Atmospheric Pollution Research 6 (2015) 572-580

Atmospheric Pollution Research

www.atmospolres.com

Comparisons of GM (1, 1), and BPNN for predicting hourly particulate matter in Dali area of Taichung City, Taiwan

Li Chen [1], Tzu–Yi Pai [2]

[1] *General Education Center, Shuzen Junior College of Medicine and Management, Luju, Kaohsiung City, 82144, Taiwan, ROC*
[2] *Master Program of Environmental Education and Management, Department of Science Education and Application, National Taichung University of Education, Taichung City, 40306, Taiwan, ROC*

ABSTRACT

This paper represents the first study to compare seven types of first–order and one–variable grey differential equation model [abbreviated as GM (1, 1)] and back–propagation artificial neural network (BPNN) for predicting hourly particulate matter (PM) including PM_{10} and $PM_{2.5}$ concentrations in Dali area of Taichung City, Taiwan. Their prediction performance was also compared. The results indicated that the minimum mean absolute percentage error (MAPE), mean squared error (MSE), and root mean squared error (RMSE) was 16.76%, 132.95, and 11.53, respectively for PM_{10} prediction. For $PM_{2.5}$ prediction, the minimum MAPE, MSE, and RMSE value of 21.64%, 40.41, and 6.36, 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) could predict the hourly PM variation precisely even comparing with BPNN.

Keywords: *Grey system theory, GM (1, 1), hourly particulate matter, back–propagation neural network*

Corresponding Author:
Tzu–Yi Pai
☎ : *+886-4-22183541*
🖷 : *+886-4-22183540*
✉ : *bai@ms6.hinet.net*

Article History:
Received: 22 June 2014
Revised: 09 November 2014
Accepted: 27 December 2014

doi: 10.5094/APR.2015.064

1. Introduction

In the past twenty years, air pollution has been reduced in most cities in Western Europe, North America as well as Asia. Most air pollution improvement has resulted from better pollution–control technologies and efficiency in facilities, power plants, and other factories (Cunningham and Cunningham, 2008). Nevertheless, the events of serious air pollution are often reported in many countries.

For all air pollutants, the particulate matter (PM) concentrations are of particular concern, because high PM concentrations not only cause human health problems, but also deteriorate environmental quality (Pope, 2000; Kara et al., 2014). Odabasi et al. (2009) reported that PM emitted from industries contained iron, iron oxides, zinc, chromium, nickel, lead, cadmium, and other metals (and metal oxides). Epidemiological researches also show an association between ambient PM pollutants and negative effects on inhabitant health (Pope, 2000). Therefore, developing the rapid–responded prediction technology for providing air pollution information to the inhabitants becomes significantly important.

Traditionally, the atmospheric condition in one area is influenced by other area, complex interrelations from other administrative boundaries and various pollutants result in the prediction difficulty of atmospheric pollution data. Many attempts to predict atmospheric pollution have been implemented (Elbir, 2002; Elbir et al., 2010; Sofowote et al., 2014; Vanoye and Mendoza, 2014). For example, linear regression methods have been vastly utilized for several decades (Ryan, 1995; Shi and Harrison, 1997; Slini et al., 2006). In addition, to precisely predict complex, non–linear behaviors and chemical processes, back–propagation neural networks (BPNN) and fuzzy logic approach have been successfully applied because they can simulate nonlinear data well (Perez et al., 2000; Kolehmainen et al., 2001; Wang et al., 2003; Slini et al., 2006; Pai et al., 2009a; Pai et al., 2009b; Pai et al., 2011a; Pai et al., 2013a).

Although BPNN can predict pollutant concentrations successfully, it requires longer time for converging the solution and a large quantity of data for establishing a model. In order to simplify statistical complexity from the observation data for predicting atmospheric pollutants, the grey system theory (GST) is an applicable method.

GST specializes in the relational analysis, modeling, and prediction of the incomplete data and has been carried out in the previous studies (Deng, 2002; Deng, 2005; Pai et al., 2007a; Pai et al., 2007b; Pai et al., 2008a; Pai et al., 2008b; Pai et al., 2008c; Pai et al., 2011b; Pai et al., 2013b; Pai et al., 2014).

There are many analysis methods in GST including grey model (GM). GM can be used to establish the relationship between several sequences of data. One advantage of GM is that it is only a process to solve a simple regression and another one is that it can resolve the problem of small amount of data. If a more efficient predicting technology could be constructed, a better response strategy could be sought for emergency.

The objectives of this study are listed as follows: (1) to 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 Dali area of Taichung City, Taiwan, (2) to compare the prediction performance of seven types of GM (1, 1) model, (3) to employ BPNN for the prediction of PM for comparison in this study.

2. Materials and Methods

2.1. Data set

Dali area, where the Dari industry district stands, has been reported atmospherically polluted due to the growth of factories and vehicles over the past decades. There are many factories which are causing air pollution in Taichung City. The observation data from air quality monitoring station of Dali area in Taichung City was chosen for study (Figure 1). The concentrations of PM_{10} and $PM_{2.5}$ from 29th of July to 16th of August 2008 were collected. The reason why chose these data was that there was no typhoon occurred during this period. They were automatically sampled for analyzing every hour and the total number was 456. For all samples, 384 samples were utilized to determine the parameters of GM (1, 1) and 72 samples were utilized as the observation data when evaluating the performance of GM (1, 1) and BPNN. The number of training data was about 5 times as that of test data. The mean value of PM_{10} and $PM_{2.5}$ was 42.11 and 29.51 μg m^{-3}, respectively. The standard deviation of both PM was 18.42 and 14.71 μg m^{-3}, respectively. GM (1, 1) simply adopts the previous (historic) data to predict the future data of the air pollution time series. The influence of meteorological conditions is contained implicitly and naturally.

2.2. Grey modeling process

When information is insufficient, GM can be created to describe the behavior of the system using fewer (at least 4) data (Deng, 2002; Deng, 2005). By implementing accumulated generating operation (AGO), the chaotic data may behave exponentially such that a first–order differential equation can be utilized to describe the system behavior. The analytic solution of the differential equation will yield a time response equation for prediction. By means of inverse accumulated generating operation (IAGO), the prediction can be transformed into the sequence of original series. Following steps describe the grey modeling process.

Assume that a data series with n observations is shown as:

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

where, the superscript (0) of $X^{(0)}$ represents the original series. Let $X^{(1)}$ be the first–order AGO of $X^{(0)}$, whose elements are generated from $X^{(0)}$:

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

where, $x^{(1)}(k) = \sum_{i=1}^{k} x^{(0)}(i)$, for $k = 1,2,\ldots,n$. If the operation of AGO continues, the r–order AGO series, $X^{(r)}$, will be yielded as:

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

where, $x^{(r)}(k) = \sum_{i=1}^{k} x^{(r-1)}(i)$, for $k = 1,2,\ldots,n$. IAGO represents the inverse operation of AGO. IAGO transforms the AGO–operational series back to a lower order series. The IAGO for the first–order series is operated as follows: $x^{(0)}(1)=x^{(1)}(1)$ and $x^{(0)}(k)=x^{(1)}(k)–x^{(1)}(k–1)$ for $k=2,3,\ldots,n$. The tendency of AGO resembles an exponential function. Thus, the grey model GM (1, 1) utilizes a first order differential equation to fit the series with AGO operation,

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

where, the coefficient *a* is the developing coefficient and *b* is the grey input. The coefficients a and b will determine the predicting trend and interception of Equation (4). In accordance with the definition, GM (1, 1) is the grey model that the order in grey differential equation is equal to 1 and defined as follows:

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

where, $z^{(1)}(k)=0.5x^{(1)}(k\text{-}1)+0.5x^{(1)}(k)$ $k=2,3,4,\ldots,n$. Expanding Equation (5), yielding,

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

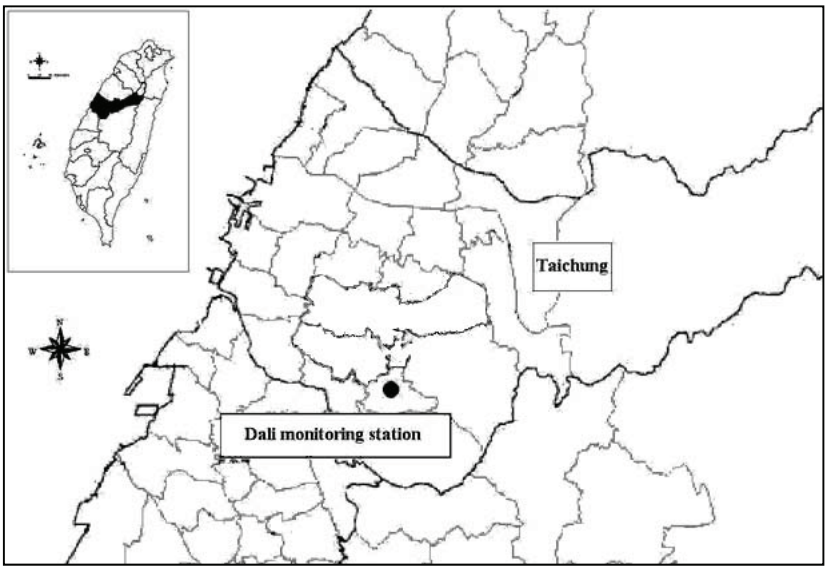

Figure 1. *Dali area.*

Transforming Equation (6) into matrix form yielding,

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

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

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

Subsequently, the whitening type of GM (1, 1) model (or in terms of GM (1, 1, *W*)) is described as:

$$\hat{x}_1^{(1)}(k+1) = \left(x^{(0)}(1) - \frac{b}{a}\right).e^{ak} + \frac{b}{a} \quad (8)$$

$$\hat{x}^{(0)}(k+1) = \hat{x}^{(1)}(k+1) - \hat{x}^{(1)}(k) \quad (9)$$

In addition, there are still several types of GM (1, 1) models derived from Equation (4) as follows.

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

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

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

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

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

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

$$x^{(0)}(k) = (1-a)x^{(0)}(k-1) \quad (12)$$

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

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

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

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

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

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

When utilizing GM (1, 1, $x^{(0)}$), GM (1, 1, a), GM (1, 1, b), and GM (1, 1, e), $x^{(0)}(2)$ can be estimated as follows:

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

All seven types of GM (1, 1) models and their denotation are summarized in Table 1. Details for derivation of these GM (1, 1) model can be referred to the references (Deng, 2002; Deng, 2005; Pai et al., 2007a; Pai et al., 2007b; Pai et al., 2008a; Pai et al., 2008b; Pai et al., 2008c; Pai et al., 2011b; Pai et al., 2013b; Pai et al., 2014).

Table 1. *Seven types of GM (1, 1) models*

Type	Denotation	Prediction equation
Whitening type	GM (1, 1, *W*)	$\hat{x}_1^{(1)}(k+1) = \left(x^{(0)}(1) - \frac{b}{a}\right).e^{-ak} + \frac{b}{a}$ $\hat{x}^{(0)}(k+1) = \hat{x}^{(1)}(k+1) - \hat{x}^{(1)}(k)$
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, $x^{(1)}$)	$x^{(0)}(k) = \beta - \alpha x^{(1)}(k-1)$ $\beta = \frac{b}{1+0.5a}, \quad \alpha = \frac{a}{1+0.5a}$
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)	$x^{(0)}(k) = x^{(0)}(3)e^{(k-3)\ln(1-\alpha)}$ $x^{(0)}(2) = \beta - \alpha x^{(0)}(1)$

2.3. Brief description on BPNN

The artificial neural network (ANN) simulates the important operation features of human nervous system to determine solutions by using information gained from historic data (Pai et al., 2009a; Pai et al., 2009b; Pai et al., 2011a; Pai et al., 2013a). To operate like a human brain, ANN uses many computational units called artificial neurons that are interrelated by various weight functions. Although each neuron can only perform a simple computation, an ANN can perform complicated calculations based on the multiple level structure of a network of connected neurons. An ANN is composed mainly of three independent layers: input, hidden, and output layers. Each layer contains many operation neurons. Input neurons accept the input values that are fed to the ANN, meanwhile the computational values in the output layer are determined by the output neurons. The hidden layers act as interfaces to relate input and output layers. Each neuron is linked to every neuron in adjacent layers by a weight function. Each neuron sums all of the values from previous inputs converts the sum to an output value. To a prediction problem, a supervised learning algorithm is often utilized to train ANN. The back propagation algorithm is commonly selected to direct ANN, i.e. BPNN. The steepest gradient descent method is commonly used to minimize the errors between the BPNN outputs and observations. The calculation of both GM (1, 1) and BPNN was also carried out using MATLAB.

3. Results and Discussion

3.1. Determination of grey parameters

The observation of PM_{10} and $PM_{2.5}$ were substituted into Equation (6) and the coefficients were determined by solving Equation (7). For PM_{10}, parameters a and b were equal to –0.00011973 and 41.236, respectively. For $PM_{2.5}$, a=–0.00021118 and b=28.379. Since the coefficients a and b represent the predicting trend and interception of Equation (4), the positive or negative values of a and b were determined by the characteristics of data set.

3.2. Determination of BPNN

The appropriate BPNN model was also designed to compare with GM (1, 1). Both BPNN models were comprised of three layers: input, hidden, and output layers. According to previous study (Pai et al., 2011b; Pai et al., 2013b), the GM (1, 1) equations of linear types outperformed others, thus the linear function was selected as the transfer function between the input, hidden, and output layers for BPNN for comparison. The hidden layer consisted of 3 and 20 operating neurons for PM_{10} and $PM_{2.5}$, respectively. The training epochs were 1 000.

3.3. Simulation of PM_{10}

Figure 2 illustrates the prediction results of PM_{10} using seven types of GM (1, 1) model and BPNN.

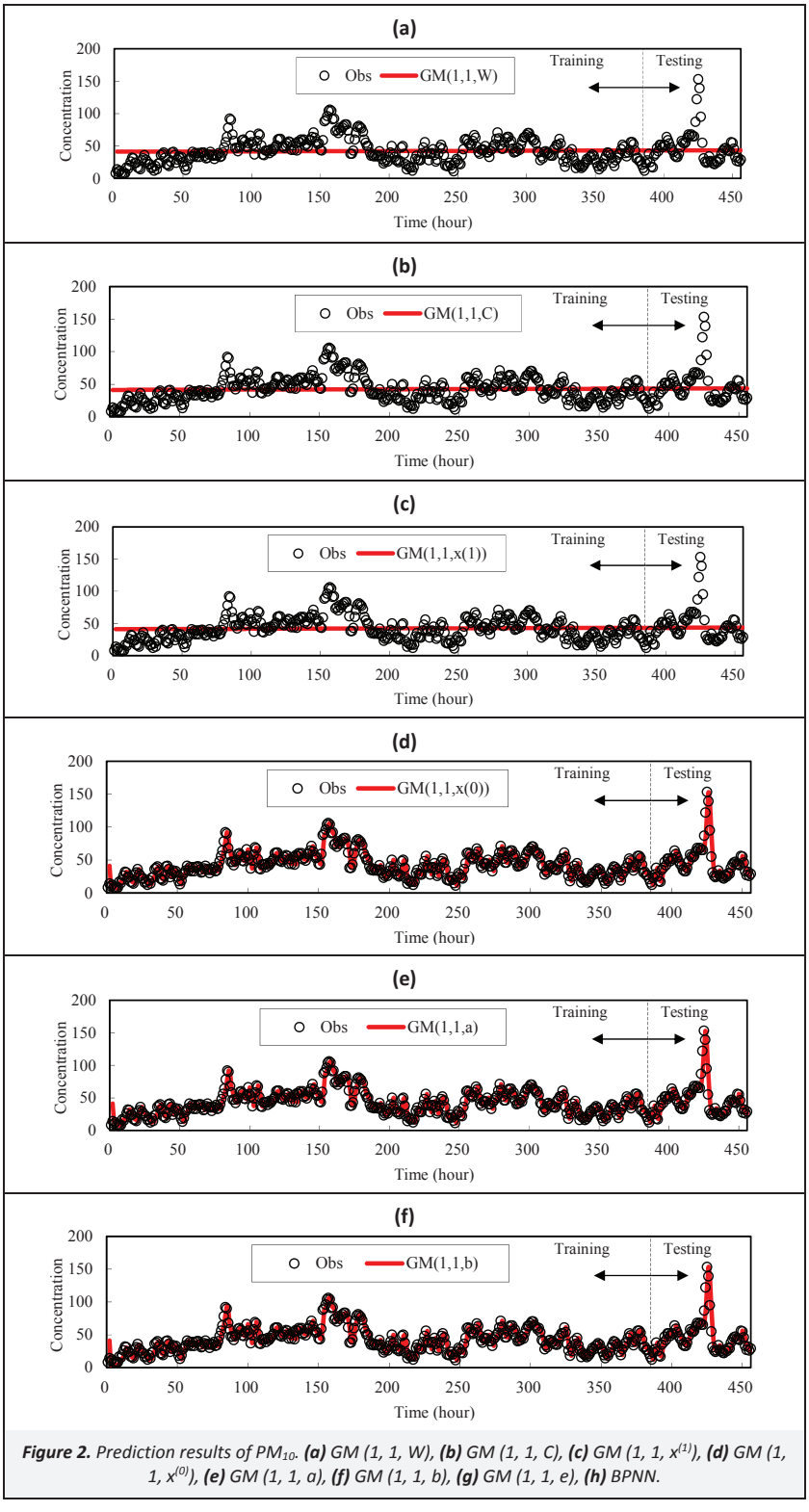

Figure 2. *Prediction results of PM_{10}.* ***(a)*** *GM (1, 1, W),* ***(b)*** *GM (1, 1, C),* ***(c)*** *GM (1, 1, $x^{(1)}$),* ***(d)*** *GM (1, 1, $x^{(0)}$),* ***(e)*** *GM (1, 1, a),* ***(f)*** *GM (1, 1, b),* ***(g)*** *GM (1, 1, e),* ***(h)*** *BPNN.*

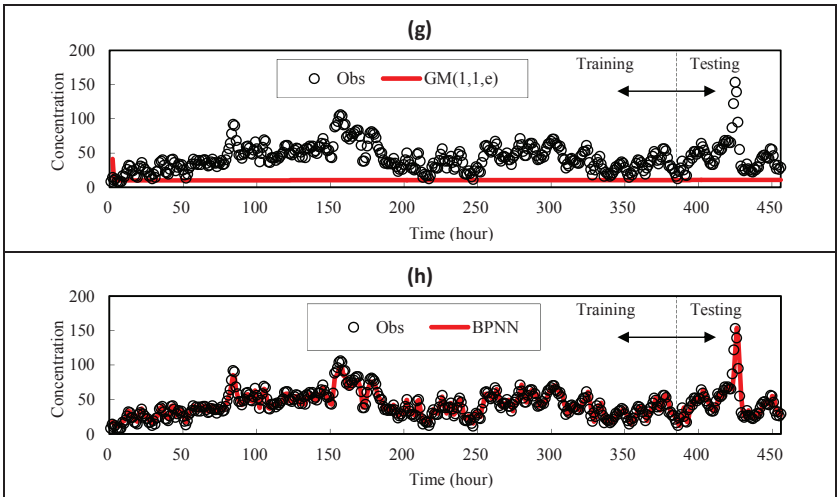

Figure 2. *Continued.*

In order to evaluate the prediction accuracy of GM (1, 1) and BPNN, the mean absolute percentage error (MAPE), mean squared error (MSE), and root mean squared error (RMSE) were adopted, given by:

$$MAPE = \frac{1}{N}\sum_{k=1}^{N}\left|\frac{x^{(0)}(k) - \hat{x}^{(0)}(k)}{x^{(0)}(k)}\right| \times 100\% \quad (17)$$

$$MSE = \frac{1}{N}\sum_{k=1}^{N}(x^{(0)}(k) - \hat{x}^{(0)}(k))^2 \quad (18)$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{k=1}^{N}(x^{(0)}(k) - \hat{x}^{(0)}(k))^2} \quad (19)$$

where N is the number of data, $x^{(0)}(k)$ is the observation value, $\hat{x}^{(0)}(k)$ is the prediction value.

All the values of MAPE, MSE, and RMSE are shown in Table 2. As shown in Table 2, when training, MAPEs of PM_{10} were between 16.19% and 16.49% using GM (1, 1, $x^{(0)}$), GM (1, 1, a), and GM (1, 1, b), but they were 47.59–70.63% using other GM (1, 1) models. When predicting, the MAPEs lay between 16.76% and 16.78% employing GM (1, 1, $x^{(0)}$), GM (1, 1, a), and GM (1, 1, b), but they were between 42.62% and 71.28% when using others.

The MSE values of 53.82–53.84 using GM (1, 1, $x^{(0)}$), GM (1, 1, a), and GM (1, 1, b) were better than those of 335.77–1 359.60 using other GMs when training. When predicting, the values of 132.95–133.29 using GM (1, 1, $x^{(0)}$), GM (1, 1, a), and GM (1, 1, b) were also better than those of 644.68–1 900.80 using other GMs. For training, the RMSE of 7.34 using GM (1, 1, $x^{(0)}$), GM (1, 1, a), and GM (1, 1, b) were lower than those of 18.32–36.87 using other GMs. The RMSE values of 11.53–11.55 using GM (1, 1, $x^{(0)}$), GM (1, 1, a), and GM (1, 1, b) were also lower than those of 25.39–43.60 using other GMs when predicting.

In the structure of GM (1, 1, $x^{(0)}$), GM (1, 1, a), and GM (1, 1, b), the point at time k is highly affected 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) exactly follow the observed pattern.

The prediction results of BPNN are also shown in Table 2. When constructing model, the MAPE of GM (1, 1, $x^{(0)}$) and GM (1, 1, a) was lower than those of BPNN. When predicting, the MAPE of GM (1, 1, $x^{(0)}$) and GM (1, 1, a) was higher than those of BPNN, but their MAPE values were very close.

The calculation time was shown in Table 3. The calculation time was 0.049 seconds when using seven types of GM (1, 1), but that of BPNN was 10.465 seconds.

Analogous observations were made by Slini et al. (2006). Slini et al. (2006) employed principal component analysis (PCA), classification and regression trees (CART), linear regression analysis (LRA), and ANN to predict daily PM_{10} concentrations. The RMSE values for PCA, CART, LRA, and ANN were 8.142, 33.55, 11.236, and 7.126, respectively.

Diaz–Robles et al. (2008) used Box–Jenkins time series (ARIMA) model, ANN, multiple linear regression (MLR), and a hybrid ARIMA–ANN model to forecast PM in urban areas. The RMSE values for ARIMA, ANN, MLR, and hybrid model were 28.46, 28.57, 28.39, and 8.80, respectively.

In this study, the RMSE values of 11.53–11.55 using GM (1, 1, $x^{(0)}$), GM (1, 1, a), and GM (1, 1, b) were obtained for predicting hourly values of PM_{10}.

3.4. Simulation of $PM_{2.5}$

Figure 3 shows the prediction results of $PM_{2.5}$. All statistical values revealed that the performance of GM (1, 1, $x^{(0)}$), GM (1, 1, a), and GM (1, 1, b) outperformed other models.

As shown in Table 4, the training MAPEs values of $PM_{2.5}$ were between 17.42% and 17.89% using GM (1, 1, $x^{(0)}$), GM (1, 1, a), and GM (1, 1, b), but they were 62.37–77.20% using other GMs. The predicting MAPEs lay between 21.64% and 21.67% when employing GM (1, 1, $x^{(0)}$), GM (1, 1, a), and GM (1, 1, b), but they were between 60.90% and 74.55% when using other GMs.

The training MSE values of 25.32–25.52 using GM (1, 1, $x^{(0)}$), GM (1, 1, a), and GM (1, 1, b) were better than those of 214.79–808.72 using other GMs. The predicting MSEs of 40.41–40.51 using GM (1, 1, $x^{(0)}$), GM (1, 1, a), and GM (1, 1, b) were also better than those of 167.47–610.88 from other GMs. The training RMSE values of 5.03–5.05 obtained from GM (1, 1, $x^{(0)}$), GM (1, 1, a), and GM (1, 1, b) were lower than those of 14.66–28.44 from other GMs. The predicting RMSE values of 6.36 from GM (1, 1, $x^{(0)}$), GM (1, 1, a), and GM (1, 1, b) were also lower than those of 12.94–24.72 from other GMs.

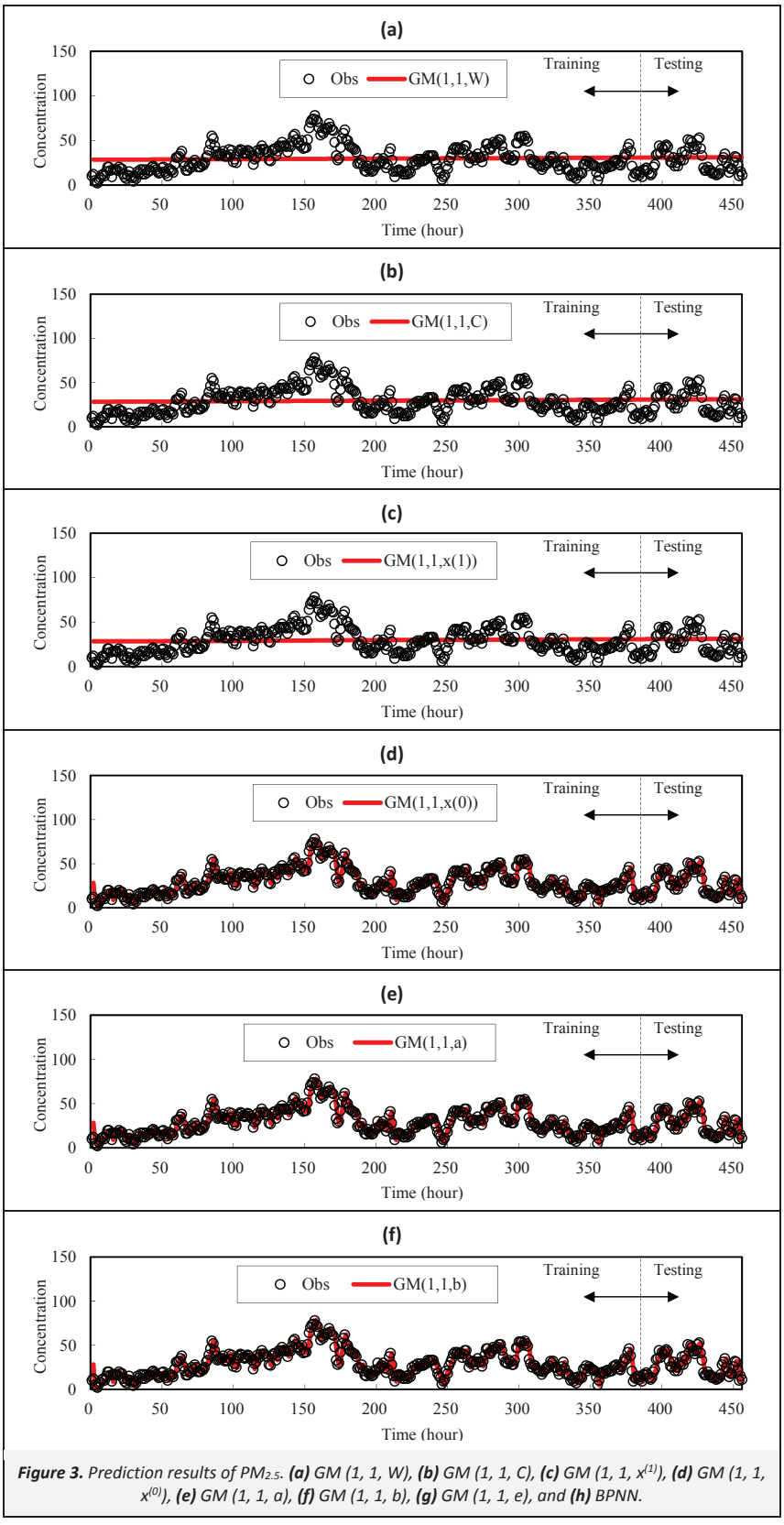

Figure 3. *Prediction results of $PM_{2.5}$.* ***(a)*** *GM (1, 1, W),* ***(b)*** *GM (1, 1, C),* ***(c)*** *GM (1, 1, $x^{(1)}$),* ***(d)*** *GM (1, 1, $x^{(0)}$),* ***(e)*** *GM (1, 1, a),* ***(f)*** *GM (1, 1, b),* ***(g)*** *GM (1, 1, e), and* ***(h)*** *BPNN.*

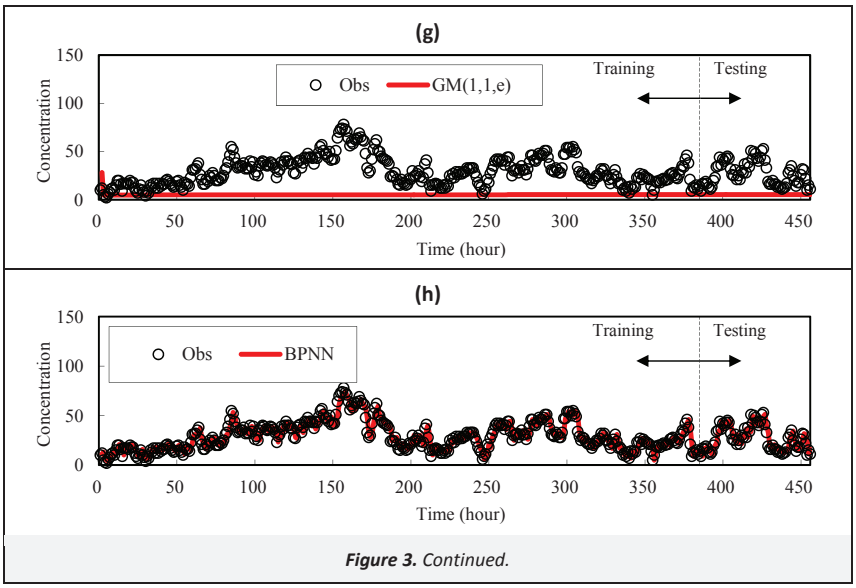

Figure 3. *Continued.*

Table 2. *The performance for PM_{10} using seven types of GM (1, 1) models and BPNN*

	MAPE		MSE		RMSE	
	Training	Testing	Training	Testing	Training	Testing
GM (1, 1, W)	47.60	42.63	335.77	644.68	18.32	25.39
GM (1, 1, C)	47.59	42.62	335.77	644.71	18.32	25.39
GM (1, 1, $x^{(1)}$)	47.60	42.65	335.82	644.72	18.33	25.39
GM (1, 1, $x^{(0)}$)	16.19	16.78	53.84	133.29	7.34	11.55
GM (1, 1, a)	16.19	16.78	53.84	133.29	7.34	11.55
GM (1, 1, b)	16.49	16.76	53.82	132.95	7.34	11.53
GM (1, 1, e)	70.63	71.28	1 359.60	1 900.80	36.87	43.60
BPNN	16.76	16.61	51.30	127.64	7.16	11.30

Table 3. *The calculation time for PM_{10} and $PM_{2.5}$ using seven types of GM (1, 1) models and BPNN*

	PM_{10}	$PM_{2.5}$
GM (1, 1, W)	0.049	0.021
GM (1, 1, C)	0.049	0.021
GM (1, 1, $x^{(1)}$)	0.049	0.021
GM (1, 1, $x^{(0)}$)	0.049	0.021
GM (1, 1, a)	0.049	0.021
GM (1, 1, b)	0.049	0.021
GM (1, 1, e)	0.049	0.021
BPNN	10.465	15.501

The performance of BPNN is also shown in Table 4. The training MAPE of GM (1, 1, $x^{(0)}$) and GM (1, 1, a) was lower than those of BPNN. The predicting MAPE of GM (1, 1, $x^{(0)}$) and GM (1, 1, a) was higher than those of BPNN, but the MAPE values were very close, too.

The calculation time for $PM_{2.5}$ was also shown in Table 3. The calculation time was 0.021 seconds when using seven types of GM (1, 1), but that of BPNN was 15.501 seconds for $PM_{2.5}$.

Perez et al. (2000) used linear perceptron model, ANN, and persistence model to forecast $PM_{2.5}$ hourly mean concentrations. The MAPE values were between 20% and 80% for these three types of model. In our study, the MAPEs lay between 21.64% and 21.67% for predicting hourly $PM_{2.5}$ concentrations using GM (1, 1, $x^{(0)}$), GM (1, 1, a), and GM (1, 1, b).

In accordance with the results, the GM (1, 1) model could result in high predictability. Besides, the coefficient calculation in GM (1, 1) model was only a process to solve a simple regression. Besides, the atmosphere condition in Dali is indeed influenced by other area and meteorological parameters. The data characteristic of the atmosphere condition in Dali reveals the influence from other area and meteorological parameters. GM (1, 1) simply adopts the previous (historic) data to predict the future data of the air pollution time series. The influence from other administrative boundary and meteorological conditions is contained implicitly and naturally. Therefore, GM could be applied successfully in predicting PM even comparing with BPNN.

4. Conclusions

The hourly PM_{10} and $PM_{2.5}$ concentrations in Dali area of Taichung City were predicted using seven types of GM (1, 1) models and BPNN. The conclusions can be drawn as follows. For PM_{10}, the minimum MAPE, MSE, and RMSE were 16.76%, 132.95, and 11.53, respectively when predicting. For $PM_{2.5}$, the minimum MAPE, MSE, and RMSE value of 21.64%, 40.41, and 6.36, respectively could be achieved for prediction. The predicting performance of GM (1, 1, $x^{(0)}$), GM (1, 1, a), and GM (1, 1, b) outperformed other models because of their equation structures. It revealed that GM (1, 1) could successfully predict the hourly PM variation even comparing with BPNN. The prediction performance of different data size and long–term prediction can be discussed in the future studies.

Table 4. *The performance for $PM_{2.5}$ using seven types of GM (1, 1) models and BPNN*

	MAPE		MSE		RMSE	
	Training	Testing	Training	Testing	Training	Testing
GM (1, 1, W)	62.39	60.98	214.79	167.75	14.66	12.95
GM (1, 1, C)	62.37	60.96	214.79	167.69	14.66	12.95
GM (1, 1, $x^{(1)}$)	62.39	60.90	214.85	167.47	14.66	12.94
GM (1, 1, $x^{(0)}$)	17.89	21.67	25.52	40.51	5.05	6.36
GM (1, 1, a)	17.89	21.67	25.52	40.51	5.05	6.36
GM (1, 1, b)	17.42	21.64	25.32	40.41	5.03	6.36
GM (1, 1, e)	77.20	74.55	808.72	610.88	28.44	24.72
BPNN	17.63	21.40	24.19	39.14	4.92	6.26

Acknowledgment

The authors are grateful to the Ministry of Science and Technology of R.O.C. for financial support under the grant number MOST 103–2621–M–142 –001.

References

Cunningham, W.P., Cunningham, M.A., 2008. *Principles of Environmental Science: Inquiry & Applications*, McGraw–Hill Company, New York.

Deng, J., 2005. *The Primary Methods of Grey System Theory*, Huazhang University of Science and Technology Press, Wuhan, China.

Deng, J., 2002. *The Foundation of Grey Theory*, Huazhang University of Science and Technology Press, Wuhan, China.

Diaz–Robles, L.A., Ortega, J.C., Fu, J.S., Reed, G.D., Chow, J.C., Watson, J.G., Moncada–Herrera, J.A., 2008. A hybrid ARIMA and artificial neural networks model to forecast particulate matter in urban areas: The case of Temuco, Chile. *Atmospheric Environment* 42, 8331–8340.

Elbir, T., 2002. Application of an ISCST3 model for predicting urban air pollution in the Izmir metropolitan area. *International Journal of Environment and Pollution* 18, 498–507.

Elbir, T., Mangir, N., Kara, M., Simsir, S., Eren, T., Ozdemir, S., 2010. Development of a GIS–based decision support system for urban air quality management in the city of Istanbul. *Atmospheric Environment* 44, 441–454.

Kara, M., Dumanoglu, Y., Altiok, H., Elbir, T., Odabasi, M., Bayram, A., 2014. Spatial distribution and source identification of trace elements in topsoil from heavily industrialized region, Aliaga, Turkey. *Environmental Monitoring and Assessment* 186, 6017–6038.

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

Odabasi, M., Bayram, A., Elbir, T., Seyfioglu, R., Dumanoglu, Y., Bozlaker, A., Demircioglu, H., Altiok, H., Yatkin, S., Cetin, B., 2009. Electric arc furnaces for steel–making: Hot spots for persistent organic pollutants. *Environmental Science & Technology* 43, 5205–5211.

Pai, T.–Y., Lo, H.–M., Wan, T.–J., Chen, L., Hung, P.–S., Lo, H.–H., Lai, W.–J., Lee, H.–Y., 2014. Predicting air pollutant emissions from a medical incinerator using grey model and neural network. *Applied Mathematical Modelling*, article in press.

Pai, T.Y., Hanaki, K., Su, H.C., Yu, L.F., 2013a. A 24–h forecast of oxidant concentration in Tokyo using neural network and fuzzy learning approach. *CLEAN–Soil Air Water* 41, 729–736.

Pai, T.Y., Hanaki, K., Chiou, R.J., 2013b. Forecasting hourly roadside particulate matter in Taipei County of Taiwan based on first–order and one–variable grey model. *CLEAN–Soil Air Water* 41, 737–742.

Pai, T.Y., Yang, P.Y., Wang, S.C., Lo, M.H., Chiang, C.F., Kuo, J.L., Chu, H.H., Su, H.C., Yu, L.F., Hu, H.C., Chang, Y.H., 2011a. Predicting effluent from the wastewater treatment plant of industrial park based on fuzzy network and influent quality. *Applied Mathematical Modelling* 35, 3674–3684.

Pai, T.Y., Ho, C.L., Chen, S.W., Lo, H.M., Sung, P.J., Lin, S.W., Lai, W.J., Tseng, S.C., Ciou, S.P., Kuo, J.L., Kao, J.T., 2011b. Using seven types of GM (1, 1) model to forecast hourly particulate matter concentration in Banciao city of Taiwan. *Water Air and Soil Pollution* 217, 25–33.

Pai, T.Y., Wan, T.J., Hsu, S.T., Chang, T.C., Tsai, Y.P., Lin, C.Y., Su, H.C., Yu, L.F., 2009a. Using fuzzy inference system to improve neural network for predicting hospital wastewater treatment plant effluent. *Computers & Chemical Engineering* 33, 1272–1278.

Pai, T.Y., Wang, S.C., Chiang, C.F., Su, H.C., Yu, L.F., Sung, P.J., Lin, C.Y., Hu, H.C., 2009b. Improving neural network prediction of effluent from biological wastewater treatment plant of industrial park using fuzzy learning approach. *Bioprocess and Biosystems Engineering* 32, 781–790.

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

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

Pai, T.Y., Chuang, S.H., Wan, T.J., Lo, H.M., Tsai, Y.P., Su, H.C., Yu, L.F., Hu, H.C., Sung, P.J., 2008c. Comparisons of grey and neural network prediction of industrial park wastewater effluent using influent quality and online monitoring parameters. *Environmental Monitoring and Assessment* 146, 51–66.

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

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

Perez, P., Trier, A., Reyes, J., 2000. Prediction of $PM_{2.5}$ concentrations several hours in advance using neural networks in Santiago, Chile. *Atmospheric Environment* 34, 1189–1196.

Pope, C.A., 2000. What do epidemiologic findings tell us about health effects of environmental aerosols? *Journal of Aerosol Medicine–Deposition Clearance and Effects in the Lung* 13, 335–354.

Ryan, W.F., 1995. Forecasting severe ozone episodes in the Baltimore metropolitan–area. *Atmospheric Environment* 29, 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, 4081–4094.

Slini, T., Kaprara, A., Karatzas, K., Moussiopoulos, N., 2006. PM_{10} forecasting for Thessaloniki, Greece. *Environmental Modelling & Software* 21, 559–565.

Sofowote, U.M., Rastogi, A.K., Debosz, J., Hopke, P.K., 2014. Advanced receptor modeling of near–real–time, ambient $PM_{2.5}$ and its associated components collected at an urban–industrial site in Toronto, Ontario. *Atmospheric Pollution Research* 5, 13–23.

Vanoye, A.Y., Mendoza, A., 2014. Application of direct regularization techniques and bounded–variable least squares for inverse modeling of an urban emissions inventory. *Atmospheric Pollution Research* 5, 219–225.

Wang, W.J., Lu, W.Z., Wang, X.K., Leung, A.Y.T., 2003. Prediction of maximum daily ozone level using combined neural network and statistical characteristics. *Environment International* 29, 555–562.