Forecasting Inbound Tourism Demand in Thailand with Grey Model

YING-FANG HUANG[1], THANH-LAM NGUYEN[2,*], MING-HUNG SHU[3]
[1,3] Department of Industrial Engineering and Management
[2] Graduate Institute of Mechanical and Precision Engineering
National Kaohsiung University of Applied Sciences
415, Chien Kung Road, Sanmin District, Kaohsiung 80778
TAIWAN, R.O.C
(* Corresponding author: green4rest.vn@gmail.com)

Abstract: - Tourism has been considered as a "smokeless" and important industry in many countries in the world because it can not only generate plenty of quality jobs but also offer great contribution to the GDP. In order to make the tourism industry grow stably, getting to know its determinants becomes critical. However, it has been found in several studies that the determinants are neither easily measured nor collected due to their availability. Hence, until now, there is no standard forecasting model that can be applied in different situations of the industry. Moreover, due to certain limitations in collecting data of inbound tourism demand, it is therefore suggested in this study to use Grey forecasting approach which has been widely employed in different areas because it can deal with the problems of uncertainty with few data points and/or poor information. In the empirical case study of annual inbound tourism demand in Thailand, after being modified with Fourier series, the traditional Grey model GM(1,1) is significantly improved and becomes FGM(1,1) which has very low value of mean absolute percentage error (MAPE) of 0.0099. It is therefore suggested to forecast the inbound tourism demand in Thailand.

Key-Words: - GM(1,1), Fourier modification, FGM(1,1), Tourism demand, Grey model, Thailand tourism

1 Introduction

Nowadays, tourism has been considered as a "smokeless" and important industry in numerous countries in the world because it can not only generate plenty of quality jobs but also offer great contribution to the GDP. In Thailand, particularly, the number of international tourist arrivals in 2012 was about triple compared to that in 1998. Based on the annual research by the World travel & tourism council [1], in 2012, the total contribution of the tourism industry to Thailand GDP was accounting for 16.7% of GDP (about THB1,896.70 billion); and it supported 12.4% of the total employment with about 4.82 million jobs. In regarding to its direct contribution, the tourism contributed 7.3% of total GDP (about THB825.6 billion) and supported more than 2 million jobs (5.2% of total employment). These figures indicate that tourism is an important industry in Thailand. In order to make proper plans for the sustainable development of the national tourism industry, accurately forecasting the tourism demand becomes mandatory.

In spite of its aforementioned importance, tourism has been considered not only as an integrated and self-contained economic activity without a strong support from economic theories but also as a complex system due to a strong inter-relationship existing among different dependable sectors in the economy such as economic, transportation, commerce, social & cultural services, political and technological changes, etc., [2]. While González and Moral [3] pointed out that tourism demand is strongly affected by various factors, including tourism price, price index, income index, marketing expenditures, demographic and cultural factors, the quality-price ratio, etc., Witt & Witt [4] considered different determinants such as population, origin country income or private consumption, own price, substitute prices, one-off events, trend, etc. Also, some marketing aspects such as tour prices, distribution channel of the travel agents, traveler's income were also suggested [5]. However, numerous researchers have concluded that many of the determinants are neither easily measured nor collected due to their availability [2, 3, 4, 6].

Furthermore, there has been no standard measure to represent "inbound tourism demand". It was suggested that inbound tourism demand be measured in terms of the number of tourist arrival, tourist expenditure (tourist receipts) or the number of nights tourists spent [4, 7]. But, due to the complexity in collecting the data of tourist expenditure and the number of nights tourists spent, tourist arrival has been widely used as an appropriate indicator of inbound tourism demand in many researches [3, 6-15]. Therefore, in this study, the annual arrival of inbound tourists to Thailand from 1997 to 2012 is used to denote the inbound tourism demand in Thailand.

Due to the above limitations in collecting relevant data of inbound tourism demand in Thailand, it is therefore suggested to use Grey forecasting approach which has been widely employed in different areas due to its ability to deal with the problems of uncertainty with few data points and/or "partial known, partial unknown" information, to predict the demand. In order to improve the accuracy of the model, its residuals series is then modified with Fourier series to form a new one. Forecasting result of the same period is then used to compare the accuracy of this modified model and ARFIMA-FIGARCH proposed by Chokethaworn et al. [16]. Finally, the best model is selected to forecast the international tourism demand in Thailand.

2 Literature Review

2.1 Grey Model

Grey theory offers a new approach to deal mainly with the problems of uncertainty with few data points and/or poor information which is said to be "partial known, partial unknown" [17, 18]. The core of the theory is the grey dynamics model which is usually called Grey model (GM). The Grey model is used to execute the short-term forecasting operation with no strict hypothesis for the distribution of the original data series [19]. The general GM model has the form of GM(d,v), where d is the rank of differential equation and v is the number of variables appeared in the equation. The basic model of Grey model is GM(1,1), a first-order differential model with one input variable which has been successfully applied in many different researches. It is obtained based on the following procedure.

Step 1: Suppose an original series with n entries is $x^{(0)}$:

$$x^{(0)} = \left\{x_1^{(0)}, x_2^{(0)}, \ldots, x_k^{(0)}, \ldots x_n^{(0)}\right\} \tag{1}$$

where $x_k^{(0)}$ is the value at time k $\left(k = \overline{1,n}\right)$.

Step 2: From the original series $x^{(0)}$, a new series $x^{(1)}$ can be generated by one time accumulated generating operation (1-AGO), which is

$$x^{(1)} = \left\{x_1^{(1)}, x_2^{(1)}, \ldots, x_k^{(1)}, \ldots x_n^{(1)}\right\} \tag{2}$$

where $x_k^{(1)} = \sum_{j=1}^{k} x_j^{(0)}$

Step 3: A first-order differential equation with one variable is expressed as:

$$\frac{dx^{(1)}}{dt} + ax^{(1)} = b \tag{3}$$

where a is called developing coefficient and b is called grey input coefficient. These two coefficients can be determined by the least square method as the following:

$$\left[a, b\right]^T = \left(B^T B\right)^{-1} B^T Y \tag{4}$$

where

$$B = \begin{bmatrix} -\left(x_1^{(1)} + x_2^{(1)}\right)/2 & 1 \\ -\left(x_2^{(1)} + x_3^{(1)}\right)/2 & 1 \\ \cdots & \cdots \\ -\left(x_{n-1}^{(1)} + x_n^{(1)}\right)/2 & 1 \end{bmatrix} \qquad Y = \begin{bmatrix} x_2^{(0)} \\ x_3^{(0)} \\ \cdots \\ x_n^{(0)} \end{bmatrix}$$

Therefore, the forecasting equation for GM(1,1) is expressed as:

$$\hat{x}_k^{(1)} = \left[x_1^{(0)} - \frac{b}{a}\right] e^{-a(k-1)} + \frac{b}{a} \qquad \left(k = \overline{1,n}\right) \tag{5}$$

Based on the operation of one time inverse accumulated generating operation (1-IAGO), the predicted series $\hat{x}^{(0)}$ can be obtained as following:

$$\hat{x}^{(0)} = \left\{\hat{x}_1^{(0)}, \hat{x}_2^{(0)}, \ldots, \hat{x}_k^{(0)}, \ldots, \hat{x}_n^{(0)}\right\} \tag{6}$$

where

$$\begin{cases} \hat{x}_1^{(0)} = \hat{x}_1^{(1)} \\ \hat{x}_k^{(0)} = \hat{x}_k^{(1)} - \hat{x}_{k-1}^{(1)} & \left(k = \overline{2,n}\right) \end{cases}$$

2.2 Fourier Residual Modification

GM(1,1) model can perform better if it is modified with Fourier series [20-24]. The procedure to obtain the modified model (hereafter called FGM(1,1)) is as the following.

Based on the predicted series $\hat{x}^{(0)}$ obtained from the GM(1,1) model, a residual series named $\varepsilon^{(0)}$ is defined as:

$$\varepsilon^{(0)} = \left\{\varepsilon_2^{(0)}, \varepsilon_3^{(0)}, \ldots, \varepsilon_k^{(0)}, \ldots, \varepsilon_n^{(0)}\right\} \tag{7}$$

where $\varepsilon_k^{(0)} = x_k^{(0)} - \hat{x}_k^{(0)}$ $\qquad \left(k = \overline{2,n}\right)$

Expressed in Fourier series, $\varepsilon_k^{(0)}$ is rewritten as:

$$\varepsilon_k^{(0)} = \frac{1}{2}a_0 + \sum_{i=1}^{F}\left[a_i \cos\left(\frac{2\pi i}{n-1}k\right) + b_i \sin\left(\frac{2\pi i}{n-1}k\right)\right] \tag{8}$$

where $F = \left[(n-1)/2 - 1 \right]$ called the minimum deployment frequency of Fourier series [24] and only take integer number [20, 21, 23]. And therefore, the residual series is rewritten as:

$$\varepsilon^{(0)} = P.C \tag{9}$$

where

$$P = \begin{bmatrix} \frac{1}{2} & \cos\left(\frac{2\pi\times 1}{n-1}\times 2\right) & \sin\left(\frac{2\pi\times 1}{n-1}\times 2\right) & \cdots & \cos\left(\frac{2\pi\times F}{n-1}\times 2\right) & \sin\left(\frac{2\pi\times F}{n-1}\times 2\right) \\ \frac{1}{2} & \cos\left(\frac{2\pi\times 1}{n-1}\times 3\right) & \sin\left(\frac{2\pi\times 1}{n-1}\times 3\right) & \cdots & \cos\left(\frac{2\pi\times F}{n-1}\times 3\right) & \sin\left(\frac{2\pi\times F}{n-1}\times 3\right) \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \frac{1}{2} & \cos\left(\frac{2\pi\times 1}{n-1}\times n\right) & \sin\left(\frac{2\pi\times 1}{n-1}\times n\right) & \cdots & \cos\left(\frac{2\pi\times F}{n-1}\times n\right) & \sin\left(\frac{2\pi\times F}{n-1}\times n\right) \end{bmatrix}$$

$$C = \left[a_0, a_1, b_1, a_2, b_2, \ldots, a_F, b_F \right]^T$$

The parameters $a_0, a_1, b_1, a_2, b_2, \ldots, a_F, b_F$ are obtained by using the ordinary least squares method (OLS) which results in the equation of:

$$C = \left(P^T P \right)^{-1} P^T \left[\varepsilon^{(0)} \right]^T \tag{10}$$

Once the parameters are calculated, the predicted series residual $\hat{\varepsilon}^{(0)}$ is then easily achieved based on the following expression:

$$\hat{\varepsilon}_k^{(0)} = \frac{1}{2} a_0 + \sum_{i=1}^{F} \left[a_i \cos\left(\frac{2\pi i}{n-1} k \right) + b_i \sin\left(\frac{2\pi i}{n-1} k \right) \right] \tag{11}$$

Therefore, based the predicted series $\hat{x}^{(0)}$ obtained from GM(1,1), the predicted series $\tilde{x}^{(0)}$ of the FGM(1,1) is determined by:

$$\tilde{x}^{(0)} = \left\{ \tilde{x}_1^{(0)}, \tilde{x}_2^{(0)}, \ldots, \tilde{x}_k^{(0)}, \ldots, \tilde{x}_n^{(0)} \right\} \tag{12}$$

where

$$\begin{cases} \tilde{x}_1^{(0)} = \hat{x}_1^{(0)} \\ \tilde{x}_k^{(0)} = \hat{x}_k^{(0)} + \hat{\varepsilon}_k^{(0)} \qquad \left(k = \overline{2,n} \right) \end{cases}$$

In order to evaluate the accuracy of the forecasting model, the residual error (ε) and its relative error ($\wp$) are used [21, 25]. ε and $\wp$ of an entry k are expressed as:

- Residual error: $\varepsilon_k = x_k^{(0)} - f_k^{(0)} \qquad \left(k = \overline{1,n} \right)$

  where $f_k^{(0)}$ is the forecasted value at the k^{th} entry

- Relative error: $\wp_k = \left| \varepsilon_k \right| / x_k^{(0)} \qquad \left(k = \overline{1,n} \right)$

However, there have been some other important indexes to be considered in evaluating the model accuracy. They are:

- The mean absolute percentage error (MAPE) [5, 17, 21, 22, 23, 24, 26, 27] (MAPE is also known as the average relative error δ [28, 29]):

$$MAPE = \frac{1}{n} \sum_{k=1}^{n} \wp_k$$

- The post-error ratio C [28, 29]:

$$C = \frac{S_2}{S_1}$$

where: $S_1 = \sqrt{ \frac{1}{n} \sum_{k=1}^{n} \left[x_k^{(0)} - \frac{1}{n} \sum_{k=1}^{n} x_k^{(0)} \right]^2 }$

$$S_2 = \sqrt{ \frac{1}{n} \sum_{k=1}^{n} \left[\varepsilon_k - \frac{1}{n} \sum_{k=1}^{n} \varepsilon_k \right]^2 }$$

The ratio C, in fact, is the ratio between the standard deviation of the original series and the standard deviation of the forecasting error. The smaller the C value, the higher accuracy the model; since smaller C value results from a larger S_1 and/or a smaller S_2.

- The small error probability P [28, 29]:

$$P = p\left\{ \left(\left| \varepsilon_k - \frac{1}{n} \sum_{k=1}^{n} \varepsilon_k \right| / S_1 \right) < 0.6745 \right\}$$

The higher the P value is, the higher accuracy the model has since P value indicates the probability of the ratio of the difference between the residual values of data points and the average residual value with the standard deviation of the original series smaller than 0.6745 [29].

- The forecasting accuracy ρ [29]: $\rho = 1 - MAPE$.

Based on the above indexes, there are four grades of accuracy as stated in Table 1.

Table 1: Four grades of forecasting accuracy

Grade level	$MAPE$	C	P	ρ
I (Very good)	< 0.01	< 0.35	> 0.95	> 0.95
II (Good)	< 0.05	< 0.50	> 0.80	> 0.90
III (Qualified)	< 0.10	< 0.65	> 0.70	> 0.85
IV (Unqualified)	≥ 0.10	≥ 0.65	≤ 0.70	≤ 0.85

3 Empirical Study

Historical data of the inbound tourism demand in Thailand from 1997 – 2012 (totally 16 observations) are obtained from the annual statistical data published by The Office of Tourism Development [30]. Based on the data from 1997-2011, GM(1,1) model for the annual inbound tourism demand in Thailand is found as:

$$\hat{x}_k^{(1)} = 132.63804 e^{0.05941(k-1)} - 125.4180$$

From GM(1,1) model, its residual series can be easily obtained, which is then modified with Fourier series to become a new model named FGM(1,1). The relevant indexes to evaluate the accuracy of the GM(1,1) and FGM(1,1) are shown in Table 2.

From Table 2, with the MAPE value of less than 1%, FGM(1,1) outperforms GM(1,1). Hence, it is strongly suggested to forecast the inbound tourism demand in Thailand.

Using FGM(1,1), the forecast value of the inbound tourism demand in Thailand in 2009 is 14.21 million arrivals. In comparison to the actual arrival of 2009, our forecast model results in an absolute percentage error of 0.424% which is extremely lower than that of the model proposed by Chokethaworn [16].

Table 2: Evaluation indexes of model accuracy

Index \ Model	GM(1,1)	FGM(1,1)
MAPE	0.0545	0.0099
S_1	15.9646	15.9646
S_2	1.2872	0.2192
C	0.0806	0.0137
P	1.0000	1.0000
ρ	0.9455	0.9901
Forecasting power	Qualified	Very good

In order to further evaluate the accuracy of our proposed model, we now compare the forecast value in 2012 with the actual observation in the same year. It is found that our model suggests 20.70 million which is 7.17% lower than the actual. The tolerance of 7.17% is acceptable in a forecasting model. So, the inbound tourism demand in Thailand in 2013 is forecasted to be 23.87 million arrivals.

4 Conclusion

Combining the Fourier residual modification to the traditional Grey forecasting model GM(1,1) results in an highly accurate forecasting model named FGM(1,1). Particularly, in the case of the inbound tourism demand in Thailand, despite of the hard assessment of relevant data about its determinants, the annual demand can still be forecasted effectively with FGM(1,1) model which has been proved to be a good choice. Precise forecasting result will not only help related organizations avoid unnecessary and unexpected costs but also assist the policy-makers from macro to micro levels in the tourism and tourism-related industries with a robust base in improving their related policies to stably develop the smokeless industry.

References:

[1] World Travel and Tourism Council, 2013 Annual Research: Key Facts. *Available from: http://www.wttc.org/research/economic-impact-research/country-reports/t/thailand/* [Access: 4 May, 2013]

[2] Lagos, D.G, Exploratory forecasting methodologies for tourism demand, *Ekistics Journal*, Vol. 396, 1999, pp. 143-154.

[3] González, P. and Moral, P., An analysis of the international tourism demand in Spain, *International Journal of Forecasting*, Vol. 11, 1995, pp. 233-251.

[4] Witt, S.F., and Witt, C.A., Forecasting tourism demand: A review of empirical research, *International Journal of Forecasting*, Vol. 11, 1995, pp. 447-475.

[5] Hsu, L.C., and Wang, C.H., The development and testing of a modified Diffusion model for predicting tourism demand, *International Journal of Management*, Vol. 25, 2008, pp. 439-445.

[6] Habibi, F., Rahim, K.A., Ramchandran, S., and Chin, L., Dynamic Model for International Tourism Demand for Malaysia: Panel Data Evidence, *International Research Journal of Finance and Economics*, Vol. 33, 2009, pp. 207-217.

[7] Ouerfelli, C., Co-integration analysis of quarterly European tourism demand in Tunisia, *Tourism Management*, Vol. 29, 2008, pp. 127-137.

[8] Morley, C.L., A dynamic international demand model, *Annals of Tourism Research*, Vol. 23, 1998, pp. 70-84.

[9] Lim, C., and McAleer, M., Monthly seasonal variations: Asian tourism to Australia, *Annals of Tourism Research*, Vol. 28, 2001, pp. 68–82.

[10] Kulendran, N., and Witt, S., Co-integration versus least squares regression, *Annals of Tourism Research*, Vol. 28, 2001, pp. 291–311.

[11] Tan, Y. F., McCahon, C., and Miller, J., Modelling tourist flows to Indonesia and Malaysia, *Journal of Travel and Tourism Marketing*, Vol. 12, 2002, pp. 63–84.

[12] Song, H. and Witt, S.F., General-to-specific modeling to international tourism demand forecasting, *Journal of Travel Research*, Vol. 42, 2003, pp. 65-74.

[13] Song, H., Witt, S. F., and Li, G., Modelling and forecasting the demand for Thai tourism, *Tourism Economics*, Vol. 9, 2003, pp. 363–387.

[14] Dritsakis, N., Co-integration analysis of German and British tourism demand for Greece, *Tourism Management*, Vol. 25, 2004, pp. 111-119.

[15] Naude, A.W., and Saayaman, A., Determinants of tourist arrivals in Africa: a panel data regression analysis, *Tourism Economics*, Vol. 11, 2005, pp. 365-391.

[16] Chokethaworn, K., Sriwichailamphan, T., Sriboonchitta, S., Chaiboonsri, C., Sriboonjit, J., and Chaitip, P., International tourist arrivals in Thailand: Forecasting with ARFIMA-FIGARCH approach, *Annals of the University*

of Petroşani, Economics, Vol. 10, 2010, pp. 75-84.

[17] Hsu, C.C., and Chen, C.Y., Applications of improved grey prediction model for power demand forecasting, *Energy Conversion and Management*, Vol. 44, 2003, pp. 2241-2249.

[18] Liu, S.F., Advances in Grey system theory and its applications in *Proceedings of 2007 IEEE International Conference on Grey Systems and Intelligent Services in Nanjing*, China, 2007.

[19] Wang, S.J., Wang, W.L, Huang, C.T., and Chen, S.C., Improving inventory effectiveness in RFID-enabled global supply chain with Grey forecasting model, *Journal of Strategic Information Systems*, Vol. 20, 2011, pp. 307-322.

[20] Hsu, L.C., Applying the grey prediction model to the global integrated circuit industry, *Technological forecasting & Social change*, Vol. 70, 2003, pp. 563-574.

[21] Guo, Z., Song, X. and Ye, J., A Verhulst model on time series error corrected for port throughput forecasting, *Journal of the Eastern Asia society for Transportation studies*, Vol. 6, 2005, pp. 881-891.

[22] Kan, M.L., Lee, Y.B. and Chen, W.C., Apply grey prediction in the number of Tourist, in *The fourth international conference on Genetic and Evolutionary computing in Shenzhen, China*, 2010, pp. 481-484.

[23] Askari, M., and Fetanat, A., Long-term load forecasting in power system: Grey system prediction-based models, *Journal of Applied Sciences*, Vol. 11, 2011, pp. 3034-3038.

[24] Huang, Y.L. and Lee, Y.H, Accurately forecasting model for the Stochastic Volatility data in tourism demand, *Modern economy*, Vol. 2, 2011, pp. 823-829.

[25] Liang, P., Zhou, H., and Zheng, J., Forecast research of droplet size based on Grey theory, *IFIP Advances in Information and Communication Technology*, Vol. 258, 2008, pp. 653-658.

[26] Chang, Y.W. and Liao, M.Y., A seasonal ARIMA model of tourism forecasting: The case of Taiwan, *Asia Pacific journal of Tourism research*, Vol. 15, 2010, pp. 215-221.

[27] Tsaur, R.C., and Kuo, T.C, The adaptive fuzzy time series model with an application to Taiwan's tourism demand, *Expert systems with Applications*, Vol. 38, 2011, pp. 9164-9171.

[28] Hua, J., and Liang, Q., Application research of the Grey forecast in the Logistics demand forecast, *First International Workshop on Education Technology and Computer Science*, Vol. 3, 2009, pp. 361-363.

[29] Ma, H., and Zhang, Z., Grey prediction with Markov-Chain for Crude oil production and consumption in China, *Advances in Intelligent and Soft Computing*, Vol. 56, 2009, pp. 551-561.

[30] Office of Tourism Development, Ministry of Tourism and Sports of Thailand (2013), http://www.tourism.go.th