Using Seven Types of GM (1, 1) Model to Forecast Hourly Particulate Matter Concentration in Banciao City of Taiwan

**Tzu-Yi Pai · Ching-Lin Ho · Shyh-Wei Chen · Huang-Mu Lo · Pao-Jui Sung ·
Shu-Wen Lin · Wei-Jia Lai · Shih-Chi Tseng · Shu-Ping Ciou · Jui-Ling Kuo ·
Jing-Tang Kao**

Abstract In this study, seven types of first-order
and one-variable grey differential equation model
(abbreviated as GM (1, 1) model) were used to
predict hourly particulate matter (PM) including
PM_{10} and $PM_{2.5}$ concentrations in Banciao City of
Taiwan. Their prediction performance was also
compared. The results indicated that the minimum
mean absolute percentage error (MAPE), mean
squared error (MSE), root mean squared error
(RMSE), and maximum correlation coefficient (R)
was 14.10%, 25.62, 5.06, and 0.96, respectively,
when predicting PM_{10}. When predicting $PM_{2.5}$, the
minimum MAPE, MSE, RMSE, and maximum R value
of 15.24%, 11.57, 3.40, and 0.93, respectively, could
be achieved. All statistical values revealed that the
predicting performance of GM (1, 1, $x^{(0)}$), GM (1, 1,
a), and GM (1, 1, b) outperformed other GM (1, 1)
models. According to the results, it revealed that GM
(1, 1) GM (1, 1) was an efficiently early warning tool
for providing PM information to the inhabitants.

Keywords Grey system theory · GM (1, 1) · Hourly
particulate matter · Air quality · PM_{10} · $PM_{2.5}$

T.-Y. Pai (✉) · H.-M. Lo · P.-J. Sung · S.-W. Lin ·
W.-J. Lai · S.-C. Tseng · S.-P. Ciou · J.-L. Kuo · J.-T. Kao
Department of Environmental Engineering and Management,
Chaoyang University of Technology,
Wufeng, Taichung 41349, Taiwan
e-mail: bai@ms6.hinet.net

C.-L. Ho
Department of Resources Engineering,
National Cheng Kung University,
Tainan 701, Taiwan

S.-W. Chen
Environmental Protection Bureau,
Taoyuan County Government,
Taoyuan 33001, Taiwan

P.-J. Sung
Dali City Administration, Taichung County Government,
Dali, Taichung 41261, Taiwan

1 Introduction

In the past two decades, air pollution has improved in
most cities in Western Europe, North American, as
well as Asia. Air pollution reductions have resulted
mainly from greater efficiency and pollution-control
technologies in factories, power plants, and other
facilities (Cunningham and Cunningham 2006). But
the problems of serious air pollution are still reported
frequently in many countries.

Among all air pollutants, the particulate matter (PM)
concentrations are of particular concern, because high
PM concentrations can cause human health problems
(Pisoni and Volta 2009). Epidemiological researches
have shown an association between ambient concen-
tration of PM pollution and negative effects on human
health (Pope 2000; Jaecker-Voirol and Pelt 2000). Pope
(2000) reported that when 24-h average ambient PM_{10}
concentrations increased by 10 $mg\,m^{-3}$, the daily all-

cause mortality would increase by 1% approximately, especially for cardiovascular and respiratory mortality. Therefore, developing an efficient forecasting and early warning system for providing air-quality information to the inhabitants becomes relatively important.

Typically, environmental data are very complex for modeling because interrelations between various components result in a complicated combination of relations. Models with reasonable accuracy have to consider physical and chemical relations among PM and other pollutants under various meteorological conditions simultaneously. However, the uncertainty problem will occur when above modeling approaches were adopted. One of the most important problems is the uncertainty from input data, including source identification, meteorological conditions, and relevant reaction mechanisms. No matter how good the inventory investigation was carried out in a large-scale modeling analysis, the uncertainties of input data in the mechanistic modeling process cannot be eliminated in reality.

Many other attempts to model the interrelations have also been carried out. Linear regression methods, for instance, have been widely employed for decades (Ryan 1995; Shi and Harrison 1997; Slini et al. 2006). Additionally, to adequately model complex, nonlinear phenomena, and chemical procedures, artificial neural networks (ANN) and fuzzy logic approach have been widely applied because they could simulate nonlinear data well (Pérez et al. 2000; Kolehmainen et al. 2001; Wang et al. 2003; Slini et al. 2006).

Although ANN could predict air pollutant concentrations successfully, it required a large quantity of data for constructing model. In order to simplify statistical complexity and gain consistent results from the investigation data for predicting air pollutant, the grey system theory (GST) is a suitable method.

The GST can resolve the problem of incomplete data and has been applied in our previous works (Deng 2002, 2005; Pai et al. 2007a, b 2008a, b, c). GST focuses on the relational analysis, model construction, and prediction of the indefinite and incomplete information. It requires only a small amount of data, and the better prediction results can be obtained.

There are many analysis methods in GST including grey model (GM). GM can be used to establish the relationship between many sequences of data, and its coefficients can be used to evaluate which sequence of data affects system significantly. If an efficient predicting method could be developed, a better control strategy could be sought.

The objectives of this study are listed as follows: (1) construct seven types of first-order and one-variable grey differential equation model (abbreviated as GM (1, 1) model) for predicting hourly PM_{10} and $PM_{2.5}$ concentrations in Banciao City of Taiwan; (2) compare the prediction performance of seven types of GM (1, 1) model.

2 Materials and Methods

2.1 Dataset

The monitoring data from air-quality monitoring station located in Banciao City was selected for study (Fig. 1). The concentrations of PM_{10} and $PM_{2.5}$ from 1st to 19th of April 2008 were investigated. They were sampled and investigated every hour, and their total number of data was 456. Among the data, 384 data points were used to obtain the coefficients of the models, and 72 data points were used as the observed values when evaluating the performance of the model. The mean value, standard deviation, maximum value, and minimum value of PM_{10} was 61.43, 28.90, 150.00, and 6.00 μg m^{-3}, respectively. That of $PM_{2.5}$ was 40.68, 20.36, 102.00, and 5.00 μg m^{-3}, respectively. The meteorological condition was ignored in this study.

2.2 Grey Modeling Process

In a situation where information is lacking, using fewer (at least four) system information, one can create a GM to describe the behavior of the few outputs. By means of accumulated generating operation (AGO), the disorderly and the unsystematic data may become exponentially behaved such that a first-order differential equation can be used to characterize the system behavior. Solving the differential equation will yield a time response solution for prediction. Through inverse accumulated generating operation (IAGO), the forecast can be transformed back to the sequence of original series. A grey modeling process is described as follows.

Assume that a series of data with n samples is expressed as:

$$X^{\{0\}} = (x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(n)), \qquad (1)$$

Fig. 1 Monitoring station of Banciao City
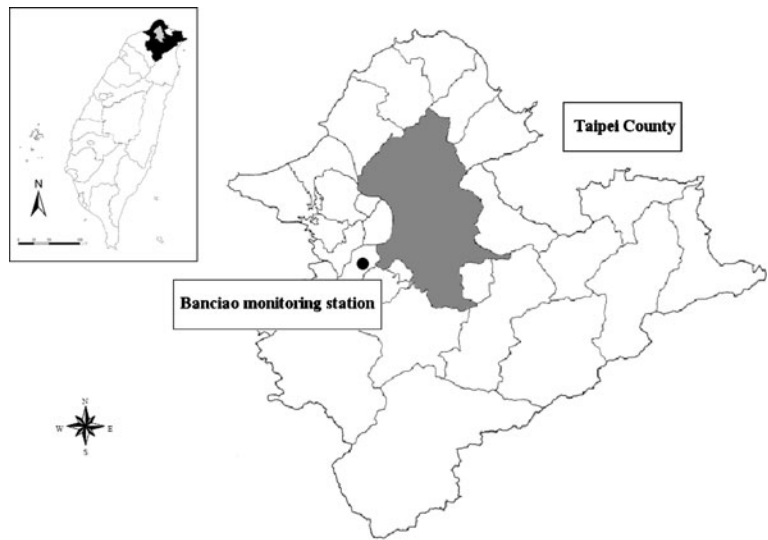
where the superscription (0) of $X^{(0)}$ represents the original series. Let $X^{(1)}$ be the first-order AGO of $X^{(0)}$, whose elements are generated from $X^{(0)}$:

$$X^{(1)} = (x^{(1)}(1), x^{(1)}(2), \cdots, x^{(1)}(n)), \qquad (2)$$

where $x^{(1)}(k) = \sum_{i=1}^{k} x^{(0)}(i)$, $for\ k = 1, 2, \cdots, n$. Further operation of AGO can be conducted to reach the r-order AGO series, $X^{(r)}$:

$$X^{\{r\}} = (x^{(r)}(1), x^{(r)}(2), \cdots, x^{(r)}(n)), \qquad (3)$$

where $x^{(r)}(k) = \sum_{i=1}^{k} x^{(r-1)}(i)$, $for\ k = 1, 2, \cdots, n$. The IAGO is the inverse operation of AGO. It transforms the AGO-operational series back to the one with a lower order. The operation of IAGO for the first-order series is defined as follows: $x^{(0)}(1) = x^{(1)}(1)$ and $x^{(0)}(k) = x^{(1)}(k) - x^{(1)}(k-1)$ $for\ k = 2, 3, \cdots, n$. After extending this representation to the IAGO of r-order series, we have $x^{(r-1)}(k) = x^{r}(k) - x^{r}(k-1)$ for k $= 2, 3, \cdots, n$. The tendency of AGO can be approximated by an exponential function. Its dynamic behavior resembles differential equation. The grey model GM (1, 1) thus adopts a first order differential equation to fit the AGO-operational series,

$$\frac{dx^{(1)}}{dt} + ax^{(1)} = b \qquad (4)$$

where the parameter a is the developing coefficient and b is the grey input. According to the definition, GM (1, 1) is that the order in grey differential equation is equal to 1 and defined as follows:

$$x^{(0)}(k) + az^{(1)}(k) = b \qquad (5)$$

where $z^{(1)}(k) = 0.5x^{(1)}(k-1) + 0.5x^{(1)}(k)$ k=2, 3, 4,…, n. Expanding Eq. 5, we have

$$\begin{array}{rcl} x^{(0)}(2) + az^{(1)}(2) & = & b \\ x^{(0)}(3) + az^{(1)}(3) & = & b \\ \vdots & \vdots & \vdots \\ x^{(0)}(n) + az^{(1)}(n) & = & b \end{array} \qquad (6)$$

Transforming Eq. 6 into matrix form, we have

$$\begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(n) \end{bmatrix} = \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \vdots & \vdots \\ -z^{(1)}(n) & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} \qquad (7)$$

Then, the coefficients can be estimated by solving matrix, $p = \begin{bmatrix} a \\ b \end{bmatrix} = (B^T B)^{-1} B^T Y$, where $p = \begin{bmatrix} a \\ b \end{bmatrix}$,

$$B = \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \vdots & \vdots \\ -z^{(1)}(n) & 1 \end{bmatrix}, \text{ and } Y = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(n) \end{bmatrix}.$$

Sometimes, singularity would be encountered when treating the increasingly accumulated data. Then, the inverse matrix could not be found. In our study, the increasingly accumulated data would not result in singularity due to their values, and numbers were not too high. Additionally, the whitening type of GM (1, 1) model (or in terms of GM (1, 1, W)) that can be used for prediction is described as:

$$\widehat{x}_1^{(1)}(k+1) = \left(x^{(0)}(1) - \frac{b}{a}\right) \cdot e^{-ak} + \frac{b}{a} \tag{8}$$

$$\widehat{x}^{(0)}(k+1) = \widehat{x}^{(1)}(k+1) - \widehat{x}^{(1)}(k) \tag{9}$$

Additionally, there still are several types of GM (1, 1) model which is derived from Eq. 4 as follows.

Connotation type of GM (1, 1): GM (1, 1, C)

$$x^{(0)}(k) = \left(\frac{1-0.5a}{1+0.5a}\right)^{k-2} \frac{b - ax^{(0)}(1)}{1+0.5a} \tag{10}$$

Grey difference type of GM (1, 1): GM (1, 1, $x^{(1)}$)

$$x^{(0)}(k) = \beta - \alpha x^{(1)}(k-1) \tag{11}$$

where $\beta = \frac{b}{1+0.5a}$ and $\alpha = \frac{a}{1+0.5a}$.

IAGO type of GM (1, 1): GM (1, 1, $x^{(0)}$)

$$x^{(0)}(k) = (1-\alpha)x^{(0)}(k-1) \tag{12}$$

Parameter-a type of GM (1, 1): GM (1, 1, a)

$$x^{(0)}(k) = \frac{1-0.5a}{1+0.5a}x^{(0)}(k-1) \tag{13}$$

Parameter-b type of GM (1, 1): GM (1, 1, b)

$$x^{(0)}(k) = \frac{x^{(1)}(k) - 0.5b}{x^{(1)}(k-1) + 0.5b}x^{(0)}(k-1) \tag{14}$$

Exponent type of GM (1, 1): GM (1, 1, e)

$$x^{(0)}(k) = x^{(0)}(3)e^{(k-3)\ln(1-\alpha)} \tag{15}$$

When adopting GM (1, 1, $x^{(0)}$), GM (1, 1, a), GM (1, 1, b), and GM (1, 1, e), $x^{(0)}$, (2) must be calculated as follows:

$$x^{(0)}(2) = \beta - \alpha x^{(0)}(1) \tag{16}$$

All seven types of GM (1, 1) model and their denotation are summarized in Table 1. Details for derivation of these GM (1, 1) model can be referred to Deng (2002, 2005).

2.3 Error Analysis

In order to evaluate the prediction accuracy of GM (1, 1), the mean absolute percentage error (MAPE), mean square error (MSE), root mean square error (RMSE), and correlation coefficient (R) were employed and described as,

$$\text{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{\text{obs}_i - \text{pre}_i}{\text{obs}_i}\right| \times 100\% \tag{17}$$

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(\text{obs}_i - \text{pre}_i)^2 \tag{18}$$

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\text{obs}_i - \text{pre}_i)^2} \tag{19}$$

$$R = \frac{\sum_{i=1}^{n}(\text{obs}_i - \overline{\text{obs}})(\text{pre}_i - \overline{\text{pre}})}{\sqrt{\sum_{i=1}^{n}(\text{obs}_i - \overline{\text{obs}})^2 \sum_{i=1}^{n}(\text{pre}_i - \overline{\text{pre}})^2}} \tag{20}$$

where obs_i is the observed value, pre_i is the prediction value, and $\overline{\text{obs}}$ and $\overline{\text{pre}}$ are the average values of observed values and prediction values, respectively.

3 Results and Discussion

3.1 Determination of Grey Parameters

For determining the parameters of GM (1, 1), the observed PM$_{10}$ and PM$_{2.5}$ data were substituted into Eq. 6 and the grey parameters were calculated by solving Eq. 7. When predicting PM$_{10}$, parameters a and b were equal to 0.00032527 and 65.452, respectively. When predicting PM$_{2.5}$, $a=0.0010902$ and $b=49.911$. According to Eq. 4, the parameter a (developing coefficient) will determine the predicting trend; meanwhile, parameter b (grey input) will determine the interception of Eq. 4.

3.1.1 Simulation of PM$_{10}$

Table 2 shows all the values of MAPE, MSE, RMSE, and R using seven types of GM (1, 1) model. The 1st to 384th data were used for constructing model, 385th

Table 1 Seven types of GM (1, 1) model

Type	Denotation	Prediction equation
Whitening type	GM (1, 1, W)	$\widehat{x}_1^{(1)}(k+1) = \left(x^{(0)}(1) - \frac{b}{a}\right) \cdot e^{-ak} + \frac{b}{a}$ $\widehat{x}^{(0)}(k+1) = \widehat{x}^{(1)}(k+1) - \widehat{x}^{(1)}(k)$
Connotation type	GM (1, 1, C)	$x^{(0)}(k) = \left(\frac{1-0.5a}{1+0.5a}\right)^{k-2} \frac{b - ax^{(0)}(1)}{1+0.5a}$
Grey difference type	GM (1, 1, $x^{(1)}$)	$x^{(0)}(k) = \beta - \alpha x^{(1)}(k-1)$ $\beta = \frac{b}{1+0.5a}, \ \alpha = \frac{a}{1+0.5a}$
IAGO type	GM (1, 1, $x^{(0)}$)	$x^{(0)}(k) = (1-\alpha)x^{(0)}(k-1)$ $x^{(0)}(2) = \beta - \alpha x^{(0)}(1)$
Parameter-a type	GM (1, 1, a)	$x^{(0)}(k) = \frac{1-0.5a}{1+0.5a}x^{(0)}(k-1)$ $x^{(0)}(2) = \beta - \alpha x^{(0)}(1)$
Parameter-b type	GM (1, 1, b)	$x^{(0)}(k) = \frac{x^{(1)}(k) - 0.5b}{x^{(1)}(k-1) + 0.5b}x^{(0)}(k-1)$ $x^{(0)}(2) = \beta - \alpha x^{(0)}(1)$
Exponent type	GM (1, 1, e)	$x^{(0)}(k) = x^{(0)}(3)e^{(k-3)\ln(1-\alpha)}$ $x^{(0)}(2) = \beta - \alpha x^{(0)}(1)$

to 456th data were used to evaluate the fitness. All statistical values revealed that the predicting performance of GM (1, 1, $x^{(0)}$), GM (1, 1, a), and GM (1, 1, b) prevailed. Figure 2a–c depict the prediction results of PM$_{10}$ using seven types of GM (1, 1) model.

As shown in Table 2, when constructing, MAPEs between the predicted and observed values of PM$_{10}$ were between 13.77% and 13.98% using GM (1, 1, $x^{(0)}$), GM (1, 1, a), and GM (1, 1, b), but they were 54.33–55.05% using other GM (1, 1) models. When predicting, the MAPEs were 14.10–14.11% when adopting GM (1, 1, $x^{(0)}$), GM (1, 1, a), and GM (1, 1, b), but they were between 56.27% and 151.14% when using other GM (1, 1) models.

The MSE values of 91.03–91.17 using GM (1, 1, $x^{(0)}$), GM (1, 1, a), and GM (1, 1, b) were lower than those of 826.57–2069.41 using other GM (1, 1) models when model constructing. When predicting, the values of 25.62–25.69 using GM (1, 1, $x^{(0)}$), GM (1, 1, a), and GM (1, 1, b) were also lower than those of 405.33–882.76 using other GM (1, 1) models. When constructing, the RMSE values of 9.54–9.55 using GM (1, 1, $x^{(0)}$), GM (1, 1, a), and GM (1, 1, b) were lower than those of 28.75–45.49 using other GM (1, 1) models. The RMSE value of 5.06–5.07 using GM (1, 1, $x^{(0)}$), GM (1, 1, a), and GM (1, 1, b) were also lower than those of 20.13–29.71 using other GM (1, 1) models when predicting.

When constructing, R value between the predicted and observed values of PM$_{10}$ was 0.95 using GM (1, 1, $x^{(0)}$), GM (1, 1, a), and GM (1, 1, b), but they were −0.01–0.08 using other GM (1, 1) models. When

Table 2 The prediction performance for PM$_{10}$ using seven types of GM (1, 1) model

	MAPE		MSE		RMSE		R	
	Constructing	Testing	Constructing	Testing	Constructing	Testing	Constructing	Testing
GM (1, 1, W)	54.96	150.03	826.76	869.16	28.75	29.48	0.08	−0.03
GM (1, 1, C)	54.99	150.10	826.76	870.03	28.75	29.50	0.08	−0.03
GM (1, 1, x(1))	55.05	151.14	826.57	882.76	28.75	29.71	0.08	−0.01
GM (1, 1, $x^{(0)}$)	13.98	14.10	91.17	25.69	9.55	5.07	0.95	0.96
GM (1, 1, a)	13.98	14.10	91.17	25.69	9.55	5.07	0.95	0.96
GM (1, 1, b)	13.77	14.11	91.03	25.62	9.54	5.06	0.95	0.96
GM (1, 1, e)	54.33	56.27	2069.41	405.33	45.49	20.13	−0.01	−0.03

Fig. 2 Prediction results of PM$_{10}$. **a** GM $(1, 1, x^{(0)})$, **b** GM $(1, 1, a)$, **c** GM $(1, 1, b)$

(a)

(b)

(c)
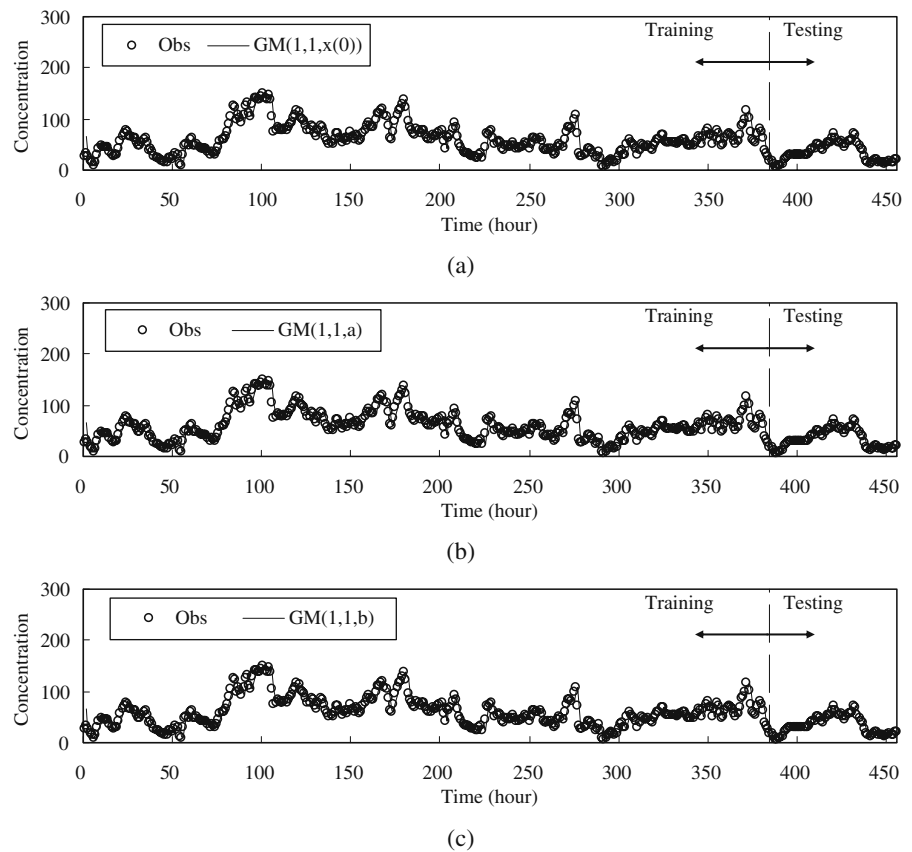
predicting, the R was 0.96 when adopting GM $(1, 1, x^{(0)})$, GM $(1, 1, a)$, and GM $(1, 1, b)$, but they were between −0.03 and −0.01 when using other GM $(1, 1)$ models.

Comparable observations were similarly made by Slini et al. (2006). Slini et al. (2006) adopted linear regression analysis (LRA), classification and regression trees (CART), principal component analysis (PCA), and neural network (NN) for predicting daily concentration levels of PM$_{10}$. The RMSE values for LRA, CART, PCA, and NN were 11.236, 33.55, 8.142, and 7.126, respectively. Díaz-Robles et al. (2008) used four types of model including multiple linear regression (MLR) model, Box–Jenkins time series (ARIMA) model, ANN, and a hybrid ARIMA-ANN model to forecast PM in urban areas. The RMSE values for MLR, ARIMA, ANN, and hybrid model were 28.39, 28.46, 28.57, and 8.80, respectively. In our study, the RMSE value of 5.06 using GM $(1, 1, x^{(0)})$, GM $(1, 1, a)$, and GM $(1, 1, b)$ could be obtained to predict hourly PM$_{10}$ concentrations.

3.1.2 Simulation of PM$_{2.5}$

The prediction results of PM$_{2.5}$ using seven types of GM $(1, 1)$ model are shown in Table 3. As the results of PM$_{10}$, all statistical values revealed that the predicting performance of GM $(1, 1, x^{(0)})$, GM $(1, 1, a)$, and GM $(1, 1, b)$ outperformed other models. Figure 3a–c shows their results.

As shown in Table 3, when constructing, MAPEs between the predicted and observed values of PM$_{2.5}$ were between 13.62% and 13.86% using GM $(1, 1, x^{(0)})$, GM $(1, 1, a)$, and GM $(1, 1, b)$, but they were 51.36–54.38% using other GM $(1, 1)$ models. When predicting, the MAPE was 15.24–15.25% when adopting GM $(1, 1, x^{(0)})$, GM $(1, 1, a)$, and GM $(1, 1, b)$, but they were between 41.53% and 109.94% when using other GM $(1, 1)$ models.

The MSE values of 42.00–42.46 using GM $(1, 1, x^{(0)})$, GM $(1, 1, a)$, and GM $(1, 1, b)$ were lower than those of 386.32–1033.61 using other GM $(1, 1)$ models when model constructing. When predicting, the values of 11.57–11.60 using GM $(1, 1, x^{(0)})$, GM

Table 3 The prediction performance for PM$_{2.5}$ using seven types of GM (1, 1) model

	MAPE		MSE		RMSE		R	
	Constructing	Testing	Constructing	Testing	Constructing	Testing	Constructing	Testing
GM (1, 1, W)	51.36	106.97	386.39	221.27	19.66	14.88	0.25	0.04
GM (1, 1, C)	51.43	107.16	386.39	222.06	19.66	14.90	0.25	0.04
GM (1, 1, x(1))	51.36	109.94	386.32	232.98	19.66	15.26	0.25	0.05
GM (1, 1, $x^{(0)}$)	13.62	15.25	42.00	11.60	6.48	3.41	0.95	0.93
GM (1, 1, a)	13.62	15.25	42.00	11.60	6.48	3.41	0.95	0.93
GM (1, 1, b)	13.86	15.24	42.46	11.57	6.52	3.40	0.95	0.93
GM (1, 1, e)	54.38	41.53	1033.61	152.55	32.15	12.35	0.16	0.04

(1, 1, a), and GM (1, 1, b) were also lower than those of 152.55–232.98 using other GM (1, 1) models. When constructing, the RMSE values of 6.48–6.52 using GM (1, 1, $x^{(0)}$), GM (1, 1, a), and GM (1, 1, b) were lower than those of 19.66–32.15 using other GM (1, 1) models. The RMSE values of 3.40–3.41 using GM (1, 1, $x^{(0)}$), GM (1, 1, a), and GM (1, 1, b) were also lower than those of 12.35–15.26 using other GM (1, 1) models when predicting.

When constructing, R value between the predicted and observed values of PM$_{2.5}$ was 0.95 using GM (1, 1, $x^{(0)}$), GM (1, 1, a), and GM (1, 1, b), but they were 0.16–0.25 using other GM (1, 1) models. When predicting, the R was 0.93 when adopting GM (1, 1, $x^{(0)}$), GM (1, 1, a), and GM (1, 1, b), but they were between 0.04 and 0.05 when using other GM (1, 1) models.

The reason why the predicting performances of GM (1, 1, $x^{(0)}$), GM (1, 1, a), and GM (1, 1, b)

Fig. 3 Prediction results of PM$_{2.5}$. **a** GM (1, 1, $x^{(0)}$), **b** GM (1, 1, a), **c** GM (1, 1, b)
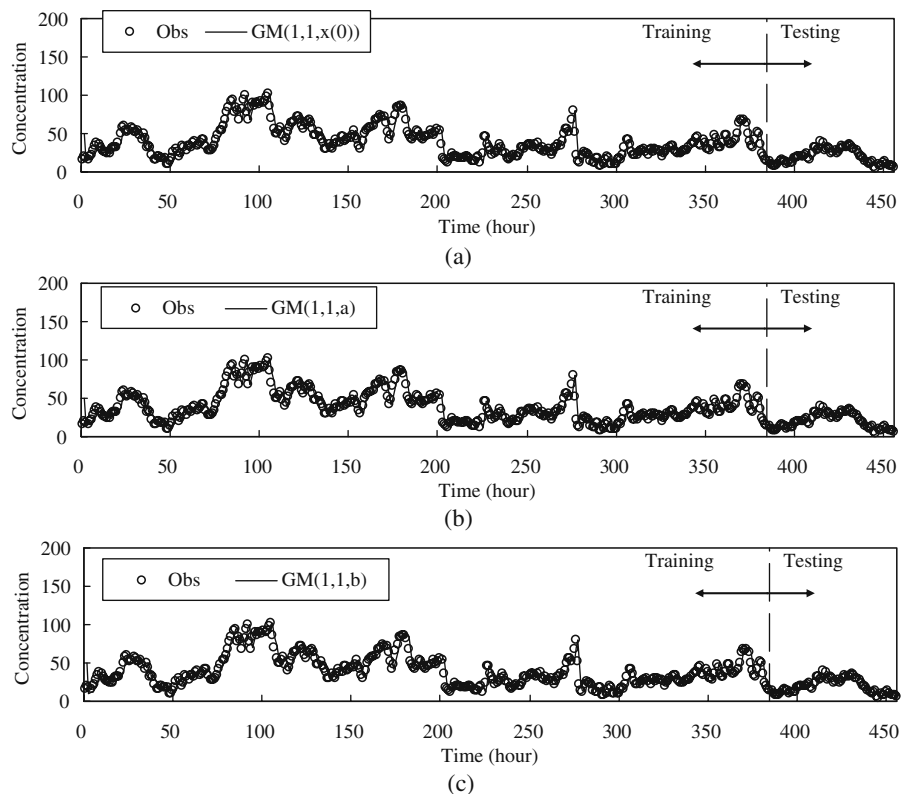
outperformed other GM (1, 1) model could be explained as follows. In the structure of GM (1, 1, $x^{(0)}$), GM (1, 1, a), and GM (1, 1, b), the point at time k is strongly influenced by the point at time $k-1$. For the time series of hourly PM, the value of PM did not vary significantly between hours. Therefore, the predicting performances of GM (1, 1, $x^{(0)}$), GM (1, 1, a), and GM (1, 1, b) outperform other GM (1, 1) model.

Additionally, in the mechanistically based modeling technique, the source identification, meteorological conditions, and relevant reaction mechanisms were adopted as the input variables. Then, the mechanistic models which consisted of lots of differential equations and kinetic reactions outputted the results which were calculated according to mass balance concept. In the GM (1, 1)-based modeling technique, the source identification, meteorological conditions, and relevant reaction mechanisms were not chosen as the input variables. Although the mechanisms were unclear, the whitening part of the GM (1, 1)-based model could serve as useful references to help observer realize more PM variation.

In the study proposed by Pérez et al. (2000), ANN, linear perceptron model, and persistence model were used to predict hourly average concentrations of $PM_{2.5}$. The MAPE values for these three types of model were between 20% and 80%. In our study, the MAPEs lay between 15.24% and 15.25% for predicting hourly $PM_{2.5}$ concentrations when adopting GM (1, 1, $x^{(0)}$), GM (1, 1, a), and GM (1, 1, b).

According to the results, the GM (1, 1) model required very small sample size, as little as four sample points, but the modeling could result in very high predictability. Furthermore, the parameter estimation in GM (1, 1) model was only a procedure to fit a simple regression. Therefore, GM could be applied successfully in predicting PM when the information was not sufficient.

4 Conclusions

Seven types of GM (1, 1) model were used to predict hourly PM_{10} and $PM_{2.5}$ concentrations in Banciao City of Taiwan. Their prediction performance was also compared. The conclusions can be drawn as follows. All statistical values revealed that the predicting performance of GM (1, 1, $x^{(0)}$), GM (1, 1, a), and

GM (1, 1, b) outperformed other models. When predicting PM_{10}, the minimum MAPE, MSE, RMSE, and maximum R was 14.10%, 25.62, 5.06, and 0.96, respectively. When predicting $PM_{2.5}$, the minimum MAPE, MSE, RMSE, and maximum R value of 15.24%, 11.57, 3.40, and 0.93, respectively, could be achieved. According to the results, it revealed that GM (1, 1) could predict the hourly PM variation even comparing with MLR and ANN. Additionally, GM (1, 1) was an efficiently early warning tool for providing PM information to the inhabitants.

References

Cunningham, W. P., & Cunningham, M. A. (2006). *Principles of environmental science. Inquiry & applications*. New York: McGraw-Hill.

Deng, J. (2002). *The foundation of grey theory*. Wuhan: Huazhang University of Science and Technology Press.

Deng, J. (2005). *The primary methods of grey system theory*. Wuhan: Huazhang University of Science and Technology Press.

Díaz-Robles, L. A., Ortega, J. C., Fu, J. S., Reed, G. D., Chow, J. C., Watson, J. G., et al. (2008). A hybrid ARIMA and artificial neural networks model to forecast particulate matter in urban areas: the case of Temuco, Chile. *Atmospheric Environment, 42*(35), 8331–8340.

Jaecker-Voirol, A., & Pelt, P. (2000). PM10 emission inventory in Ile de France for transport and industrial sources: PM10 re-suspension, a key factor for air quality. *Environmental Modelling and Software, 15*(6–7), 575–581.

Kolehmainen, M., Martikainen, H., & Ruuskanen, J. (2001). Neural networks and periodic components used in air quality forecasting. *Atmospheric Environment, 35*, 815–825.

Pai, T. Y., Hanaki, K., Ho, H. H., & Hsieh, C. M. (2007). Using grey system theory to evaluate transportation on air quality trends in Japan. *Transportation Research Part D: Transport and Environment, 12*(3), 158–166.

Pai, T. Y., Tsai, Y. P., Lo, H. M., Tsai, C. H., & Lin, C. Y. (2007). Grey and neural network prediction of suspended solids and chemical oxygen demand in hospital wastewater treatment plant effluent. *Computers & Chemical Engineering, 31*(10), 1272–1281.

Pai, T. Y., Chiou, R. J., & Wen, H. H. (2008). Evaluating impact level of different factors in environmental impact assessment for incinerator plants using GM (1, N) model. *Waste Management, 28*(10), 1915–1922.

Pai, T. Y., Chuang, S. H., Ho, H. H., Yu, L. F., Su, H. C., & Hu, H. C. (2008). Predicting performance of grey and neural network in industrial effluent using online monitoring parameters. *Process Biochemistry, 43*(2), 199–205.

Pai, T. Y., Chuang, S. H., Wan, T. J., Lo, H. M., Tsai, Y. P., Su, H. C., et al. (2008). Comparisons of grey and neural network prediction of industrial park wastewater effluent using influent quality and online monitoring parameters.

Environmental Monitoring and Assessment, 146(1–3), 51–66.

Pérez, P., Trier, A., & Reyes, J. (2000). Prediction of PM2.5 concentrations several hours in advance using neural networks in Santiago, Chile. *Atmospheric Environment, 34*(8), 1189–1196.

Pisoni, E., & Volta, M. (2009). Modeling Pareto efficient PM10 control policies in Northern Italy to reduce health effects. *Atmospheric Environment, 43*(20), 3243–3248.

Pope, C. A. (2000). What do epidemiological findings tell us about health effects of environmental aerosols? *Journal of Aerosol Medicine, 55*, 1350–1354.

Ryan, W. F. (1995). Forecasting severe ozone episodes in the Baltimore metropolitan area. *Atmospheric Environment, 29*(17), 2387–2398.

Shi, J. P., & Harrison, R. M. (1997). Regression modelling of hourly NO_x and NO_2 concentrations in urban air in London. *Atmospheric Environment, 31*(24), 4081–4094.

Slini, T., Kaprara, A., Karatzas, K., & Moussiopoulos, N. (2006). PM10 forecasting for Thessaloniki, Greece. *Environmental Modelling and Software, 21*, 559–565.

Wang, W., Lu, W., Wang, X., & Leung, A. Y. T. (2003). Prediction of maximum daily ozone level using combined neural network and statistical characteristics. *Environment International, 1049*, 1–8.