



NORTH-HOLLAND

Applied Hybrid Grey Model to Forecast Seasonal Time Series

FANG-MEI TSENG, HSIAO-CHENG YU, and GWO-HSIUNG TZENG

ABSTRACT

The grey forecasting model has been successfully applied to finance, physical control, engineering, economics, etc. However, no seasonal time series forecast has been tested. The authors of this paper proved that GM(1,1) grey forecasting model is insufficient for forecasting time series with seasonality. This paper proposes a hybrid method that combines the GM(1,1) grey forecasting model and the ratio-to-moving-average deseasonalization method to forecast time series with seasonality characteristics. Three criteria, i.e., the mean squares error (MSE), the mean absolute error (MAE), and mean absolute percentage error (MAPE) were used to compare the performance of the hybrid model against other four models, i.e., the seasonal time series ARIMA model (SARIMA), the neural network back-propagation model combined with grey relation, the GM(1,1) grey model with raw data, the GM(1,N) grey model combined with grey relation. The time series data of the total production value of Taiwan's machinery industry (January 1994 to December 1997) and the sales volume of soft drink reported from Montgomery's book were used as test data sets. Except for the out-of-sample error of the Taiwan machinery production value time series, the MSE, the MAE, and the MAPE of the hybrid model were the lowest. © 2001 Elsevier Science Inc.

Introduction

The grey system theory was initially presented by Deng (1982) [1–3]. The grey forecasting model adopts the essential part of the grey system theory, and it has been successfully used in finance, physical control, engineering and economics [4–6]. The advantages of the grey forecasting model include: (a) it can be used in circumstances with relatively little data; as low as four observations were reported [7] to estimate the outcome of an unknown system; and (b) it can use a first-order differential equation to characterize a system. Therefore, only a few discrete data are sufficient to characterize an unknown system. This leads to the suggestion that the grey forecasting model is suitable for forecasting the competitive environment where decision makers can reference only limited historical data. However, no application of the grey forecasting model was reported on time series with seasonality. Hwang et al. [8] only used the grey relation

FANG-MEI TSENG is Associate Professor of Department of Finance, Hsuan Chuang University, Taiwan.

HSIAO-CHENG YU is Associate Professor of Institute of Management of Technology, College of Management, National Chiao Tung University, Taiwan.

GWO-HSIUNG TZENG is Professor of Energy and Environmental Research Group, Institute of Management of Technology and Institute of Traffic and Transportation, College of Management, National Chiao Tung University, Taiwan.

Address reprint requests to Fang-Mei Tseng, No. 12, Lane 10, Ming Yeo 2nd Street, Hsinchu 300, Taiwan.

Technological Forecasting and Social Change 67, 291–302 (2001)

© 2001 Elsevier Science Inc. All rights reserved.

655 Avenue of the Americas, New York, NY 10010

0040-1625/01/\$–see front matter

PII S0040-1625(99)00098-0

to select the influential factors for power-load forecasting and used the neural network back-propagation model to build the forecasting model. The authors of this paper proved that the GM(1,1) grey forecasting model is insufficient for forecasting a time series with seasonality. This research proposed a hybrid forecasting model that combined the GM(1,1) grey forecasting model and the ratio-to-moving-average deseasonalization method to remove the seasonality characteristic from a seasonal time series.

Three criteria—i.e., the MSE, the MAE, and the MAPE—were used to compare the performance of the hybrid model against four other models—i.e., the SARIMA model, the neural network back-propagation model combined with the grey relation, the GM(1,1) grey model (i.e., first-order, one variable grey model) with originally seasonal time series data, the GM(1, N) grey model (i.e., first-order, N variables grey model) combined with the grey relation. The time series data of the total production value of Taiwan's machinery industry (January 1994 to December 1997) and the sales volume of soft drink reported from Montgomery's book [19] were used as test data sets. The results showed that, except for the out-of-sample error of the Taiwan's machinery production time series, the hybrid method, i.e., the GM(1,1) grey model with deseasonalized data, outperformed other models.

The remainder of this paper is organized as follows: Section 2 describes the deseasonalizing method and the grey system theory. In Section 3, details of the hybrid models are discussed. Section 4 describes the evaluation methods used for comparing the performance of forecasting techniques. Section 5 describes the procedures of the four forecasting models, i.e., the GM(1,1) grey model with deseasonalized data, the SARIMA model, the back-propagation model combined with grey relation, and the GM(1, N) grey model combined with the grey relation. Section 6 compares the results obtained from the five forecasting models. Section 7 presents conclusions.

A Deseasonalizing Method and the Grey System Theory

The purpose of this paper is to proposed the hybrid method that combined the GM(1,1) model and the ratio-to-moving-average method. To explained the method, the ratio-to-moving-average method and the grey system theory are described in the following.

THE RATIO-TO-MOVING-AVERAGE METHOD

Deseasonalizing means removing the seasonality from a time series data set [10]. After we do so, period-to-period comparisons are more meaningful, and can help identify whether a trend exists.

The ratio-to-moving-average method is the method to calculate the seasonal indexes and remove the seasonal factor from a time series data set. That includes six steps [11]: (1) compute a k period moving average of the time series from the first data to the last data (assume the seasonal time period is k); (2) calculate the centered moving average. Take the average of the two moving averages from step (1), i.e., the one is at the beginning of the period and the other one is the end of the period; (3) divide the actual value of each period by the centered moving average; (4) calculate the median of the ratios for each time period; (5) adjust each period's mean in order to make the mean value of the k medians equal to 100; and (6) remove the seasonal factor by dividing the actual value of each period by its corresponding adjusted mean.

THE GREY SYSTEM THEORY

The theory of the grey system is based on the assumption that a system is uncertain, and that the information regarding the system is insufficient to build a relational analysis or to construct a model to characterize the system [12]. The grey system puts each stochastic variable as a grey quantity that changes within a given range. It does not rely on statistical method to deal with the grey quantity. It deals directly with the original data, and searches the intrinsic regularity of data [9]. The grey system theory include the following fields: (a) grey generating, (b) grey relational analysis, (c) grey forecasting, (d) grey decision making, and (e) grey control. In the following, the basic ideas of grey relational analysis and grey forecasting are introduced because they are directly related to this study.

Grey Relational Analysis

The grey relational analysis is used to determine the relationship between two sequences of stochastic data in a grey system. The procedure may bear some similarity to the pattern recognition technology. One sequence of data is called the “reference pattern” or “reference sequence,” and the correlation of the other sequence to the reference sequence is to be identified [13, 15, 17, 18]. Let \mathbf{x}_0 be the reference pattern with n entries

$$\mathbf{x}_0 = (x_0(1), x_0(2), \dots, x_0(n)) \quad (1)$$

and \mathbf{x}_i be one of the m patterns with n entries to be compared with \mathbf{x}_0 (each \mathbf{x}_i has the same n number of entries as \mathbf{x}_0). The \mathbf{x}_i is written as

$$\mathbf{x}_i = (x_i(1), x_i(2), \dots, x_i(n)) \quad 1 \leq i \leq m. \quad (2)$$

The set of the sequence \mathbf{x}_i is generally the influencing factor to \mathbf{x}_0 . The grey relational coefficient between the compared pattern \mathbf{x}_i , and the reference pattern \mathbf{x}_0 at the k th entry ($k = 1, 2, \dots, n$) is defined

$$\begin{aligned} \gamma_{ik} &= \gamma_{ik}(x_0(k), x_i(k)) \\ &= \frac{\min_i \min_k |x_0(k) - x_i(k)| + \zeta \max_i \max_k |x_0(k) - x_i(k)|}{|x_0(k) - x_i(k)| + \zeta \max_i \max_k |x_0(k) - x_i(k)|} \end{aligned} \quad (3)$$

where $\zeta \in [0, 1]$ is the distinguishing factor to control resolution scale, typically taken as 0.5 [17, 18]. When $x_0(k)$ equals $x_i(k)$, the coefficient of grey relation is $\gamma_{ik} = 1$. This indicates that $x_i(k)$ is highly related to $x_0(k)$. When $|x_0(k) - x_i(k)|$ is the maximum value over all k entries, γ_{ik} reaches the minimum value over all k entries. The grey relational factor $\gamma_i(\mathbf{x}_0, \mathbf{x}_i)$ between the reference pattern \mathbf{x}_0 and the compared pattern \mathbf{x}_i is taken as the average of γ_{ik} over all k entries. When the coefficients of grey relation are equally important at all entries,

$$\gamma_i = \gamma_i(\mathbf{x}_0, \mathbf{x}_i) = \frac{1}{n} \sum_{k=1}^n \gamma_{ik}(x_0(k), x_i(k)). \quad (4)$$

If the coefficients of grey relation are not equally important to the factor of grey relation, the weighting factor β_k should be included in Equation (4). The factor of grey relation is then given as

$$\gamma = \gamma_i(x_0, x_i) = \frac{1}{n} \sum_{k=1}^n \beta_k \gamma_{ik}(x_0(k), x_i(k)). \quad (5)$$

The factor of grey relation γ_i is the similarity indicator of the pattern x_0 and the pattern x_i . If $\gamma_i > \gamma_j$, then the pattern x_i has characteristics closer to those of the reference pattern x_0 than the pattern x_j .

Grey Forecasting

This section briefly reviews the operation of grey forecasting. Grey forecasting model (GM) has three basic operations: (1) accumulated generation, (2) inverse accumulated generation, and (3) grey modeling. 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) grey model, i.e., a single variable first-order grey model, is summarized as follows [13]: (a) Step 1: the initial sequence is

$$\mathbf{x}^{(0)} = (x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(i), \dots, x^{(0)}(n)), \quad (6)$$

where $x^{(0)}(i)$ is the time series data at time i . (b) Step 2: based on the initial sequence $\mathbf{x}^{(0)}$, a new sequence $\mathbf{x}^{(1)}$ is generated by the accumulated generating operation (AGO), where

$$\mathbf{x}^{(1)} = (x^{(1)}(1), x^{(1)}(2), \dots, x^{(1)}(n)) \text{ and is } x^{(1)}(k)$$

derived as follows:

$$x^{(1)}(k) = \sum_{i=1}^k x^{(0)}(i). \quad (7)$$

(c) Step 3: the following first-order differential equation holds true:

$$\frac{dx^{(0)}}{dt} + ax^{(1)} = u \quad (8)$$

(d) Step 4: from Step 3, we have

$$\hat{x}^{(1)}(k+1) = \left(x^{(0)}(1) - \frac{u}{a} \right) e^{-ak} + \frac{u}{a}, \quad (9)$$

$$\hat{x}^{(0)}(k+1) = \hat{x}^{(1)}(k+1) - \hat{x}^{(1)}(k), \quad (10)$$

where

$$\hat{\mathbf{a}} = \begin{bmatrix} a \\ u \end{bmatrix} = (\mathbf{B}^T \mathbf{B})^{-1} \mathbf{B}^T \mathbf{y}_N, \quad (11)$$

$$\mathbf{B} = \begin{bmatrix} -0.5(x^{(1)}(1) + x^{(1)}(2)) & 1 \\ -0.5(x^{(1)}(2) + x^{(1)}(3)) & 1 \\ \vdots & \vdots \\ -0.5(x^{(1)}(n-1) + x^{(1)}(n)) & 1 \end{bmatrix} \quad (12)$$

$$\mathbf{y}_N = (x^{(0)}(2), x^{(0)}(3), \dots, x^{(0)}(n))^T \quad (13)$$

$\hat{x}^{(1)}(k+1)$ is the predicted value of $x^{(1)}(k+1)$ at time $k+1$.

The GM(1,1) grey model can be easily extended to a GM(1, N) grey model. Note that the second index in the GM(1, N) grey model stands for N variables ($\mathbf{x}_1^{(0)}, \mathbf{x}_2^{(0)}, \dots, \mathbf{x}_N^{(0)}$), and the differential equation can be written as follows:

$$\frac{dx_1^{(0)}}{dt} + ax_1^{(1)} = \sum_{i=2}^N b_{i-1}x_i^{(1)} \quad (14)$$

where $a, b_1, b_2, \dots, b_{N-1}$ are unknown parameters. According to Step 4 of the GM(1,1) grey model, these parameters can be estimated as follows:

$$\hat{\mathbf{a}} = (\hat{a}, \hat{b}_1, \hat{b}_2, \dots, \hat{b}_{N-1}) = (\mathbf{B}^T \mathbf{B})^{-1} \mathbf{B}^T \mathbf{y}_N \quad (15)$$

where

$$\mathbf{B} = \begin{bmatrix} -0.5(x_1^{(1)}(1) + x_1^{(1)}(2)) & x_2^{(1)}(2) & \cdots & x_N^{(1)}(2) \\ -0.5(x_1^{(1)}(2) + x_1^{(1)}(3)) & x_2^{(1)}(3) & \cdots & x_N^{(1)}(3) \\ \vdots & \vdots & \cdots & \vdots \\ -0.5(x_1^{(1)}(n-1) + x_1^{(1)}(n)) & x_2^{(1)}(n) & \cdots & x_N^{(1)}(n) \end{bmatrix} \quad (16)$$

$$\mathbf{y}_N = (x^{(0)}(2), x^{(0)}(3), \dots, x^{(0)}(n))^T. \quad (17)$$

The forecasts of $x_1^{(1)}$ are as follow:

$$\hat{x}_1^{(1)}(k+1) = (x_1^{(0)}(1) - \sum_{i=2}^N \frac{b_{i-1}}{a} x_i^{(1)}(k+1))e^{-ak} + \sum_{i=2}^N \frac{b_{i-1}}{a} x_i^{(1)}(k+1). \quad (18)$$

The authors used the posterror test to evaluate the accuracy of the grey forecasting. The forecasting errors were defined as $q^{(0)}(k) = x^{(0)}(k) - \hat{x}^{(0)}(k)$, $k = 1, 2, \dots, n$. The mean and the standard deviation of the forecasting errors are \bar{q} and S_2 . The mean and the standard deviation of the original time series are

$$\bar{m} = \frac{\sum_{k=1}^n x^{(0)}(k)}{n}, S_1 = \sqrt{\frac{\sum_{k=1}^n (x^{(0)}(k) - \bar{m})^2}{n-1}}.$$

The posterror ratio C is derived by dividing S_2 by S_1 , $C = S_2/S_1$. The lower the C is, the better the model is. The posterror ratio can indicate the change rate of the forecasting error. Probability of small error is defined as $p = \text{prob.}\{|q^{(0)}(k) - \bar{q}| \leq 0.6745S_1\}$, where $k = 2, 3, \dots, n$ [13]. P is another indicator of forecasting accuracy. This shows the probability that the relative bias of the forecasting error is lower than 0.6745. p is commonly required to be larger than 0.95. The pairs of the forecasting indicators p and C can characterize four grades of forecasting accuracy [13], as shown in Table 1.

The Hybrid Forecasting Model

The u and a are parameters derived from Equation (11). Let $P = x^{(0)}(1) - u/a$ be a constant. Equation (10) can be expressed as follows:

TABLE 1
The Grades of Forecasting Accuracy

Grade	Forecasting indicators	
	p	C
Good	>0.95	<0.35
Qualified	>0.8	<0.5
Just	>0.7	<0.65
Unqualified	≤ 0.7	≥ 0.65

Source: [13].

$$\hat{x}^{(0)}(k+1) = Pe^{-ak}(e^{-a} - 1) \quad (19)$$

Hence, the forecast of the GM(1,1) model will always generate either exponentially increasing or decreasing series. Based on this argument, the authors claimed that GM(1,1) model is insufficient to forecast time series with seasonality. Therefore, any seasonal time series data must be preprocessed to remove the seasonality component before building a GM(1,1) grey model.

In this section, the authors propose to use the ratio-to-moving-average method to remove the seasonality from a seasonal time series before building a GM(1,1) grey model.

A time series may be decomposed into three separate components of trend, cycle, and seasonal factors. The seasonal factor could be either additive or multiplicative to other components of a time series. The ratio-to-moving-average method is one way to characterize the seasonal factor. In this study, the authors used this method to deseasonalize the original time series data before building a GM(1,1) grey model.

Forecast Evaluation Methods

Yokum and Armstrong [14] conducted an expert opinion survey to select evaluation criteria for forecasting techniques. Accuracy was the most important criterion, followed by the cost savings generated from improved decisions. In particular, execution issues such as ease of interpretation and ease of use were also highly rated. In this study, there are three criteria used to evaluate forecasting models.

The first measurement is mean square error, MSE:

$$MSE = \sum_{t=1}^T (P_t - Z_t)^2 / T \quad (20)$$

where P_t is the predicted value at time t , Z_t is the actual value at time t , and T is the number of predictions.

The second criterion is the mean absolute error, MAE:

$$MAE = \sum_{t=1}^T |P_t - Z_t| / T \quad (21)$$

A third criterion is the mean absolute percent error, MAPE:

$$MAPE = \frac{1}{T} \sum_{t=1}^T \left| \frac{P_t - Z_t}{Z_t} \right| \quad (22)$$

Experimental Results

The above forecasting models were applied to two time series: one is the total production value of Taiwan's machinery industry, and the other is the sales volume of soft drink quoted from Montgomery's book [19]. In the following section the characteristics of the production value of Taiwan's machinery industry are described. Subsequent sections report the forecasting results of the machinery output time series based on the SARIMA model, the neural network model combined with grey relation, GM(1,N) grey model combined with grey relation and the hybrid forecasting model that combines the ratio-to-move-average model and the GM(1,1) grey model. Another section describes the soft drinks time series. Then subsections report the forecasting results of the soft drink time series.

THE TAIWAN MACHINERY INDUSTRY TIME SERIES

The machinery industry in Taiwan has made steady progress over the past decade. It has played a critical supporting role as the foundation for the whole manufacturing

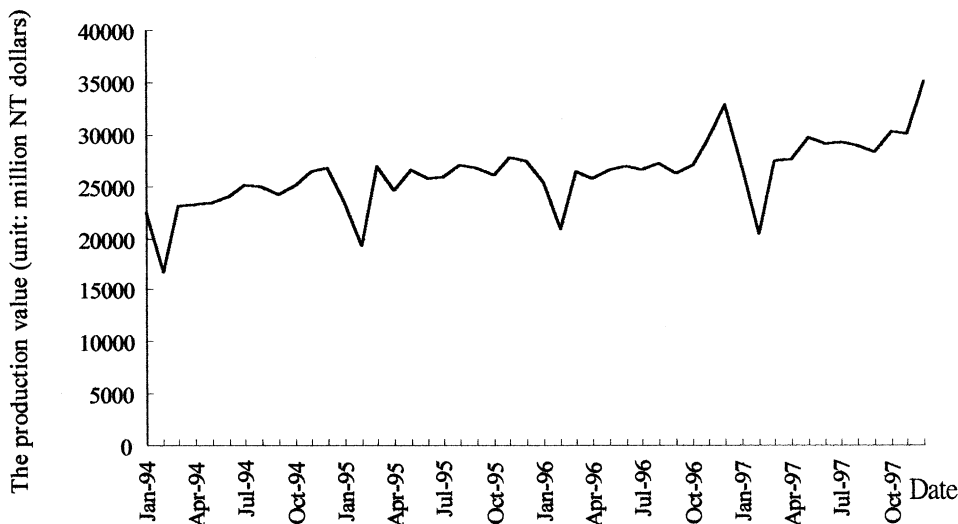


Fig. 1. Total production value of Taiwan's machinery industry (January 1994 to December 1997).

industry in Taiwan, and it is a major exporting industry itself. The time series data of the total production revenues of Taiwan's machinery industry in the period from January 1994 to December 1997 showed strong seasonality and growth trend, as shown in Figure 1. The sharp drop in each year generally happens in January or February, due to the plant shutting down for the Chinese New Year holiday. The time series data from January 1994 to December 1996 were used as known data (i.e., in-sample data) to forecast the 1997 data (i.e., out-of-sample data). The difference between the actual and the forecast for the 12 months of 1997 were used to evaluate the accuracy of the forecasting model.

Building SARIMA Model

The time series data was preprocessed by taking first-order regular differencing and first-order seasonal differencing to remove the growth trend and seasonality. The authors used SAS statistical package to implement the SARIMA model. Akaike Information Criterion (AIC) [20, 21] was used to determine the best model. The derived model is $ARIMA(0,1,0)(0,1,1)_{12}$, and the equation is

$$(1 - B)(1 - B^{12})Z_t = (1 - 0.7664B)a_t \quad (23)$$

Building the Neural Network Back-Propagation Model Combined With the Grey Relation

Hwang et al. [8] showed that grey relation analysis is appropriate to identify the most influential factors for power-load forecasting. In this study, the authors used similar ideas to build a neural network back-propagation model where the choice of input nodes was derived from the grey relation.

The authors first initialized the raw data and calculated the grey relation of the reference series $(W_{t-1}, W_{t-2}, \dots, W_{t-12})$. The first and second rank of the grey relation is W_{t-12}, W_{t-9} . The authors thus decided to choose the input nodes as Z_{t-12}, Z_{t-9} (12-month and 9-month lags) to build the neural network back-propagation model. At the training stage, the number of neurons in the hidden layer, the learning rate and the momentum were tried. The best result was produced at three hidden neurons, learning rate 0.5, momentum 0.95, and iteration 10,000.

Building the GM(1,3) Grey Model Combined With the Grey Relation

From the results of the previous grey relation, the grey variables x_t, x_{t-12} , and x_{t-9} , were used to build the GM(1,3) model. The parameters of Equation (15) are $\hat{a} = (1.673, 0.698, 1.025)$. The function is

$$\hat{x}_1^{(1)}(k+1) = (23370 - 0.417 * x_2^{(1)}(k+1) - 0.613 * \hat{x}_3^{(1)}(k+1))\exp(1.673k) + 0.417 * x_2^{(1)}(k+1) + 0.613 * \hat{x}_3^{(1)}(k+1). \quad (24)$$

The forecasts are estimated through one operation of the inverse of the accumulated generating operation. Due to the fact that $c = 0.67$ and $p = 0.75$, the model was characterized as an unqualified model.

Building the GM(1,1) Model With Deseasonalized Data

The authors used the ratio-to-moving-average method to deseasonalize the original time series, and the deseasonalized time series is used as input to build the GM(1,1) grey model. The deseasonalized data went through one operation of the accumulated generating operation. Using the least-squares method, the parameters are $\hat{a} = (-0.00539, 23003.99)$. The function is

$$\hat{x}^{(1)}(k+1) = (23571.91 + 23003.99/0.00539)\exp(0.00539k) - 23003.99/0.00539 \quad (25)$$

The forecasts were estimated through one operation of the inverse of the accumulated generating operation. Due to the fact that $c = 0.3176$ and $p = 0.97$, the model was characterized as making accurate forecasts.

THE SOFT DRINK TIME SERIES

To demonstrate the performance of the five models, the authors applied these models to another time series, which was the monthly sales volume of soft drink from Montgomery's book "Forecasting and Time Series Analysis," ([19] p. 364). The time series demonstrates growth trend and seasonality, as is shown in Figure 2.

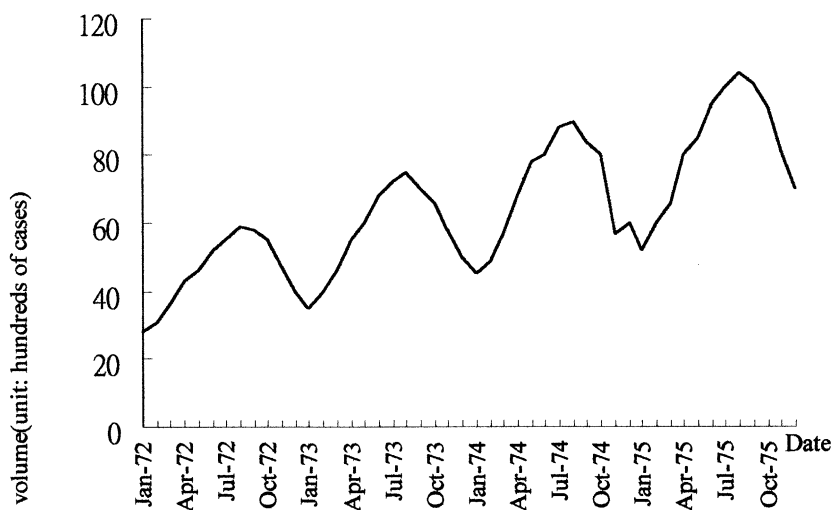


Fig. 2. Monthly sales volume of soft drinks.

Building the SARIMA Model

The time series data was preprocessed by the logarithmic transformation, first-order regular differencing and first-order seasonal differencing to fix the variance and remove the growth trend and seasonality. The authors used SAS statistical package to formulate the SARIMA model. Akaike Information Criterion (AIC) [20, 21] was used to determine the best model. The derived model is $ARIMA(1,1,0)(0,1,0)_{12}$, and the equation is

$$(1 + 0.73B)(1 - B)(1 - B^{12})Z_t = a_t \quad (26)$$

Building the Neural Network Back-Propagation Model Combined With the Grey Relation

The authors initialized the raw data and calculated the grey relation of the reference series $(W_{t-1}, W_{t-2}, \dots, W_{t-12})$. The first and second rank of the grey relation is W_{t-12}, W_{t-11} . The authors thus decided to choose the input nodes Z_{t-12}, Z_{t-11} (12-month and 11-month lags) to build the neural network back-propagation model. At the training stage, the number of neurons in the hidden layer, the learning rate, and the momentum were tried thoroughly. The best result was produced at four hidden neurons, learning rate 0.3, momentum 0.95, and iteration 10,000.

Building the GM(1,3) Grey Model Combined with the Grey Relation

Using the same grey relation derived in the previous subsection, the grey variables x_t, x_{t-12} , and x_{t-11} were used to build the GM(1,3) model. The parameters of Equation (15) were $\hat{a} = (1.082, -0.145, 1.484)$. The function is

$$\begin{aligned} \hat{x}_1^{(1)}(k+1) = & (35 + 0.134 * x_2^{(1)}(k+1) - 1.372 * \hat{x}_3^{(1)}(k+1))\exp(-1.082k) \\ & - 0.134 * x_2^{(1)}(k+1) + 1.372 * \hat{x}_3^{(1)}(k+1). \end{aligned} \quad (27)$$

The forecasts were generated through one operation of the inverse of the accumulated generating operation. Due to the fact that $c = 0.48$ and $p = 0.83$, the model was characterized as making accurate forecasts.

Building the GM(1,1) Grey Model With Deseasonalized Data

The authors used the ratio-to-moving-average method to deseasonalize the original time series, and used the deseasonalized data as input to build the GM(1,1) grey model. The deseasonalized data went through one operation of the accumulated generating operation. Using the least-squares method, the parameters are $\hat{a} = (-0.0164, 42.32678)$. The function is

$$\hat{x}^{(1)}(k+1) = (41.11 + 42.32678/0.0164)\exp(0.0164k) - 42.32678/0.0164. \quad (28)$$

The forecasts were estimated through one operation of the inverse of the accumulated generating operation. Due to the fact that $c = 0.169$ and $p = 0.97$, the model was characterized as making accurate forecasts.

Evaluations and Comparisons

In this section, the performance of the previous models in forecasting the Taiwan's machinery industry output as well as the sales volume of soft drinks are reported. The measurement criteria include MAPE, MAE, and MSE.

In addition to MAPE, MAE, and MSE measurements, T-value was also used to test the hypothesis that the GM(1,1) grey model with deseasonalized data, the SARIMA model, the GM(1,1) grey model with raw data, the GM(1,N) grey model, and the neural

TABLE 2
Performance Measurements of the Forecast of the Machinery Production Time Series

Data type	SARIMA (differenced)	GM(1,1) (raw)	GM(1,3) + grey relation (raw)	Back-propagation + grey relation (raw)	GM(1,1) + ratio-to- moving-average (deseasonalized)
In-sample errors					
MAE ^a	1,084.97	1,380.84	1,421.49	803.35	620.90
MSE ^b	1,580,679.00	4,609,697.00	4,237,156.00	1,476,153.00	749,363.30
MAPE ^c	3.91	5.88	5.34	2.91	2.41
T-value ^d	2.45**	2.89**	2.49**	1.56	—
Out-of-sample errors					
MAE ^a	941.56	1,673.86	1,744.56	1,912.40	996.41
MSE ^b	1,584,786.00	7,833,501.00	5,125,347.00	7,156,162.00	1,852,473.00
MAPE ^c	3.26	6.45	6.41	6.86	3.45
T-value ^d	-0.64	1.12	1.97*	1.40	—

^a Mean Absolute Error; ^b Mean Squares Error; ^c Mean Absolute Percent Error; ^d Null hypothesis of the existence of the same means of the forecast absolute errors generated by the GM(1,1) grey model with deseasonalized data and other models.

** Denotes rejected at five percent significance level; * denotes rejected at the 10% significance level.

network back-propagation model with grey relation have the same means of absolute errors. If the hypothesis is statistically rejected, we would have demonstrated a better model. The results are shown in Tables 2 and 3.

The in-sample error in Tables 2 and 3 indicate that GM(1,1) grey model with deseasonalized data outperformed the SARIMA model, the GM(1,3) grey model, the neural network back-propagation model, and the GM(1,1) grey model with raw data for both the machinery production time series and the soft drink time series. For the machinery production time series, the MSE of the GM(1,1) grey model with raw data, the neural network back-propagation model, the GM(1,3) grey model, and the SARIMA model were 4,609,697, 1,476,153, 4,237,156, and 1,580,679, respectively, while the MSE of the GM(1,1) grey model with deseasonalized data was considerably lower at

TABLE 3
Performance Measurements of the Forecast of the Soft Drink Time Series

Data type	SARIMA (differenced)	GM(1,1) (raw)	GM(1,3) + grey relation (raw)	Back-propagation + grey relation (raw)	GM(1,1) + ratio-to- moving-average (deseasonalized)
In-sample errors					
MAE ^a	3.67	9.55	6.35	1.94	1.62
MSE ^b	31.69	123.53	51.11	8.58	7.30
MAPE ^c	5.43	17.44	10.67	3.07	2.83
T-value ^d	2.05**	8.56**	4.73**	0.18	—
Out-of-sample errors					
MAE ^a	12.19	12.75	11.77	5.47	2.71
MSE ^b	172.96	234.15	189.23	38.01	9.93
MAPE ^c	14.33	17.92	15.08	6.98	3.14
T-value ^d	7.18**	3.60**	3.96**	2.98**	—

^a Mean Absolute Error; ^b Mean Squares Error; ^c Mean Absolute Percent Error; ^d Null hypothesis of the existence of the same means of the forecasted absolute errors generated by the GM(1,1) grey model with deseasonalized data and other models.

** Denotes rejected at 5% significance level; * denotes rejected at the 10% significance level.

749,363.30. The MAPE of the GM(1,1) grey model with deseasonalized data was only 2.41, which is at least half a percentage point better than the other models. The MAE of the GM(1,1) grey model with deseasonalized data is also better than those of the other models. For the in-sample error of the soft drink time series shown in Table 3, the GM(1,1) grey model with deseasonalized data still has the lowest MAE, MSE, and MAPE. T-tests also showed rejection of the hypothesis that the mean absolute errors of the GM(1,1) grey model with deseasonalized data is the same as those of the other models except for the neural network model.

For out-of-sample error comparisons, Table 2 and Table 3 indicate that the GM(1,1) grey model with deseasonalized data outperformed the GM(1,3) grey model, the neural network back-propagation model, and the GM(1,1) with raw data for both the machinery production time series and the soft drinks time series. For out-of-sample forecast of the machinery production time series, the SARIMA model outperformed the GM(1,1) model with deseasonalized data but not with statistical significance. The hybrid model still outperforms the SARIMA model for the out-of-sample forecast of the soft drink time series. For the machinery production time series, the MAE of the GM(1,1) grey model with deseasonalized data, the GM(1,3) grey model, the neural network back-propagation model, and the GM(1,1) with raw data were 996.41, 1744.56, 1,912.40, and 1,673.86, respectively, while the MAE for the SARIMA model was lowest at 941.56. Nevertheless, the T-tests could not reject the hypothesis that the mean absolute errors of the GM(1,1) grey model with deseasonalized data are the same as the SARIMA models. The MSE and MAPE of the GM(1,1) grey model with deseasonalized data are better than those of the other models except for the SARIMA model. For the soft drink time series, the GM(1,1) grey model with deseasonalized data has the lowest MAE, MSE, and MAPE. T-tests also showed rejection of the hypothesis that the MAE of the GM(1,1) grey model with deseasonalized data is the same as the MAEs of the other models.

Conclusions

The authors proved that the GM(1,1) grey model is insufficient for forecasting time series with seasonality. Therefore, a seasonal time series should be deseasonalized first before building a GM(1,1) model. The authors proposed to use the ratio-to-moving average method to remove the seasonality out of a seasonal time series.

For the in-sample forecast, the GM(1,1) with deseasonalized data outperformed the SARIMA model, the GM(1, N) grey model combined with grey relation, the GM(1,1) grey model with raw data, and the neural network back-propagation model combined grey relation, for both the Taiwan machinery output time series and the soft drink time series. For the out-of-sample, the GM(1,1) grey model with deseasonalized data outperformed the SARIMA model, the GM(1, N) grey model combined with grey relation, the GM(1,1) grey model with raw data, and the neural network back-propagation model combined grey relation for the soft drink time series only.

The out-of-sample errors for the SARIMA model are less than those of the GM(1,1) grey model with deseasonalized data. Nevertheless, the hypothesis that the means of the model are different cannot be rejected with statistical significance.

References

1. Deng J. L.: Introduction to Grey System Theory, *The Journal of Grey System* 1(1) 1–24 (1989).
2. Deng J. L.: Properties of Multivariable Grey Model GM(1, N), *The Journal of Grey System* 1(1) 125–141 (1989).

3. Deng J. L.: Control Problems of Grey Systems, *System and Control Letters* 1(1) 288–294 (1989).
4. Shih, N. Z., and Liou, D. K.: An Evaluation Study of Future Indexes Hedging Strategies in Grey System—Applied on Volume Weighted Index and Future Index, *1997 Second National Conference on Grey Theory and Applications* (Taiwan), 1997, pp. 16–33.
5. Chen, J. Y., and Lin, Y. H.: Design of Fuzzy Sliding Mode Controller With Grey Predictor, *The Journal of Grey System* 8(2) 147–164 (1996).
6. Chiao, J. H., Wang, W. Y., and Lu, M. J.: A Study for Applying Grey Forecasting to Improve the Reliability of Product, *1997 Second National Conference on Grey Theory and Applications* (Taiwan) 1997, pp. 202–206.
7. Homg, C. M., Chiang, C. T., and Lin, S. C.: Design of Grey Prediction Controller Based on Fuzzy Reasoning, *The Second National Conference on Fuzzy Theory and Application* (Taiwan), 1994, pp. 66–71.
8. Hwang, R. C., Huang, H. C., Chen, Y. J., Hsieh, J. G., Chen, H. C., and Chuang, C. W.: Selection of Influencing Factors for Power Load Forecasting by Grey Relation Analysis, *1997 Second National Conference on Grey Theory and Applications* (Taiwan), 1997, pp. 109–113.
9. Huang, Y. P., Huang, C. C., and Hung, C. H.: Determination of the Preferred Fuzzy Variables and Applications to the Prediction Control by the Grey Modeling, *The Second National Conference on Fuzzy Theory and Application* (Taiwan), 1994, pp. 406–409.
10. Mansfield, E.: *Statistics for Business and Economics: Methods and Application*, 5th ed. W. W. Norton and Company, New York, 1994, p. 580.
11. Wheelwright, S. C., and Makridakis, S.: *Forecasting Methods for Management*. John Wiley & Sons, Inc., New York, 1985, pp. 84–103.
12. Tseng, F. M., Tzeng, G. H., Yu, C. H., and Yuan, B.: The Comparison of Four Kinds of Prediction Methods: ARIMA, Fuzzy Time Series, Fuzzy Regression Time Series and Grey Forecasting—An Example of the Production Value Forecast of Machinery Industry in Taiwan, *1997 Second National Conference on Grey Theory and Applications* (Taiwan), 1997, pp. 53–68.
13. Deng, J. L.: *Grey Forecasting and Decision*. Huazhong University of Science and Technology Press, Wuhan, 1986, pp. 97–134.
14. Yokum, J. T., and Armstrong, J. S.: Beyond Accuracy: Comparison of Criteria Used to Select Forecasting Methods, *International Journal of Forecasting* 11(4) 591–597 (1995).
15. Wu, H. S., Deng, J. L., and Wen, K. L.: *Introduction of Grey Analysis*. Gau-Li Publication Inc., Taiwan, 1996, pp. 12–52.
16. Chiang, J. S., Wu, P. L., Chiang, S. D., Chang, T. J., Chang, S. T., and Wen, K. L.: *The Introduction of Grey System Theory*, Gau-Li Publication Inc., Taiwan, 1998, pp. 57–69.
17. Tzeng, G. H., and Tsaor, S. H.: The Multiple Criteria Evaluation of Grey Relation Model, *The Journal of Grey System* 6(3) pp. 87–104 (1994).
18. Mon, D. L., Tzeng, G. H., and Lu, H. C.: Grey Decision Making in Weapon System Evaluation, *Journal of Chung Institute of Technology* 26(1) 73–84 (1995).
19. Montgomery, D. C., Johnson, L. A., and Gardiner, J. S.: *Forecasting and Time Series Analysis*. McGraw-Hill, Inc., New York, 1990, p. 364.
20. Akaike, H.: A New Look at the Statistical Model Identification, *IEEE Transaction on Automatic Control* AC-19, 716–723 (1974).
21. Harvey, A. C.: *Time Series Models*. John Wiley & Sons, Inc., New York, 1981.

Received 9 July 1999; accepted 12 November 1999