Predictor Design of a Novel Grey Model PGM21
Using Pseudo Second-Order Information*

Keh-Tsong LI** and Yon-Ping CHEN***

A novel grey model, called pseudo-GM(2,1) or PGM21 for simplicity, is developed to improve the prediction accuracy for a non-monotone sequence by modifying the GM(1,1) model with a remedial term. Since the remedial term consists of two development coefficients, it can be treated as a kind of second-order information, called the pseudo second-order information in this paper. With the remedial term, the PGM21 only costs a little higher computation time than the original GM(1,1) model; however, its prediction accuracy is highly increased. The important features of the PGM21 are demonstrated by numeric simulation results.

Key Words: Grey Model, Prediction Accuracy, Non-Monotone Sequence, Pseudo Second-Order Information

1. Introduction

In 1982, Professor Deng first proposed the grey system theory to deal with systems possessing poor and incomplete information(1)-(2). The grey system theory is primarily classified into two categories: grey relational analysis and grey prediction. Up to now, the grey prediction had been successfully applied to diverse fields, such as earthquakes, industry, economics, and control(3)-(11).

The first-order single-variable grey model, denoted as GM(1,1), is the most popular approach for prediction. In many literatures(8)-(11), the GM(1,1) was applied to the controller design as a predictor. Although the GM(1,1) model takes the advantages of simplicity and quickness, the limitation of its prediction accuracy is still arguable. To further improve the prediction accuracy, some investigators have proposed their modified versions of the GM(1,1), such as Yeh and Lu(12) and Lin and Hsu(13). In addition, some investigators have paid their attentions to the optimization of the background value of the GM(1,1) in order to improve its prediction accuracy(14),(15). Actually, these modified versions only improve the prediction accuracy on a small scale, especially for a non-monotone sequence of data. It is because they are still based on the GM(1,1) model.

Intuitively, in order to increase the prediction accuracy, a higher order grey model, like the GM(2,1), is required. Unfortunately, the use of a higher order grey model implies a much more complicated algorithm and an increase of the computation time. In this paper, a novel grey model, called the pseudo-GM(2,1) or PGM21 for simplicity, is developed by adding a remedial term, which is related to the difference between the development coefficients of two latest consecutive subsequences modeled by the original GM(1,1). Since the remedial term consists of two development coefficients, it is second-order-like, called the pseudo second-order information in this paper. Significantly, the PGM21 model not only highly reduces the prediction error but also costs only a little higher computation time than the GM(1,1) model.

Next, the conventional GM(1,1) model will be briefly introduced in section 2 and the design concept and procedure of the modified version PGM21 model will be presented in section 3. By using the numeric simulation, the features of the PGM21 model are shown in section 4. Finally, section 5 gives the concluding remarks.

2. First-Order Single-Variable Grey Model GM(1,1)

This section will briefly introduce the first-order single-variable grey model, denoted as GM(1,1), which is adopted to obtain the predictive data following a given positive data sequence. It is known that three fundamental operations are required to establish the GM(1,1) model, which are the accumulated generating operation (AGO), the mean operation (MEAN), and the inverse accumulated operation (IAGO). Let the positive data sequence be given.
where $x^{(0)}(k) > 0$, $k = 1, 2, \ldots, n$ and $n \geq 4(2)$. The three fundamental operations are then defined as

- **AGO** - $x^{(1)}(k) = \sum_{i=1}^{k} x^{(0)}(l), \quad k = 1, 2, \ldots, n \quad (2)$
- **MEAN** - $z^{(1)}(k) = \frac{1}{2} [x^{(1)}(k) + x^{(1)}(k - 1)], \quad k = 2, 3, \ldots, n \quad (3)$
- **IAGO** - $x^{(0)}(1) = x^{(1)}(1), \quad x^{(0)}(k) = x^{(1)}(k) - x^{(1)}(k - 1), \quad k = 2, 3, \ldots, n \quad (4)$

Clearly, two new positive data sequences $x^{(1)}(k)$ and $z^{(1)}(k)$ are generated from the original sequence (1). Since $x^{(1)}(k)$ accumulates the data from $x^{(0)}(1)$ to $x^{(0)}(k)$, it is easy to verify that $x^{(1)}(k) > x^{(1)}(k - 1)$ and $z^{(1)}(k) > z^{(1)}(k - 1)$. As to the IAGO operation, it recovers the data $x^{(0)}(k)$ from $x^{(1)}(k)$ via the inverse operation (4).

With these fundamental operations, the GM(1,1) model is commonly constructed as the following grey differential equation (1):

$$x^{(0)}(k) + ax^{(1)}(k) = b, \quad k = 2, 3, \ldots, n \quad (5)$$

where $a$ is called the development coefficient, $b$ is treated as the grey input and $z^{(1)}(k)$ is defined in (3). Both $a$ and $b$ are constant and unknown, which needs to be further determined. Rewriting (5) into a matrix form yields

$$\begin{bmatrix} y \\ \vdots \\ y^{(n)} \end{bmatrix} = \begin{bmatrix} x^{(0)}(2) \\ \vdots \\ x^{(0)}(n) \end{bmatrix}, \quad B = \begin{bmatrix} -z^{(1)}(2) & 1 \\ \vdots & \vdots \\ -z^{(1)}(n) & 1 \end{bmatrix}$$

where

$$a = \begin{bmatrix} a \\ b \end{bmatrix} = (B^T B)^{-1} B^T y \quad (7)$$

which completes the establishment of the grey differential equation (5). Theoretically, the GM(1,1) model employs (5) to imitate the first-order ordinary differential equation

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

which is treated as its whitening equation. By directly modifying the solution of (8), the term $x^{(1)}(k)$ is estimated as

$$\hat{x}^{(1)}(k) = \left(x^{(0)}(1) - \frac{b}{a}\right) e^{-ak} + \frac{b}{a}, \quad k = 1, 2, \ldots \quad (9)$$

Further using the IAGO in (4) yields

$$\hat{x}^{(0)}(k) = (1 - e^a) \left(x^{(0)}(1) - \frac{b}{a}\right) e^{-ak} + \frac{b}{a}, \quad k = 2, 3, \ldots \quad (10)$$

Clearly, $\hat{x}^{(0)}(k)$ for $k > n$ are the so-called predictive data of the sequence (1), which can be expressed as

$$\hat{x}^{(0)}(n+p) = (1 - e^a) \left(x^{(0)}(1) - \frac{b}{a}\right) e^{-a(n+p)} \quad (11)$$

For $p = 1$, the one-step-ahead predictive value could be obtained by

$$\hat{x}^{(0)}(n+1) = (1 - e^a) \left(x^{(0)}(1) - \frac{b}{a}\right) e^{-an} \quad (12)$$

which is the first predictive value coming after the sequence (1). Obviously, the resulted predictive values (11) are composed of the term $e^{-a(n+p)}$, which means the GM(1,1) model is suitable for predicting a sequence with a single exponential rate. For a sequence not in such case, such as a non-monotone sequence, a higher order grey model is needed to reduce the prediction error. Next, a novel grey model, called the PGM21, is introduced to fulfill a higher order prediction by simply modifying the GM(1,1) model.

3. Design of Pseudo Second-Order Grey Model PGM21

In this section, the restriction of the GM(1,1) model will be indicated and then a novel modified pseudo-second-order grey model, called pseudo-GM(2,1) or PGM21, will be proposed to improve the drawback.

First, for the resulted sequence $\hat{x}^{(0)}(k)$ in (10), let’s check the following ratio between two consecutive data

$$\frac{\hat{x}^{(0)}(k+1)}{\hat{x}^{(0)}(k)} = e^{-a}, \quad k \geq 2 \quad (13)$$

Clearly, the ratios for $k \geq 2$ are all the same and equal to $e^{-a}$. It shows that the sequence $\hat{x}^{(0)}(k)$ in (10) decreases or increases monotonously with an exponential rate $a$. In other words, the GM(1,1) model is mainly suitable for monotone sequences approximately possessing a single exponential rate. Unfortunately, most of the physical sequences are changeable and not of single exponential rate. This implies the GM(1,1) model may not well predict most of the physical sequences. Figure 1 depicts two sequences not possessing a single exponential rate, where

![Fig. 1 Two sequences not of a single exponential rate](image-url)
Sequence-(a) is non-monotone with one turning point and Sequence-(b) is monotone but with more than one exponential rates. To reduce their prediction errors, some investigators employ a higher order grey model and some others try to modify the original GM(1,1) model. Here, we will focus on the development of a modified GM(1,1) model, the PGM21 model, which is not only as simple as the GM(1,1) model but also allows the predictive data to possess two exponential rates similar to the GM(2,1) model.

In the fields of the series forecasting and the system control, the GM(1,1) model usually adopts the privileging fresh information rationale\(^{(i)}\). To put it more simply, the GM(1,1) model adopts the latest \(n\) data which have been observed. Let’s consider the following data sequence

\[
x^{(0)}(1), x^{(0)}(2), \ldots, x^{(0)}(m), \quad m > n \geq 4
\]  

(14)

where all the data are observed in order. When the first \(n\) data \(x^{(0)}(1), \ldots, x^{(0)}(n)\) are obtained, the GM(1,1) model starts to predict the data \(x^{(0)}(n+p), p = 1, 2, \ldots,\) coming after \(x^{(0)}(n)\). For convenience, the first \(n\) data are set to be \(x^{(0)} = \{x^{(0)}(1), \ldots, x^{(0)}(n)\}\) and the predictive data of \(x^{(0)}(n+p)\) are represented by \(\hat{x}^{(0)}(n+p), p = 1, 2, \ldots\). Next, after the datum \(x^{(0)}(n+1)\) is observed, the GM(1,1) model immediately applies to the latest \(n\) data \(x^{(0)}(2), \ldots, x^{(0)}(n+1)\), which are grouped as \(x^{(0)} = \{x^{(0)}(2), \ldots, x^{(0)}(n+1)\}\). Similarly, the predictive data of \(x^{(0)}(n+1+p)\) are represented by \(\hat{x}^{(0)}(n+1+p), p = 1, 2, \ldots\). Step by step, the GM(1,1) model is processed to the \(i\)-th set of latest data, denoted as

\[
x^{(0)} = (x^{(0)}(1), x^{(0)}(2), \ldots, x^{(0)}(n)), \quad i = 1, 2, \ldots, m - n + 1
\]  

(15)

where \(x^{(0)}(k) = x^{(0)}(k+i-1), k = 1, 2, \ldots, n\). The predictive data of \(x^{(0)}(n+i+p)\) are represented by \(\hat{x}^{(0)}(n+i+p), p = 1, 2, \ldots\). The above step-by-step process is the so-called rolling procedure.

According to the GM(1,1) model, the predictive data of the \(i\)-th subsequence \(x^{(i)}\) in (15), given as \(x^{(i)} = (x^{(i)}(1), x^{(i)}(2), \ldots, x^{(i)}(n+i-1))\), can be achieved from (11) and expressed as

\[
\hat{x}^{(i)}(n+p) = (1 - e^{-a_i}) \cdot \left( x^{(i)}(1) - \frac{b_i}{a_i} \right) e^{-a_i(n+p-1)},
\]  

(16)

where the development coefficient \(a_i\) and grey input \(b_i\) could be solved from (7) using the least-square method. Trace back to the \((i-1)\)-th sequence, i.e.,

\[
x^{(i-1)} = (x^{(i-1)}(1), x^{(i-1)}(2), \ldots, x^{(i-1)}(n+i-2))
\]  

(17)

whose predictive data are

\[
\hat{x}^{(i-1)}(n+p) = (1 - e^{-a_{i-1}}) \cdot \left( x^{(i-1)}(1) - \frac{b_{i-1}}{a_{i-1}} \right) e^{-a_{i-1}(n+p-1)},
\]  

(18)

Clearly, these two sequences have different development coefficients \(a_i\) and \(a_{i-1}\), even though there are \(n-1\) data, \(x^{(0)}(i), x^{(0)}(i+1), \ldots, x^{(0)}(i+n-2)\), overlapped in these two sequences. In case that \(a_i\) and \(a_{i-1}\) are not quite distinct, i.e., the difference between \(a_i\) and \(a_{i-1}\) is small, (16) and (18) will be good results for predicting \(x^{(i)}(0)\) and \(x^{(i-1)}(0)\). However, if the difference between \(a_i\) and \(a_{i-1}\) is increased to a certain level, then it reveals that \(x^{(i)}(0)\) and \(x^{(i)}(0)\) are at least related to two exponential rates. Intuitively, the GM(2,1) model should be a better choice for such situation. However, the GM(2,1) model is much more complicated than the GM(1,1) model. In order to keep the simplicity of the GM(1,1) model, here a modification is proposed on (16) as below:

\[
\hat{x}^{(0)}(n+p) = (1 - e^{-a_i}) \cdot \left( x^{(0)}(1) - \frac{b_i e^{-a_i(a_i-a_{i-1})}}{a_i} \right) e^{-a_i(n+p-1)},
\]  

(19)

which changes \(b_i\) in (16) into \(b_i e^{-a_i(a_i-a_{i-1})}\). Obviously, a remedial term \(e^{-a_i(a_i-a_{i-1})}\) is added to the grey input \(b_i\). As a result, the predictive data of \(x^{(i)}(0)\) could have two exponential rates \(a_i\) and \(a_{i-1}\) as expected. It is also true for \(x^{(i)}(0)\) since its predictive data are possessed of two exponential rates \(a_i\) and \(a_{i-1}\).

In addition to stressing the use of two exponential rates, the remedial term \(e^{-a_i(a_i-a_{i-1})}\) also emphasizes that the variation between \(x^{(i)}(0)\) and \(x^{(i)}(0)\) is compensated by the difference \((a_i-a_{i-1})\). Furthermore, it is noticed that the remedial term \(e^{-a_i(a_i-a_{i-1})}\) is only attached to \(b_i\) in (16) since \(b_i\) is the so-called grey input and could potentially consist of the information from the previous sequence \(x^{(i-1)}\). Clearly, a sub-term \(e^{-a_i(a_i-a_{i-1})}\) related to the former sequence \(x^{(i-1)}\) is included in the modified grey input \(b_i e^{-a_i(a_i-a_{i-1})}\). The correctness of employing the term \(e^{-a_i(a_i-a_{i-1})}\), not \(e^{-a_i(a_i-a_{i-1})}\), can be explained from the ratio in (13), rearranged as \(\hat{x}^{(i)}(k) = e^{\cdot \hat{x}^{(i)}(k+1)}\). Apparently, the datum \(\hat{x}^{(i)}(k)\) before \(\hat{x}^{(i)}(k+1)\) should be multiplied by \(e^{\cdot \hat{x}^{(i)}(k+1)}\). In other words, if a term achieves information from the previous sequence, then it may contain a term multiplied by \(e^{\cdot \hat{x}^{(i)}(k+1)}\). Hence, it is reasonable to add a remedial term multiplied by \(e^{\cdot \hat{x}^{(i)}(k+1)}\) to the grey input \(b_i\), which is assumed to have information from the former sequence \(x^{(i-1)}\).

Because the remedial term \(e^{-a_i(a_i-a_{i-1})}\) is composed of information concerning two exponential rates, it can be treated as consisting a kind of second-order information, or simply called the pseudo second-order information. With this pseudo second-order information, the GM(1,1) rolling model can be easily extended to the so-called pseudo-GM(2,1) or PGM21 for simplicity. The procedure of implementing PGM21 is now summarized as follows:

**Step 1:** Rearrange the sequence (14) into a rolling sequence as

\[
x^{(i)} = (x^{(i)}(1), x^{(i)}(2), x^{(i)}(n)),
\]  

\(i = 1, 2, \ldots, m - n + 1\)

where \(x^{(i)}(k) = x^{(i)}(k+i-1), k = 1, 2, \ldots, n,\) and \(n \geq 4\).
Step 2: Determine the predictive step \( p, p \geq 1 \).

Step 3: Apply the GM(1,1) model to \( x_1^{(0)} \) and obtain the predictive data as
\[
\hat{x}_1^{(0)}(n + p) = (1 - e^{a_1}) \cdot \left( x_1^{(0)}(1) - \frac{b_1}{a_1} \right) \cdot e^{-a_1(n + p - 1)}
\]
where \( a_1 \) and \( b_1 \) are calculated from (7).

Step 4: Let \( i = 2 \).

Step 5: Apply the GM(1,1) model to \( x_i^{(0)} \) and obtain the predictive data as
\[
\hat{x}_i^{(0)}(n + p) = (1 - e^{a_i}) \cdot \left( x_i^{(0)}(1) - \frac{b_i}{a_i} \right) \cdot e^{-a_i(n + p - 1)}
\]
where \( a_i \) and \( b_i \) are calculated from (7).

Step 6: Modify the predictive data of \( x_i^{(0)} \) into
\[
\hat{x}_i^{(0)}(n + p) = (1 - e^{a_i}) \cdot \left( x_i^{(0)}(1) - \frac{b_i e^{(a_i-a_{i-1})p}}{a_i} \right) \cdot e^{-a_i(n + p - 1)}
\]

Step 7: If \( i = m - n + 1 \), then stop, otherwise \( i = i + 1 \) and go to Step 5.

By adding the remedial term \( e^{-(a_i-a_{i-1})p} \) to the grey input, the PGM21 indeed contains two exponential rates and highly improves the prediction accuracy. Most significantly, when comparing to the original GM(1,1) model, the computation time of PGM21 is only increased a little in calculating the remedial term \( e^{-(a_i-a_{i-1})p} \). Some numeric examples will be used to demonstrate the advantages of the PGM21 in the next section.

4. Simulation Results

In this section, three cases are used as examples to illustrate the excellence of the PGM21 proposed in section 3. For comparison, the original GM(1,1) model described in section 2 also applies to these cases. The functions and predictive steps adopted for the three cases are listed as below:

Case 1 - \( x(t) = 2 + 3e^{-0.5t} \sin(2t) \) for \( 0 \leq t \leq 10 \) with \( p = 1 \).

Case 2 - \( x(t) = 4 + \sin(2t) + \cos(3t) \) for \( 0 \leq t \leq 10 \) with \( p = 1 \).

Case 3 - \( x(t) = 4 + \sin(2t) + \cos(3t) \) for \( 0 \leq t \leq 10 \) with \( p = 2 \).

These three functions are all positive and sampled every \( 0.1 \) sec. The sampled data are denoted as \( x(kT) \), where \( T = 0.1 \) sec and \( k = 1, 2, \ldots, 100 \). Clearly, each function results in a data sequence just like (14), where \( x^{(0)}(k) = x(kT) \) with \( k = 1, 2, \ldots, m \), and \( m = 100 \). Following the procedure shown in section 3, the numeric simulation results of PGM21 for Case 1, Case 2, and Case 3 are obtained and demonstrated in Fig. 2 to Fig. 7. For all these three cases, the rolling sequence is set to be \( x^{(0)} = (x^{(0)}_1(1), x^{(0)}_2(2), x^{(0)}_3(3), x^{(0)}_4(4)), i.e., n = 4 \). Besides, each case consists of two kinds of simulation results, which are the predictive data and prediction error. The prediction error is defined as

\[
e(k) = \frac{|x(kT) - \hat{x}^{(0)}(k)|}{x(kT)} = \frac{|x^{(0)}(k) - \hat{x}^{(0)}(k)|}{x^{(0)}(k)}
\]

which is the ratio of the difference \( |x(kT) - \hat{x}^{(0)}(k)| \) to the real data \( x^{(0)}(k) \).

In Figs. 2, 4 and 6, the predictive data show that the PGM21 is better than the original GM(1,1) model in all the three cases. Note that all the curves in these figures start from the moment that the first predictive data are obtained. Figures 3, 5 and 7 show the prediction errors. It could be found that both the original GM(1,1) model and the PGM21 have small prediction errors around the data ranges of monotonously increasing or decreasing. However, when the data sequence is in a trend of non-monotone, especially with a turning point, the prediction error of the original GM(1,1) becomes much larger than the PGM21. It demonstrates that the remedial term \( e^{-(a_i-a_{i-1})p} \) indeed aids the PGM21 to reduce the prediction errors of non-monotone data sequences.

In order to further show the effect caused by the predictive step, an average prediction error is defined as
From the above results, it could be found that $E_{\text{model}}^{\text{GM(1,1)}}_{\text{Case } i} < E_{\text{model}}^{\text{PGM21}}_{\text{Case } i}$ for different models. The truth of $E_{\text{model}}^{\text{GM(1,1)}}_{\text{Case } i} < E_{\text{model}}^{\text{PGM21}}_{\text{Case } i}$ for different models is related to the frequency of the functions in Case 1 and Case 2. Since the function in Case 2 changes faster, its average prediction error becomes larger than that of Case 1. The truth of $E_{\text{model}}^{\text{PGM21}}_{\text{Case } i} < E_{\text{model}}^{\text{GM(1,1)}}_{\text{Case } i}$ means the PGM21 is better than the original GM(1,1) model, which has been explained by Figs. 2, 4 and 6.

5. Conclusion

In this paper, a novel grey model PGM21 is developed by using a remedial term containing pseudo second-order information. With this remedial term, the computation time of the PGM21 only increases a little higher than the original GM(1,1) model. However, its prediction error is reduced on a large scale. The excellent predictive ability of the PGM21 has been demonstrated by the numeric simulation results.

Acknowledgements

This work was supported by the MOE Program for Promoting Academic Excellent of Universities under the grant number 91-E-FA06-4-4.

References