Expert Systems with Applications 36 (2009) 8597–8603

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

Forecasting the turning time of stock market based on Markov–Fourier grey model

Yen-Tseng Hsu [a], Ming-Chung Liu [b], Jerome Yeh [a,*], Hui-Fen Hung [b]

[a] Department of Computer Science and Information Engineering, National Taiwan University of Science and Technology, Taipei, Taiwan, ROC
[b] Department of Electrical Engineering, National Taiwan University of Science and Technology, Taipei, Taiwan, ROC

ARTICLE INFO

Keywords:
Grey model
Forecast
Stock market

ABSTRACT

This paper presents an integration prediction method including grey model (GM), Fourier series, and Markov state transition, known as Markov–Fourier grey model (MFGM), to predict the turning time of Taiwan weighted stock index (TAIEX) for increasing the forecasting accuracy. There are two parts of forecast. The first one is to build an optimal grey model from a series of data, the other uses the Fourier series to refine the residuals produced by the mentioned model. Finally, the Markov state scheme is used for predicting the possibility of location results to promote the intermediate results generated by the Fourier grey model (FGM). It is evident that the proposed approach gets the better result performance than that of the other methods.

1. Introduction

In general, people are highly interested in forecasting future tendency of some events, such as investment in stock market, which is necessary to be forecasted for obtaining higher profit and reducing the investment risk. Since the prediction is mainly used to reduce the uncertainty or risk in making decision, therefore prediction accuracy is crucial. For the consideration of the long-term prediction, the most commonly used forecasting method are useful for prediction of periodical data, but not good enough for the non-linear data including the statistical regression, ARMA, etc. Therefore, many researchers proposed some AI-based methods, such as neural networks (Baba & Kozaki, 1992; Chi, Chen, & Chend, 1999; Hsu, Cheng, & Wu, 1997; Jang, Lai, Wang, & Parng, 1993), fuzzy systems (Hsu, Lin, & Cheng, 1999; Tseng & Tzeng, 2002), neural-fuzzy paradigms (Chang & Su, 1993; Jang, 1991; Lin & Lee, 1991), and grey systems (Lin, Su, & Hsu, 2001; Lin & Man, 1989; Liu, Yeh, Chen, & Hsu, 2004; Su, Lin, & Hsu, 2002) to resolve this problem.

In recent years, in spite of the grey models (Lin et al., 2001; Lin & Man, 1989) and/or its variants (Cheng, Hsu, & Wu, 1997; Tan & Chang, 1996; Tan & Lu, 1996; Wang, 2002) can obtain good performance on the various prediction applications, but for acquiring the best forecasting results, an effective method, based on the grey model, Fourier series, and Markov state transition matrices, termed as Markov–Fourier grey Model is proposed in our research.

The rest of this paper is organized as follows. In Section 2, the grey theory is described. In Section 3, an approach that adopts Fourier series to increase the forecast accuracy rate of grey model is discussed. The Markov state transition matrices is briefly illustrated in Section 4. Section 5 shows the Markov–Fourier grey prediction system. Experimental results and comparisons are shown in Section 6. Finally, the conclusions are discussed in Section 7.

2. Grey theory

The grey theory that was first proposed by Professor Deng (1982, 1989) for dealing with those systems with poor information. Recently, the two hot topics of grey theory, grey relational analysis (GRA) and grey model GM(1,1) prediction, have been successfully employed in various fields and has proven good performance (Hsu, Chen, & Lin, 2000; Hsu, Yeh, & Chang, 2000). Therefore, in this section, we illustrate the basic function of the grey system including grey relational analysis, grey model and grey metabolizing checking.

2.1. Grey relational analysis

Grey relational analysis is an effective method for analyzing the relationship among different series function or pattern. The essence of this method is to analyze the similarity or difference of the geometric shapes of different curves (Deng, 1988). The algorithm of grey relational analysis is illustrated as follows:

Step 1: Calculate grey relational gradesLet x_0 be the referential series with k entities,

$$x_0 = (x_0(1), x_0(2), \dots, x_0(k), \dots, x_0(n)) \tag{1}$$

and x_i the compared series,

$$x_i = (x_i(1), x_i(2), \dots, x_i(k), \dots, x_i(n)) \quad \text{for } i = 1, 2, \dots, m \tag{2}$$

* Corresponding author. Tel.: +886 2 2733 3141/7299; fax: +886 2 27301081.
E-mail address: jerome@mail.ntust.edu.tw (J. Yeh).

0957-4174/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2008.10.075

The grey relational coefficients between the referential series x_0 and the compared series x_i at the kth point is defined as follows:

$$\gamma_{0i}(k) = \frac{\Delta_{\min} + \zeta\Delta_{\max}}{\Delta_{0i}(k) + \zeta\Delta_{\max}} \tag{3}$$

where $\Delta_{0i}(k)$ is the absolute value of difference between x_0 and x_i at the kth point, that is $\Delta_{0i}(k) = \| x_0(k) - x_i(k)\|$, and $\Delta_{\min} = \min_i \min_k \Delta_{0i}(k)$, $\Delta_{\max} = \max_i \max_k \Delta_{0i}(k)$, $\zeta \in [0,1]$ is the distinguishing coefficient, typically taken as 0.5.

The grey relational grade Γ_{0i}(GRG) for series x_0 to x_i in terms of weight w_k is given as follows:

$$\Gamma_{0i} = \sum_{k=1}^{n} w_k \gamma_{0i}(k) \tag{4}$$

where w_k is the kth weight of γ_{0i}. If it is not necessary to care the weight, then the weight can be taken as $w_k = (1/n)$ for averaging.

Step 2: Grey relational generationBefore calculating the grey relational coefficients, these data series can be treated by the following three kinds of situations so that the linearity of normalization may be maintained (Hsia & Wu, 1997).There are:

(a) Upper bound effectiveness measurement

$$x_i^*(k) = \frac{x_i^{(0)}(k) - \min_k x_i^{(0)}(k)}{\max_k x_i^{(0)}(k) - \min_k x_i^{(0)}(k)} \tag{5}$$

where $\min_k x_i^{(0)}(k)$ is the minimum value of entity k and $\max_k x_i^{(0)}(k)$ is the maximum value of entity k.

(b) Lower bound effectiveness measurement

$$x_i^*(k) = \frac{\max_k x_i^{(0)}(k) - x_i^{(0)}(k)}{\max_k x_i^{(0)}(k) - \min_k x_i^{(0)}(k)} \tag{6}$$

(c) Moderate effectiveness measurement

$$x_i^*(k) = 1 - \frac{|x_i^0(k) - x_{ob}(k)|}{\max\{\max_k x_i^0(k) - x_{ob}(k) x_{ob}(k) - \min_k x_i^0(k)\}} \tag{7}$$

Step 3: Ordering the concerned series according to grey relational grade.From the ordering arrangement the main-factor and subfactor can be identified.

In this paper, the method of GRA is used to find out the important technical indices from various MAP (moving average of price) for constructing the GM(1,1) model to predict the turning time of TAIEX.

2.2. Grey models

Using less data, usually 3–5 points, the grey model can be constructed. The first step of grey model GM(1,1) accumulates the selected data so that the preprocessed data sequence may be more regular than original one. Then the grey differential equation is constructed to predict the regulated data. According to the solution of the grey differential equation, the prediction value for the regulated data can be obtained through a given prediction step. As we obtain the prediction value, the inverse accumulated operation is applied to get the prediction value of the original data. A brief description of the procedure of GM(1,1) is given as follows.

Assume that $x^{(0)}$ stands for the raw data series of turning time of TAIEX 24MAP highs/lows, namely,

$$x^{(0)} = (x^{(0)}(1), x^{(0)}(2), x^{(0)}(3), \ldots, x^{(0)}(n)) \tag{8}$$

where n is the sample size. By 1-AGO (one time Accumulated Generating Operation) $x^{(0)}$, the preprocessed series, $x^{(1)}$

$$x^{(1)} = (x^{(1)}(1), x^{(1)}(2), x^{(1)}(3), \ldots, x^{(1)}(n)) \tag{9}$$

where $x^{(1)}(k) = \sum_{m=1}^{k} x^{(0)}(m), \text{for} k = 1, 2, \ldots, n$.

By mean operation on $x^{(1)}$, the series $z^{(1)}$

$$z^{(1)} = (z^{(1)}(1), z^{(1)}(2), z^{(1)}(3), \ldots, z^{(1)}(n)) \tag{10}$$

where $z^{(1)}(k) = 0.5x^{(1)}(k) + 0.5x^{(1)}(k-1)$

Thus, from grey system theory (Deng, 1988) the grey differential equation of GM(1,1) and its whitening equation are obtained, respectively, as follows:

$$x^{(0)}(k) + az^{(1)}(k) = b, \quad k = 2, 3, \ldots, n$$

$$\frac{dx^{(1)}}{dt} + ax^{(1)} = b \tag{11}$$

where a and b is the developing coefficient and grey input, respectively. Let $\hat{\theta}$ be the parameters vector. By using least square method (Hsia, 1979), the parameters a and b can be obtain as

$$\hat{\theta} = (X^T X)^{-1} X^T Y = \begin{bmatrix} a \\ b \end{bmatrix} \tag{12}$$

where

$$X = \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \vdots & \vdots \\ -z^{(1)}(n) & 1 \end{bmatrix} \quad \text{and} \quad Y = \begin{bmatrix} X^{(1)}(2) \\ X^{(1)}(3) \\ \vdots \\ X^{(1)}(n) \end{bmatrix} \tag{13}$$

and X denotes the accumulated matrix and Y represents the constant vector. The approximate relation can be obtained by substituting the $\hat{\theta}$ into the differential equation, and solving equation (11):

$$\hat{x}^{(1)}(k+1) = \left[x^{(0)}(1) - \frac{b}{a}\right] e^{-ak} + \frac{b}{a} \tag{14}$$

where $\hat{x}^{(0)}(1) = x^{(0)}(1)$. By IAGO (inverse AGO) Eq. (14) the recovered value $\hat{x}^{(0)}(k)$ is:

$$\hat{x}^{(0)}(k) = \hat{x}^{(1)}(k) - \hat{x}^{(1)}(k-1) = (1 - e^{a}) \left[x^{(0)}(1) - \frac{b}{a}\right] e^{-a(k-1)} \tag{15}$$

Given $k = 1, 2, \ldots, n$, the predictive value is

$$\hat{x}^{(0)} = (\hat{x}^{(0)}(1), \hat{x}^{(0)}(2), \hat{x}^{(0)}(3), \ldots, \hat{x}^{(0)}(n)) \tag{16}$$

Finally, the forecasting data is further examined to see if it meets the residual error checking procedure. Usually, the following equation is utilized (Deng, 1988), where the residual error $e(k)$ is between the actual $x^{(0)}(k)$and the predicted $\hat{x}^{(0)}(k)$.

$$e(k) = \left| \frac{x^{(0)}(k) - \hat{x}^{(0)}(k)}{x^{(0)}(k)} \right| \times 100\%, \quad k = 2, 3, 4, \ldots, n \tag{17}$$

2.3. Grey metabolizing checking

Let $x^{(0)}$ be raw series of turning time of TAIEX 24MAP highs/lows,

$$x^{(0)} = (x^{(0)}(1), x^{(0)}(2), x^{(0)}(3), \ldots, x^{(0)}(n)) \tag{18}$$

where $x^{(0)}(k) \in \{x^{(0)}(\text{m}) — \text{m} = 1, \ldots, n\}$ and $x_i^{(0)}$ stands for the ith metabolizing series of $x^{(0)}$, namely,

$$x_i^{(0)} = (x^{(0)}(1_i), x^{(0)}(2_i), x^{(0)}(3_i), \ldots, x^{(0)}(n_i)) \tag{19}$$

$x_{i+1}^{(0)}$ Stands for the $i+1$th metabolizing series of $x^{(0)}$

$$x_{i+1}^{(0)} = (x^{(0)}(1_{i+1}), x^{(0)}(2_{i+1}), x^{(0)}(3_{i+1}), \ldots, x^{(0)}(n_{i+1})) \tag{20}$$

when

$$x^{(0)}(1_{i+1}) = x^{(0)}(2_i)$$
$$\text{Pot} K = \text{const}$$
$$K = \{1, 2, \ldots, \text{n}_\text{i}\} \quad \forall i \in I$$

where I and K are the metabolizing index and time index set of $x^{(0)}$, respectively, then $x_i^{(0)} \Rightarrow x_{i+1}^{(0)}\ \forall i \in I$ is so-called metabolizing, and $\{x_i^{(0)} | i \in I\} = \{x_i^{(0)}\}$ is the metabolizing subset of $x^{(0)}$.

By GM(1,1) modeling $x_i^{(0)}$ as follows:

$$\mathrm{IAGO}_o\ \mathrm{GM}_o\ \mathrm{AGO}: x_i^{(0)} \rightarrow \hat{x}^{(0)}(n_i+1) \quad n_i+1 = i+j, \quad 4 \leqslant j \leqslant n-1 \tag{21}$$

where $\hat{x}^{(0)}(n_i+1)$ is the predictive data. By calculating the residual error $q^{(0)}(i)$ which is between the actual $x^{(0)}(n_i+1)$ and the predicted $\hat{x}^{(0)}(n_i+1)$, we can check if the GM(1,1) metabolizing modeling is good enough or not

$$q^{(0)}(i) = \frac{x^{(0)}(i+j) - \hat{x}^{(0)}(i+j)}{x^{(0)}(i+j)} \times 100\% \tag{22}$$

$$P = \begin{bmatrix} \frac{1}{2} & \cos\left(\frac{2\pi\cdot 2}{T}\right) & \sin\left(\frac{2\pi\cdot 2}{T}\right) & \cos\left(\frac{2\cdot 2\pi\cdot 2}{T}\right) & \sin\left(\frac{2\cdot 2\pi\cdot 2}{T}\right)\cdots & \cos\left(\frac{k_a\cdot 2\pi\cdot 2}{T}\right) & \sin\left(\frac{k_a\cdot 2\pi\cdot 2}{T}\right) \\ \frac{1}{2} & \cos\left(\frac{2\pi\cdot 3}{T}\right) & \sin\left(\frac{2\pi\cdot 3}{T}\right) & \cos\left(\frac{2\cdot 2\pi\cdot 3}{T}\right) & \sin\left(\frac{2\cdot 2\pi\cdot 3}{T}\right)\cdots & \cos\left(\frac{k_a\cdot 2\pi\cdot 3}{T}\right) & \sin\left(\frac{k_a\cdot 2\pi\cdot 3}{T}\right) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{1}{2} & \cos\left(\frac{2\pi\cdot n}{T}\right) & \sin\left(\frac{2\pi\cdot n}{T}\right) & \cos\left(\frac{2\cdot 2\pi\cdot n}{T}\right) & \sin\left(\frac{2\cdot 2\pi\cdot n}{T}\right)\cdots & \cos\left(\frac{k_a\cdot 2\pi\cdot n}{T}\right) & \sin\left(\frac{k_a\cdot 2\pi\cdot n}{T}\right) \end{bmatrix} \tag{26}$$

$$q^{(0)} = \frac{1}{\mathrm{Pot}\, I_{i\in \mathrm{I}}} \sum |q^{(0)}(i)|$$
$$I = \{1, 2, \ldots, n-j\}$$

3. Fourier series

Due to the simplicity of the grey model prediction, the modeling accuracy may not be very high even though it has surpassed some prediction approaches in time series prediction. As stated in Bingqian (1990), grey models may have worse curve fitting effects in case of random data or center-symmetry situation. Various remedies have been discussed in several literatures (Cheng et al., 1997; Tan & Chang, 1996; Tan & Lu, 1996; Wang, 2002). Many methods have been proposed to increase accuracy by modeling the residuals of the original grey prediction. In fact, residual diagnosis is a commonly used correction approach for time series prediction. One of those residual correction approaches is to use Fourier series (Tan & Lu, 1996). Owing to the good performance of the Fourier correction approach, it is adopted in our research to improve the modeling performance of grey models. Notice that the Fourier correction approach is to increase prediction capability from the considered input data set and does not change the local characteristics of the grey model prediction.

Let the residual series E_r be defined as

$$E_r = \{E_r(2)E_r(3)\ldots E_r(n)\}^T \tag{23}$$

where

$$E_r(k) = X^{(0)}(k) - \hat{X}^{(0)}(k)$$

for $k = 2, 3, \ldots, n$.

Fourier series can approximate the residual series as Tan and Chang (1996).

$$E(k) = \frac{1}{2}a_0 + \sum_{i=1}^{k_a}\left[a_i \cos\left(\frac{i\cdot 2\pi}{T}k\right) + b_i \sin\left(\frac{i\cdot 2\pi}{T}k\right)\right] \tag{24}$$

for $k = 2, 3, \ldots, n$

where $T = n-1$ and $k_a = [(n-1)/2 - 1$. Here $[(n-1)/2] - 1$is the integer portion of $(n-1)/2$. The idea of using Fourier series is to transform the residuals into frequency spectra and then select low-frequency terms. In our later simulation, we will show that exact fitting models in fact may result in very poor prediction for real cases. Fourier approaches, on the other hand, can filter out high-frequency terms, which are supposed to be noise, and then have very nice performance. The parameters a_0, a_i, and b_i for $i = 1, 2, \ldots, k_a$ in (24) are estimated by applying the least square approach and obtained as

$$C = (P^TP)^{-1}P^TE_r \tag{25}$$

where $C = [a_0\ a_1\ b_1\ a_2\ b_2 \cdots a_{ka}\ b_{ka}]^T$ and P is shown in the following equation:

Therefore, the original prediction can be corrected as

$$\hat{X}(1) = X^{(0)}(1) \quad \text{and} \quad \hat{X}^{(0)}(k) = \hat{X}^{(0)}(k) + E(k)$$
$$\text{for } k = 2, 3, \ldots, n, n+1, \ldots \tag{27}$$

This prediction method is called the Fourier grey model. In our implementation, FGM is performed for the most recent five data points to predict the value of the next step. From the simulations the FGM approach indeed can significantly improve the prediction accuracy of grey models.

4. Markov chains theory

A Markov chain is a particular type of stochastic process $\{X(t), t \in T\}$, where T is a subset of $(-\infty, +\infty)$ and is treated as the time parameter set. If the index set is discrete, e.g. $T = \{0, 1, 2, \ldots\}$, then we have a discrete-time Markov chain; otherwise, if T is continuous, e.g., $T = \{t{:}0 \leqslant t \leqslant +\infty\}$, we call the process a continuous-time Markov chain. Note that, a Markov chain must satisfy the Markov property, namely, the future evolution of the conditional probability depends only on the current state of the system and not on its history.

Most of the books dealing with Markov processes concentrate either on the exposition of their mathematical properties (Chung, 1967; Iosifescu, 1980), or provide their introductory properties (Medhi, 1982; William, 1994). Besides, Markov process has been used extensively to construct stochastic models in a variety of disciplines like biology, electrical engineering, physics, sociology, finance etc. Moreover, Markov process is a forecasting method which can be used to predict the future by these occurred events. The most popular one is regression analysis.

4.1. Transition probability and matrix

Let the state space of a Markov chain $\{X_m\}$ be S, the current state be i and the next state be j, then the transition probability is written as

$$P_{ij} = \mathrm{Prob}\{X_{m+1} = j | X_m = i\} \quad (i, j \in S, m = 0, 1, 2, \ldots) \tag{28}$$

where the P_{ij} is independent of m.

The matrix P, formed by placing P_{ij} in row i and column j, for all i and j, is called the transition probability matrix or chain matrix

(William, 1994). Note that the elements of the matrix P satisfy the following two properties:

(1) $P_{ij} \geqslant 0 \quad \forall i,j \in S$

(2) $\sum_{j \in S} P_{ij} = 1 \quad \forall i \in S$

4.2. Stationary finite state Markov chain

In a Markov chain, if the states of possible appearance all come from a state space $S = \{1, 2, \ldots, k\}$ $(k < \infty)$of a finite set, the Markov chain is called a finite state Markov chain. Its transition probability matrix can be expressed as follows:

$$P = \begin{bmatrix} P_{11} & P_{12} & \cdots & P_{1k} \\ P_{21} & P_{22} & \cdots & P_{2k} \\ \vdots & \vdots & \vdots & \vdots \\ P_{k1} & P_{k2} & \cdots & P_{kk} \end{bmatrix} \tag{29}$$

Note that, if the transition probability$P_{ij}$is independent of time, then we call the Markov chain a stationary finite state Markov chain.

5. Markov–Fourier grey prediction system

The major purpose of this research attempts to develop an effective prediction model to promote the prediction accuracy of stock market. Firstly, the grey relational analysis is developed to extract the most important quantitative technical indices for constructing the grey model. Secondly, in order to find out the best modeling data, the technique of different size of sliding window is used in the grey model. Then, the Fourier series to correct the residual is adopted in our research for improving the modeling performance of grey model. Finally, a powerful prediction model is generated by integrating the FGM and Markov model, termed as MFGM, for promoting the prediction accuracy.

5.1. GRA process

For the consideration of the forecast of medium-term trend of TAIEX, the short-term indices such as 3MAP, 5MAP and so on, are neglected. That is, this study emphasizes the indices including 24MAP, 72MAP, 144MAP and 288MAP. The method of GRA is used to filter the major technical indices from various MAP for prediction. The number of experiment data of daily TAIEX, collected from October 1996 to August 2007, was totally 2878 points during the 131 months. There are two ways to analyze the grey relational grade between the referential and compared series. One is the whole moving average price, and the other is different size of sliding window such as 6, 12, 24, and 72. This research adopts the latter method to compare the referential and compared series. There are totally 2807 sliding windows to analyzed, each sliding window is composed of 6, 12, 24, and 72 historical data, respectively, among them, the data of the first slide window is the newest k (k = 6, 12, 24, and 72) one, and the data of the last sliding window is the oldest one. The comparison is from the first sliding window till the last one. During the comparison process, sliding window model should be operated in accordance with the principle of keeping the same dimension of data series. That means, a new data is attached on the rear of the k-size of sliding window and the first data in the k-size of sliding window should be removed before the next comparison. For example, suppose the first sliding window of the referential series is $s_1(k)$, where k = 6, then the first sliding window is

$$s_1 = (s_1(2807), s_1(2806), \ldots, s_1(2802))$$

The second sliding window $s_2(k)$ is

$$s_2 = (s_2(2806), s_2(2805), \ldots, s_2(2801))$$

And so on, the last sliding window $s_{2807}(k)$ is

$$s_{2807} = (s_{2807}(77), s_{2807}(76), \ldots, s_{2807}(72))$$

Similarly, we can get the different size of sliding window for comparison.

After the reference data series and all of the comparative data series are established, we utilize the following steps to find out the optimal compared series for grey prediction model to forecast the turning time of TAIEX highs/lows.

(1) Normalize all of the data series by the Eq. (5).
(2) Calculate the GRG (grey relationship grade) between the referential series and compared series by the Eq. (3) and Eq. (4).
(3) Order the technical indices according to their respective degrees of relationship from the biggest to the smallest. For instance, $\Gamma_{01} \succ \Gamma_{05} \succ \Gamma_{02} \succ \Gamma_{09} \succ \ \ldots \succ \Gamma_{0n}$.
(4) Compute and accumulate the frequency of the biggest GRG of each comparative series and the referential series. For example, $\Gamma_{01(\max)} = 123, \Gamma_{05(\max)} = 88, \Gamma_{02(\max)} = 99, \ldots, \Gamma_{0n(\max)} = 66$.
(5) Repeat step 2 to step 4 till end.
(6) Find out the comparative series which keep the maximum frequency of the biggest GRG.

The most suitable technical index will be selected as the input of the grey prediction model for finding out the best modeling data. The frequency of the maximum GRG between various moving average of price and TAIEX, are presented in Table 1. As Table 1 shows, the maximum GRG is 24MAP which will be utilized in the grey prediction model to find out the optimal modeling data and forecast the turning time of highs/lows.

5.2. Grey prediction model

After the optimal compared series, namely 24MAP, is obtained, the grey model GM(1,1) with different size of sliding window is utilized to find the optimal number of modeling data for forecasting the next turning time of highs/lows. The numbers of historical highs (lows) of 24MAP, ranged from October 1996 to August 2007, was 28. The index and time of TAIEX highs/lows based on 24MAP are listed in Table 2, where the time points as an initial data series $x^{(0)}$, means the time sequence used in various models proposed for training and testing. The simulation results are shown in Table 3, in which the optimal number of modeling data for highs and lows is 5. After obtaining the optimal modeling data, the first 21 and last 12 points of 24MAP highs and lows are used as the training data set and testing data set, respectively. Table 4 shows the prediction results of grey prediction model. Despite its superior forecasting capacity, GM(1,1) can be further improved by integrating the Fourier series in the grey prediction model. For evaluating the accuracy of prediction model, the mean residual error (MRE) is calculated as

$$\mathrm{MRE} = \frac{1}{n} \sum_{k=1}^{n} e(k) \tag{30}$$

Table 1
The frequency of maximum GRG among various MAP and TAIEX.

Size of sliding window	24MAP	72MAP	144MAP	288MAP
6	974	672	569	592
12	973	675	557	602
24	977	651	574	605
72	1527	427	385	468

Table 2
The related data of 24MAP highs/lows.

Highs			Lows		
Date (mm/dd/yy)	Index	Time point	Date (mm/dd/yy)	Index	Time point
05/03/97	8511.38	152	06/05/97	8122.37	180
08/27/97	9908.05	248	11/19/97	7630.65	311
01/03/98	8209.18	346	01/22/98	7935.85	361
04/10/98	9034.31	415	06/30/98	7540.08	476
07/31/98	7874.10	500	09/21/98	6787.32	541
12/15/98	7166.64	604	03/02/99	6000.31	654
07/13/99	8345.77	754	08/13/99	7354.43	780
09/14/99	8087.02	804	11/18/99	7606.82	851
03/03/00	9807.06	923	03/24/00	9278.96	939
04/21/00	9627.07	960	05/30/00	8665.60	989
06/16/00	8861.69	1002	01/04/01	5041.31	1156
02/27/01	5861.03	1188	08/14/01	4380.93	1302
09/04/01	4499.51	1317	10/23/01	3675.36	1347
04/26/02	6256.06	1470	10/22/02	4201.93	1594
12/10/02	4712.07	1629	01/07/03	4613.96	1648
02/12/03	4848.61	1668	04/04/03	4433.89	1704
04/22/03	4512.26	1716	05/28/03	4279.41	1741
11/18/03	6035.75	1862	12/23/03	5842.28	1887
03/19/04	6771.06	1943	04/15/04	6560.37	1962
04/29/04	6667.90	1972	08/18/04	5386.42	2050
10/13/04	5943.24	2087	11/16/04	5928.04	2109
01/11/05	5999.24	2150	02/14/05	8928.04	2168
03/17/05	6146.53	2189	05/19/05	5852.70	2233
08/15/05	6370.82	2292	11/11/05	5841.76	2355
02/14/06	6623.87	2415	03/29/06	6497.04	2446
05/19/06	7179.74	2480	08/15/06	6453.63	2542
02/27/07	7822.59	2671	03/28/07	7697.24	2693
08/03/07	9326.65	2780	09/12/07	8759.60	2710

Table 3
Mean residual error of GM(1,1) with different window.

Highs		Lows	
Window	MRE	Window	MRE
3	4.69501	3	4.86698
4	4.11194	4	3.83229
5	**4.05503**	**5**	**3.70827**
6	4.30577	6	3.90561
7	5.17762	7	4.24385
8	5.07843	8	4.25993
9	5.25012	9	4.39234
10	5.08137	10	4.36413
11	4.83091	11	5.03271
12	5.65264	12	5.89553
13	6.76734	13	6.50489
14	8.05526	14	8.04379
15	9.48184	15	8.99285
16	10.32272	16	9.62426
17	10.92800	17	10.05236
18	11.97854	18	10.94158
19	12.93997	19	11.71338
20	13.41286	20	12.50948
21	14.23777	21	13.09337
22	14.90806	22	13.45139
23	14.94963	23	13.52845
24	15.13262	24	14.04106
25	15.78577	25	14.43637
26	14.52639	26	14.11066
27	14.60432	27	17.78598

Table 4
Prediction results of 24MAP highs/lows based on GM(1,1).

Original	GM(1,1)	$e(k)$	Original	GM(1,1)	$e(k)$
Training results of highs			*Training results of lows*		
604	626	3.64	654	662	1.22
754	723	4.11	780	786	0.77
804	910	13.18	851	916	7.64
923	967	4.77	939	1005	7.03
960	1053	9.69	989	1061	7.28
1002	1059	5.69	1156	1081	6.49
1188	1088	8.42	1302	1252	3.84
1317	1251	5.01	1347	1452	7.8
1470	1471	0.07	1594	1527	4.2
1629	1671	2.58	1648	1729	4.92
1668	1809	8.45	1704	1821	6.87
1716	1842	7.34	1741	1867	7.24
1862	1821	2.2	1887	1799	4.66
1943	1916	1.39	1962	1943	0.97
1972	2054	4.16	2050	2066	0.78
2087	2092	0.24	2109	2172	2.99
MRE		**5.06**	MRE		**4.67**
Testing results of highs			*Testing results of lows*		
2150	2149	0.05	2168	2197	1.34
2189	2229	1.83	2233	2246	0.58
2292	2283	0.39	2355	2296	2.51
2415	2348	2.77	2446	2425	0.86
2480	2496	0.65	2542	2550	0.31
2671	2602	2.58	2693	2658	1.3
2780	2781	0.04	2693	2658	1.3
MRE		**1.19**	MRE		**1.46**

5.3. Fourier grey model (FGM)

As mentioned earlier, owing to the Fourier correction approach can increase the accuracy by modeling the residuals of the time series prediction, it is adopted in our research to improve the modeling performance of grey model. Notice that the Fourier correction approach is to increase prediction capability from the considered input data set and does not change the local characteristics of the grey model prediction. The prediction results of 24MAP highs/lows based on FGM is shown in Table 5. Obviously, according to the MRE of GM(1,1) and FGM, it is evident that the FGM has better performance than that of the grey prediction model GM(1,1). Despite its superior performance, another approach, namely Markov state transition matrices, is proposed to incorporate FGM so that the prediction can be more accurate.

Table 5
Prediction results of 24MAP highs/lows based on GM(1,1) and FGM.

Original	GM(1,1)	$e1(k)$	FGM	$e2(k)$	Original	GM(1,1)	$e1(k)$	FGM	$e2(k)$
Training results of highs					*Training results of lows*				
604	626	3.64	623	3.15	654	662	1.22	669	2.29
754	723	4.11	739	1.99	780	786	0.77	785	0.64
804	910	13.18	857	6.59	851	916	7.64	894	5.05
923	967	4.77	998	8.13	939	1005	7.03	1002	6.71
960	1053	9.69	1001	4.27	989	1061	7.28	1037	4.85
1002	1059	5.69	1056	5.39	1156	1081	6.49	1108	4.15
1188	1088	8.42	1109	6.65	1302	1252	3.84	1220	6.3
1317	1251	5.01	1221	7.29	1347	1452	7.8	1425	5.79
1470	1471	0.07	1482	0.82	1594	1527	4.2	1601	0.44
1629	1671	2.58	1679	3.07	1648	1729	4.92	1645	0.18
1668	1809	8.45	1776	6.47	1704	1821	6.87	1831	7.45
1716	1842	7.34	1833	6.82	1741	1867	7.24	1837	5.51
1862	1821	2.2	1830	1.72	1887	1799	4.66	1831	2.97
1943	1916	1.39	1869	3.81	1962	1943	0.97	1878	4.28
1972	2054	4.16	2031	2.99	2050	2066	0.78	2084	1.66
2087	2092	0.24	2127	1.92	2109	2172	2.99	2163	2.56
MRE		**5.06**		**4.44**			**4.67**		**3.8**
Testing results of highs					*Testing results of lows*				
2150	2149	0.05	2124	1.21	2168	2197	1.34	2203	1.61
2189	2229	1.83	2213	1.1	2233	2246	0.58	2230	0.13
2292	2283	0.39	2300	0.35	2355	2296	2.51	2310	1.91
2415	2348	2.77	2354	2.53	2446	2425	0.86	2408	1.55
2480	2496	0.65	2471	0.36	2542	2550	0.31	2560	0.71
2671	2602	2.58	2653	0.67	2693	2658	1.3	2679	0.52
2780	2781	0.04	2752	1.01	2710	2800	3.32	2756	1.7
MRE		**1.19**		**1.03**			**1.46**		**1.16**

5.4. Markov–Fourier grey model (MFGM)

Furthermore, Markov state transition matrices are used to improve the accuracy of FGM mentioned. This approach is called the Markov–Fourier grey model (MFGM), which combines the advantages of FGM (Tan & Chang, 1996) and the Markov forecasting model (Tan & Lu, 1996). Markov transition matrices can accumulate all information for all training data or only accumulate information for a period of time. The algorithm is first to use the prediction results of FGM as the estimated data. Next, we partition the state space into several states for the estimated data and find the most fitting number of states by simulation program. After comparing the original data with FGM estimated data, the state number of the original data is determined and the state transition

Table 6
Prediction results of 24MAP highs based on GM(1,1), FGM and MFGM.

Original	GM(1,1)	$e1(k)$	FGM	$e2(k)$	MFGM	$e3(k)$
Training results of highs						
604	626	3.64	623	3.15	604	0
754	723	4.11	739	1.99	754	0
804	910	13.18	857	6.59	806	0.25
923	967	4.77	998	8.13	923	0
960	1053	9.69	1001	4.27	961	0.1
1002	1059	5.69	1056	5.39	1003	0.1
1188	1088	8.42	1109	6.65	1103	7.15
1317	1251	5.01	1221	7.29	1319	0.15
1470	1471	0.07	1482	0.82	1489	1.29
1629	1671	2.58	1679	3.07	1654	1.53
1668	1809	8.45	1776	6.47	1669	0.06
1716	1842	7.34	1833	6.82	1723	0.41
1862	1821	2.2	1830	1.72	1821	2.2
1943	1916	1.39	1869	3.81	1925	0.93
1972	2054	4.16	2031	2.99	1960	0.61
2087	2092	0.24	2127	1.92	2084	0.14
MRE		**5.06**		**4.44**		**0.93**
Testing results of highs						
2150	2149	0.05	2124	1.21	2145	0.23
2189	2229	1.83	2213	1.1	2213	1.1
2292	2283	0.39	2300	0.35	2288	0.17
2415	2348	2.77	2354	2.53	2354	2.53
2480	2496	0.65	2471	0.36	2471	0.36
2671	2602	2.58	2653	0.67	2653	0.67
2780	2781	0.04	2752	1.01	2780	0
MRE		**1.19**		**1.03**		**0.72**

Table 7
Prediction results of 24MAP lows based on GM(1,1), FGM and MFGM.

Original	GM(1,1)	$e1(k)$	FGM	$e2(k)$	MFGM	$e3(k)$
Training results of lows						
654	662	1.22	669	2.29	656	0.31
780	786	0.77	785	0.64	781	0.13
851	916	7.64	894	5.05	854	0.35
939	1005	7.03	1002	6.71	942	0.32
989	1061	7.28	1037	4.85	990	0.1
1156	1081	6.49	1108	4.15	1102	4.67
1302	1252	3.84	1220	6.3	1293	0.69
1347	1452	7.8	1425	5.79	1311	2.67
1594	1527	4.2	1601	0.44	1593	0.06
1648	1729	4.92	1645	0.18	1645	0.18
1704	1821	6.87	1831	7.45	1703	0.06
1741	1867	7.24	1837	5.51	1745	0.23
1887	1799	4.66	1831	2.97	1858	1.54
1962	1943	0.97	1878	4.28	1963	0.05
2050	2066	0.78	2084	1.66	2063	0.63
2109	2172	2.99	2163	2.56	2120	0.52
MRE		**4.67**		**3.8**		**0.78**
Testing results of lows						
2168	2197	1.34	2203	1.61	2170	0.09
2233	2246	0.58	2230	0.13	2230	0.13
2355	2296	2.51	2310	1.91	2322	1.4
2446	2425	0.86	2408	1.55	2444	0.08
2542	2550	0.31	2560	0.71	2560	0.71
2693	2658	1.3	2679	0.52	2692	0.04
2710	2800	3.32	2756	1.7	2728	0.66
MRE		**1.46**		**1.16**		**0.44**

probability matrices are established. According to the state transition probability matrices, we can produce the prediction table and calculate new forecasting results by averaging the upper bound and lower bound. Finally, a more accurate forecasting result is produced after MFGM process is finished. Tables 6 and 7 show the results based on the GM(1,1), FGM and MFGM.

6. Experiment results

The experiments focus on the TAIEX ranged from October 1996 to August 2007. There are seven experiment results shown as Tables 1–7. Table 1 shows the frequency of maximum relational degree among concerned MAP and TAIEX. Clearly the 24MAP having maximum GRG is used for modeling. The historical highs (lows) of 24MAP and the corresponding time points are listed in Table 2, where the time points are used to construct an original data series and to forecast the next turning time of highs/lows. Table 3 shows the mean residual error about different dimension of grey prediction model GM(1,1). It is clear in Table 3, that both the optimal number of modeling data for highs and lows are 5. Table 4 shows the prediction results of highs/lows based on grey prediction model GM(1,1). Due to the mean residual error of training/testing data of highs/lows is 5.06/1.19 and 4.67/1.46, respectively, both have better performance. From Table 5, it is evident that the performance of FGM is better than that of the GM(1,1). As for the last two tables, namely Tables 6 and 7, list the prediction results of all prediction models. From these two tables, it is obvious that the performance of MFGM is the best among all the prediction models.

7. Conclusions

In this paper, the MFGM approach is presented to forecast the turning time of TAIEX for improving the prediction accuracy. The proposed Markov–Fourier grey model prediction approach uses the grey model to roughly predict the next datum from a set of the most recent data and the Fourier series to fit the residual errors produced by the grey model. Finally, Markov state transition matrices are employed to promote the predict results generated by the FGM. From Tables 6 and 7, it is evident that the proposed approach MFGM has a higher forecasting accuracy than other methods. While MFGM generates precise forecasting result, it is only suitable for long-term operation. To create a universally applicable model, we may combine long-term with short-term operational strategies. The short-term operational strategy will be the topic in our future research.

References

Baba, N., & Kozaki, M. (1992). An intelligent forecasting system of stock price using neural networks. In *Proceedings of IJCNN* (Vol. I, pp. 371–377).

Bingqian, L. (1990). The grey modeling for central symmetry sequence. *The Journal of Grey System, 2*(2), 95–103.

Chang, W. B., & Su, S. F. (1993). Using linear forms for the consequences in neural-network-based fuzzy logic controllers. In *Proceedings of the 1993 international symposium on artificial neural networks*.

Cheng, C. S., Hsu, Y. T., & Wu, C. C. (1997). Grey modeling using neural network. *The Journal of Grey System, 3*, 219–232.

Chi, S.-C., Chen, H.-P., & Chend, C.-H. (1999). A forecasting approach for stock index future using grey theory and neural networks. In *Proceedings of the IJCNN* (Vol. 6, pp. 3850–3855).

Chung, K. L. (1967). *Markov chains with stationary transition probabilities* (2nd ed.). Berlin: Springer.

Deng, J. (1982). Control problems of grey system. *System and Control Letter, 1*, 288–294.

Deng, J. (1988). *Essential method of grey system* (pp. 17–34). WuHan: HUST Press.

Deng, J. (1989). Introduction to grey system theory. *The Journal of Grey System, 1*, 1–24.

Hsia, T. C. (1979). *System identification: Least square methods*. Davis: University of California.

Hsia, K. H., & Wu, J. H. (1997). A study on the data preprocessing in grey relational analysis. *Journal of Chinese Gery System, 1*, 47–53.

Hsu, Y. T., Cheng, C. S., & Wu, C. C. (1997). Grey modeling using neural network. *The Journal of Grey System, 9*(3), 219–232.

Hsu, Y. T., Chen, H. C., & Lin, C. B. (2000). A long-term prediction using GMs. *The Journal of Grey System, 12*(1), 41–54.

Hsu, Y. T., Lin, C. B., & Cheng, H. C. (1999). Grey model selections based on fuzzy trend prediction. *Journal of Fuzzy Mathematics, 7*(4), 897–905.

Hsu, Y. T., Yeh, J., & Chang, H. C. (2000). Grey relational analysis for image compression. *The Journal of Grey System, 12*(2), 131–138.

Iosifescu, M. (1980). *Finite Markov processes and their applications*. New York: John Wiely.

Jang, J. S. R. (1991). Adaptive-network-based fuzzy inference systems. *IEEE Transactions on Systems, Man, & Cybernetics, 23*, 665–685.

Jang, G. S., Lai, F. P., Wang, B. W., & Parng, T. M. (1993). Intelligent stock trading system with price trend prediction and reversal recognition using dual module neural network. *Journal of Applied Intelligence, 3*, 225–248.

Lin, C. T., & Lee, C. S. G. (1991). Neural-network-based fuzzy logic control and decision systems. *IEEE Transactions on Systems, Man, & Cybernetics, 40*, 1414–1427.

Lin & Man (1989). An application of the GM(1,1) model: The prediction of flight safety. *The Journal of Grey System, 1*(1).

Lin, C. B., Su, S. F., & Hsu, Y. C. (2001). High precision forecast using grey models. *International Journal of Systems Science, 32*(5), 609–619.

Liu, T. Y., Yeh, J., Chen, C. M., & Hsu, Y. T. (2004). A grey ART system for grey information processing. *Neurocomputing, 56*, 407–414.

Medhi, J. (1982). *Stochastic processes*. San Francisco: Holden-Day.

Su, S. F., Lin, C. B., & Hsu, Y. T. (2002). A high precision global prediction approach based on local prediction approaches. *IEEE Transactions on Systems, Man and Cybernetics, Part C: Applications and Review, 32*(4), 416–425.

Tan, C. L., & Chang, S. P. (1996). Residual correction method of Fourier series to GM(1,1) model. In *1996 first national conference on grey theory applications* (pp. 93–101).

Tan, C. L., & Lu, B. F. (1996). Grey Markov chain forecasting model. In *1996 first national conference on grey theory applications* (pp. 157–162).

Tseng, F.-M., & Tzeng, G.-H. (2002). A fuzzy seasonal ARIMA model for forecasting. *Fuzzy Sets and Systems, 126*, 367–376.

Wang, Y. T. (2002). Predicting stock price using fuzzy grey prediction system. *Expert System With Application, 22*, 33–39.

William, J. S. (1994). *Introduction to the numerical solution of Markov chains*. Princeton, New Jersey: Princeton University Press.