

$$\hat{x}_{\mathrm{o}}^{(0)}(k) = \hat{x}^{(1)}(k) - \hat{x}^{(1)}(k-1), \ k = 2, 3, \ldots, n.$$

2.3. Markov residual modified model

To improve the GM(1,1) model, this work uses the Markov-chain to modify the residual errors of GM(1,1).

The residual modification is established as follows (Hsu & Wen, 1998):

Step 1: Define the residual series $q^{(0)}$,

$$q^{(0)} = [q^{(0)}(2), q^{(0)}(3), \ldots, q^{(0)}(n)],$$

$$\text{where } q^{(0)}(k) = x^{(0)}(k) - \hat{x}_{\mathrm{o}}^{(0)}(k), \ k = 2, 3, \ldots, n.$$

Step 2: Denote the absolute values of the residual series as $\varepsilon^{(0)}$,

$$\varepsilon^{(0)} = [\varepsilon^{(0)}(2), \varepsilon^{(0)}(3), \ldots, \varepsilon^{(0)}(n)],$$

$$\text{where } \varepsilon^{(0)}(k) = |q^{(0)}(k)|, \ k = 2, 3, \ldots, n.$$

Step 3: A GM(1,1) model of $\varepsilon^{(0)}$ can be established:

$$\hat{\varepsilon}(k) = \left[\varepsilon^{(0)}(1) - \frac{u_\varepsilon}{a_\varepsilon}\right](1 - e^{a_\varepsilon})\, e^{-a_\varepsilon(k-1)}$$

where a_ε, u_ε is estimated using OLS.

Step 4: Assume that the sign of the kth data residual is in state 1 when it is positive, and in state 2 when it is negative. A one-step transition probability P is associated with each possible transition from state i to state j, and P can be estimated using $P_{ij} = M_{ij}/M_i$, $i = 1, 2$, $j = 1, 2$, where M_i denotes the number of years whose residuals are in state i, and M_{ij} is the number of transitions from state i to state j that have occurred. These P_{ij} values can be presented as a transition matrix R.

Step 5: These P values can be arranged as a transition matrix;

$$R = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix},$$

where R can be estimated by examining the signs of residual for all years from 1996 to 1999.

Step 6: Denote the initial state distribution by the vector: $\pi^{(0)} = [\pi_1^{(0)}, \pi_2^{(0)}]$, where

$$\begin{cases} \pi_1 \text{ is state } 1(+) \\ \pi_2 \text{ is state } 2(-) \end{cases}$$

probabilities. The state probabilities after n' transitions are given by $\pi^{(n')} = \pi^{(0)}\hat{R}^{n'}$, where $\pi^{(n')} = [\pi_1^{(n')}, \pi_2^{(n')}]$. $\pi^{(n')}$ is $(n+n')$th year residual state probabilities. Let the sign of the $(n+n')$th year residual be represented as follows:

$$\sigma(n+n') = \begin{cases} +1 & \text{if } \pi_1^{(n')} > \pi_2^{(0)} \\ -1 & \text{if } \pi_1^{(n')} < \pi_2^{(n')} \end{cases}, \quad n = 1, 2, \ldots$$

Step 7: An improved grey model with residual modification and Markov-chain sign estimation can be formulated. That namely

$$\hat{x}_r^{(0)}(k) = \hat{x}_o^{(0)}(k) + \sigma(k)\left[\varepsilon^{(0)}(1) - \frac{u_\varepsilon}{a_\varepsilon}\right](1 - e^{a_\varepsilon})\, e^{-a_\varepsilon(k-1)}$$

where $\hat{x}_r^{(0)}(1) = x^{(0)}(1)$, $\sigma(k) = \pm 1$;

a, u, a_ε and u_ε is parameter value, which is estimated using OLS.

3. Forecasting accuracy measurement

To examine the accuracy of the different models, this work compares the forecasting results of the value of 2001. One of the evaluation standards is used to test the accuracy of the forecast model, namely the relative percentage error (RPE), which compares the real value and forecast value to evaluate the short-term results. RPE is defined as $\text{RPE} = \lfloor |x^{(0)}(k) - \hat{x}^{(0)}(k)|/x^{(0)}(k) \rfloor 100\%$, where RPE is the absolute value of error rate, $x^{(0)}(k)$ is the actual value, and $\hat{x}^{(0)}(k)$ is the predicted value.

4. The empirical application using the sample of tourism arrivals to Taiwan

The data used were arrivals to Taiwan from the major markets of Hong Kong, United States and Germany from 1989 to 2000 (Table 1). Forecasting tourism arrivals to Taiwan used the fuzzy time series and grey theory. An evaluation of performance of forecast model used the error evaluation technique discussed above. The following subsection will show the detailed stages.

4.1. The empirical result of fuzzy time series

We herein follow the fuzzy time series theory then operate the mathematic steps given the raw data.

Step 1: The variation in tourism from Hong Kong data from 1991 to 2000 yield the maximum $D_{\max} = 41492$ and minimum $D_{\min} = -18260$. For simplicity, $D_1 = 4120$ and $D_2 = 4126$ are chosen; thus, the universe U is given by $U = [-22380, 45620]$.

Step 2: The universe U is divided into five intervals with equal length, according to the average distribution method. The parameters $u_1, u_2, \ldots, u_5$ are used for each interval:

$u_1 = [-22380, -8780]$, $u_2 = [-8780, 4820]$,
$u_3 = [4820, 18420]$, $u_4 = [18420, 32020]$, and
$u_5 = [32020, 45620]$

Step 3: Five linguistic values must be determined to define fuzzy sets on the universe U. Let A_1 = "decrease", A_2 = "slow decrease", A_3 = "gentle slow increase", A_4 = "slow increase", A_5 = "increase", are the possible values. Meanwhile, each fuzzy set is composed of the elements u and their corresponding memberships. Hence, all the fuzzy sets $A_i (i = 1, 2, \ldots, 5)$ are expressed as follows:

$$A_1 = \{1/u_1,\ 0.5/u_2,\ 0/u_3,\ 0/u_4,\ 0/u_5\},$$

$$A_2 = \{0.5/u_1,\ 1/u_2,\ 0.5/u_3,\ 0/u_4,\ 0/u_5\},$$

$$A_3 = \{0/u_1,\ 0.5/u_2,\ 1/u_3,\ 0.5/u_4,\ 0/u_5\},$$

$$A_4 = \{0/u_1,\ 0/u_2,\ 0.5/u_3,\ 1/u_4,\ 0.5/u_5\},$$

$$A_5 = \{0/u_1,\ 0/u_2,\ 0/u_3,\ 0.5/u_4,\ 1/u_5\}.$$

Step 4. Fuzzify the historical data. Consider for example, the value change of 18260 in 1992. The degree of 1992 belongs to fuzzy set A_1, meaning that the medium value decreases from 1991 to 1992 (reference Table 1).

Step 5: Make a fuzzy forecast of the value changed of tourism from Hong Kong in 2000. Different window

Table 1
Forecast result of fuzzy time series

Year	Hong Kong		United State		Germany	
	Actual value	$w=5$ forecast	Actual value	$w=3$ forecast	Actual value	$w=6$ forecast
1989	211,804		220,594		25,002	
1990	193,544		224,915		24,320	
1991	181,765		240,375		25,798	
1992	193,523		259,145		28,969	
1993	213,953		269,110		28,644	
1994	241,775		286,713	283,557	31,334	
1995	246,747		290,138	301,160	32,944	
1996	262,585	258,367	289,900	298,285	33,914	
1997	259,664	274,205	303,634	291,747	34,660	34,919
1998	279,905	264,484	308,407	318,081	35,343	35,665
1999	319,814	305,125	317,801	310,254	34,190	36,348
2000	361,308	351,834	359,533	332,248	34,829	33,860
2001*	392,552	400,128	339,390	386,580	33,716	35,834

*2001 is forecast value of out-of-sample.

base will result in the different forecasting values given in Hwang's model. In order to find the smallest error rate, we need put different w bases into Huarng (2001) model, in this work we take the $w=3$ and $t=2000$, for example. $F(2000)$ is calculated using: $O^4(2000)$, $C(2000)$, and $R(2000)$, as follows:

$$O^4(2000) = \begin{bmatrix} \text{fuzzy variation of 1998} \\ \text{fuzzy variation of 1997} \\ \text{fuzzy variation of 1996} \end{bmatrix} = \begin{bmatrix} A_4 \\ A_2 \\ A_3 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 0 & 0.5 & 1 & 0.5 \\ 0.5 & 1 & 0.5 & 0 & 0 \\ 0 & 0.5 & 1 & 0.5 & 0 \end{bmatrix}$$

$$C(2000) = [\text{fuzzy variation of 1999}] = [A_5] = \begin{bmatrix} 0 & 0.5 & 1 & 0.5 & 1 \end{bmatrix},$$

$$R(2000) = \begin{bmatrix} 0\times 0 & 0\times 0 & 0\times 0.5 & 0.5\times 1 & 1\times 0.5 \\ 0\times 0.5 & 0\times 1 & 0\times 0.5 & 0.5\times 0 & 1\times 0 \\ 0\times 0 & 0\times 0.5 & 0\times 1 & 0.5\times 0.5 & 1\times 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0.5 & 0.5 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.25 & 0 \end{bmatrix},$$

$$F(1996) = \begin{bmatrix} 0 & 0 & 0 & 0.5 & 0.5 & 0 & 0 & 0 \end{bmatrix}.$$

Step 6: The matrix $F(2000)$ has more than one entry of the maximum value. The maximum value is in u_4 and u_5. Consequently, the new forecast value for 2000 is the average of u_4 and u_5. The relevant calculation is $(25220+38820)/2 = 32020$.

Step 7: The forecast value 351834 in 2000 is the actual value 319814 for 1999 plus the forecast variation 32020 in 2000.

When the window bases are equal to 3,4,5, and 6, the average value of RPE of Hong Kong are 5.04%, 4.3%, 3.64%, and 4.54%, respectively; the average value of RPE of United States are 3.54%, 3.95%, 3.98%, and 4.25%, respectively; Furthermore, the average value of RPE of Germany are 3.82%, 2.85%, 2.7%, and 2.69%, respectively. Under the minimum of average RPE value rule, the adequate window base of Hong Kong is 5, United States, 3, and Germany, 6. Table 1 shows the forecast value of three countries using the above window bases.

4.2. The empirical result of grey model

The grey model GM(1,1) is a time series prediction model. It is not necessary to employ all the data from the original series to construct GM(1,1), but the potency of the series must be more than 4. The work takes the tourism flows from Hong Kong to be the calculating example.

Step 1: The original total tourism series are $X^{(0)} = \{211804,\ 193544,\ \ldots, 361308\}$.

Step 2: Based on the initial sequence $X^{(0)}$, a new sequence is generated by the AGO. $X^{(1)} = [21804,\ 405348,\ 587113,\ \ldots, 2966387]$.

Step 3: This work used $\alpha = 0.999$ to be the example and obtained the following result:

$$B = \begin{bmatrix} -211823 & 1 \\ -405366 & 1 \\ \vdots & \vdots \\ -2605115 & 1 \end{bmatrix}, \quad Y = \begin{bmatrix} 193544 \\ 181765 \\ \vdots \\ 361308 \end{bmatrix},$$

$$\hat{\theta} = (B^{\mathrm{T}} B)^{-1} B^{\mathrm{T}} Y = \begin{bmatrix} -0.068 \\ 161763.1 \end{bmatrix}.$$

Step 4: Taking $\hat{\theta}$ into equation

$$\hat{x}_{\mathrm{o}}^{(0)}(k) = \left[x^{(0)}(1) - \frac{u}{a}\right](1-\mathrm{e}^{a})\,\mathrm{e}^{-a(k-1)}.$$

The forecast equation of GM(1,1) is namely as below.

$$\hat{x}_{\mathrm{o}}^{(0)}(k) = \left[211804 - \frac{1617631}{(-0.068)}\right](1 - \mathrm{e}^{-0.068}) \times \mathrm{e}^{-(-0.068)(k-1)}, \quad k = 2, 3, \ldots, n, \ n+1 \ldots.$$

When $k = 11$, GM(1,1) generated the forecast value of tourism arrival Taiwan being 359825.

4.3. The empirical result of Markov modified model

We present herein an improved grey model-Markov model. The procedure of estimation is the same as grey model.

(1) Denote the residual series, the differences between the actual value and forecast value is obtained from the absolute values
$\varepsilon^{(0)} = [11250.59, \ 13355.54, \ldots, 1483.098]$.

(2) A new sequence generated by the AGO
$\varepsilon^{(1)} = [11250.59, \ 24606.13, \ldots, 158704.8]$.

(3) Using $\alpha = 0.5$ (Hsu & Wen, 1998), we get the following functions:

$$B = \begin{bmatrix} -5625.3 & 1 \\ -17928.4 & 1 \\ \vdots & \vdots \\ -157963 & 1 \end{bmatrix}, \quad Y = \begin{bmatrix} 11250.59 \\ 13355.54 \\ \vdots \\ 1483.098 \end{bmatrix},$$

$$\hat{\theta} = \begin{bmatrix} a_\varepsilon \\ u_\varepsilon \end{bmatrix} (B^{\mathrm{T}} B)^{-1} B^{\mathrm{T}} Y = \begin{bmatrix} -0.04167 \\ 11413.03 \end{bmatrix}.$$

(4) Incorporated a_ε and u_ε into the estimation function, we can achieve the Markov model for Hong Kong.

$$\hat{x}_{\mathrm{M}}^{(0)}(k) = \left[211804 - \frac{161763.1}{(-0.068)}\right](1 - \mathrm{e}^{-0.068}) \times \mathrm{e}^{-(-0.068)(k-1)} + \delta(k)\left[0 - \frac{11413.03}{(-0.04167)}\right] \times (1 - \mathrm{e}^{-0.04167})\mathrm{e}^{-(-0.04167)(k-1)}$$
$$k = 2, 3, \ldots, n, \ n+1 \ldots.$$

The empirical result of Markov model of tourists for Hong Kong arrival Taiwan is 377504 using $k = 11$.

Given the same procession, we are first present the GM(1,1) equation of tourists for United States (using $\alpha = 0.0001$)

$$\hat{x}_{\mathrm{o}}^{(0)}(k) = \left[220594 - \frac{217794.9}{(-0.03747)}\right](1 - \mathrm{e}^{-0.03747}) \times \mathrm{e}^{-(-0.03747)(k-1)}.$$

The Markov model of tourists for United States is presented as follows:

$$\hat{x}_{\mathrm{M}}^{(0)}(k) = \left[220594 - \frac{217794.9}{(-0.03747)}\right](1 - \mathrm{e}^{-0.03747}) \times \mathrm{e}^{-(-0.03747)(k-1)} + \delta(k)\left[0 - \frac{6432.819}{(-0.05586)}\right] \times (1 - \mathrm{e}^{-0.05586})\,\mathrm{e}^{-(-0.05586)(k-1)}$$
$$k = 2, 3, \ldots, n, \ n+1 \ldots.$$

Secondly, the GM(1,1) equation of tourists for Germany is presented as follows ($\alpha = 0.0001$):

$$\hat{x}_{\mathrm{o}}^{(0)}(k) = \left[25002 - \frac{24666.89}{(-0.03309)}\right](1 - \mathrm{e}^{-0.03309}) \times \mathrm{e}^{-(-0.03309)(k-1)}$$

The Markov model of tourists for Germany is namely as follows:

$$\hat{x}_{\mathrm{M}}^{(0)}(k) = \left[25002 - \frac{24666.89}{(-0.03309)}\right](1 - \mathrm{e}^{-0.03309}) \times \mathrm{e}^{-(-0.03309)(k-1)} + \delta(k)\left[0 - \frac{1331.857}{(-0.0139)}\right] \times (1 - \mathrm{e}^{-0.0139})\,\mathrm{e}^{-(-0.0139)(k-1)},$$
$$k = 2, 3, \ldots, n, \ n+1 \ldots.$$

Table 2 lists the forecast value of Hong Kong, United States and Germany using GM(1,1) model and Markov model.

4.4. Error measurement and analysis

Table 3 lists the 2001 error of three forecast models of 2001 in three countries using RPE method. These figures depict that the minimum error rates, which is calculated by GM(1,1) model, for Hong Kong and United States tourists are 1.88% and −2.492%, respectively. Meanwhile, the minimum error rate of German is −5.999% using the Markov improved model.

Figs. 1–3 illustrates the actual value, the forecast from the Markov model, GM(1,1) and the fuzzy time series among Hong Kong, United States and German markets, respectively. These figures show the trends of tourism demand from 1989 to 2001.

5. Conclusions

Concerning the tourism demand prediction research, multi-variant regression and ARIMA were the major prediction tools. The multi-variant regression models face the issues of how to decide independent variables in a regression model and whether multi-collinearily exists among the independent variables and of how to collect the large amount time series data required.

Table 2
GM(1,1) and Markov model

Year	Hong Kong		United State		Germany	
	GM (1,1)	Markov	GM (1,1)	Markov	GM (1,1)	Markov
1989	211,804	211,804	220,594	220,594	25,002	25,002
1990	182,293	193,948	230,349	223,733	25,921	24,580
1991	195,121	182,970	239,144	246,140	26,793	25,433
1992	208,850	196,183	248,275	255,673	27694	29,073
1993	223,546	210,340	257,754	265,577	28,626	30,024
1994	239,276	253,044	267,595	275,868	29,589	31,007
1995	256,113	241,759	277,812	286,560	30,584	32,022
1996	274,134	259,170	288,420	297,670	31,613	33,071
1997	293,424	277,822	299,432	309,213	32,677	34,155
1998	314,071	297,805	310,864	300,521	33,776	35,275
1999	336,170	319,213	322,733	311,796	34,913	33,393
2000	359,825	377,504	335,056	346,622	36,087	34,546
2001*	385,144	366,713	347,848	360,079	37,301	35,739

*2001 is out of sample.

Table 3
The analysis of error in 2001

Model	Country		
	Hong Kong (%)	United States (%)	Germany (%)
Fuzzy	1.93	13.9	6.28
GM(1,1)	1.887	−2.492	−10.634
Markov	6.582	−6.096	−5.999

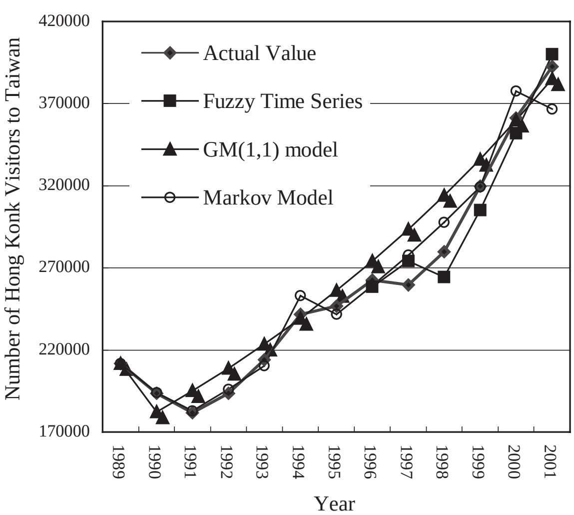

Fig. 1. Trend of actual value and forecast values in Hong Kong.

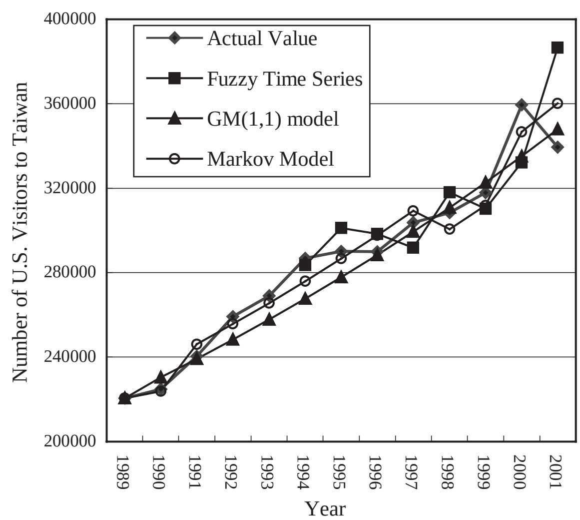

Fig. 2. Trend of actual value and forecast values in United States.

The ARIMA forecasting model uses historical time series data to construct the mathematical function. Furthermore, the ARIMA method has been widely applied to different fields and has achieved accuracy in forecast results in data possessing stable trend conditions.

Recently, in order to overcome the constraints of time series, scholars have developed the neural network methodology, which is a computation technique that can imitate a human brain. Unfortunately, the need to gather a large amount raw data for the learning process of the neural model is a serious problem in tourism demand prediction.

Having discussed the traditional forecasting techniques for tourism demand, the contribution of this work is to present an AI methods that does not make many assumptions when forming the forecasting model, namely that of grey theory and fuzzy time series to predict tourism demand arrival to Taiwan. The original GM(1,1) model was further improved by using the technique that combines residual modifications with Markov chain sign estimations.

The empirical results indicate that fuzzy time series is suitable for the tourism demand forecasting of Hong Kong arrival to Taiwan, the GM(1,1) model

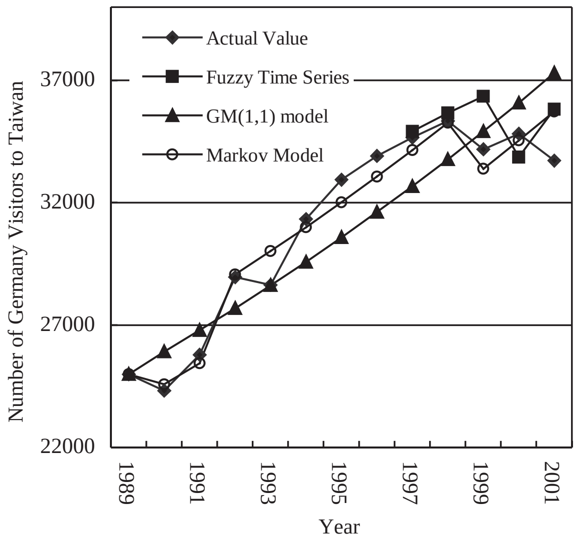

Fig. 3. Trend of actual value and forecast values in Germany.

appropriate the tourism demand forecasting for Hong Kong and United States arrival figure for Taiwan, and Markov-improved model is the best for German tourism demand estimation. From the prediction, tourism demand will continuously grow in the future. This implies a need for the Taiwanese government to plan public sightseeing physical facilities.

References

Chen, S.-M. (1996). Forecasting enrollment based on fuzzy time series. *Fuzzy Sets and Systems*, *81*, 311–319.

Deng, J.-L. (1982). Control problem of grey system. *Systems and Control Letters*, *1*, 288–294.

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

Hsu, L.-C. (2001). The comparison of three residual modification model. *Journal of the Chinese Grey System Association*, *4*(2), 97–110.

Hsu, L.-C. (2003). Applying the grey prediction model for the global integrated circuit industry. *Technological Forecasting and Social Change*, forthcoming.

Hsu, L.-C., & Wang, C.-H. (2002). Grey forecasting the financial ratios. *The Journal of Grey System*, *14*(4), 399–408.

Hsu, C.-I., & Wen, Y.-U. (1998). Improved grey prediction models for trans-pacific air passenger market. *Transportation Planning and Technology*, *22*, 87–107.

Huarng, K.-H. (2001). Effective lengths of intervals to improve forecasting in fuzzy time series. *Fuzzy Sets and Systems*, *123*, 387–394.

Hwang, J., Chen, S.-M., & Lee, C.-H. (1998). Handling forecasting problems using fuzzy time series. *Fuzzy Sets and Systems*, *100*, 217–228.

Law, R. (2000). Back-propagation learning in improving the accuracy of neural network-based tourism demand forecasting. *Tourism Management*, *21*, 331–340.

Law, R., & Au, N. (1999). A neural network model to forecast Japanese demand for travel to Hong Kong. *Tourism Management*, *20*, 89–97.

Lim, C., & McAleer, M. (2002). Time series forecasts of international travel demand for Australia. *Tourism Management*, *23*, 389–396.

Song, Q., & Chissom, B. S. (1993a). Forecasting enrollments with fuzzy time series—Part I. *Fuzzy Sets and Systems*, *54*, 1–9.

Song, Q., & Chissom, B. S. (1993b). Fuzzy time series and its models. *Fuzzy Sets and Systems*, *54*, 269–277.

Song, Q., & Chissom, B. S. (1994). Forecasting enrollments with fuzzy time series—Part II. *Fuzzy Sets and Systems*, *62*, 1–8.

Wen, J.-C., Huang, K.-H., & Wen, K.-L. (2000). The Study of α in GM(1,1) Model. *Journal of the Chinese Institute of Engineers*, *23*(5), 583–589.