A Novel Weighted Fractional GM(1,1) Model and Its Applications

Qin-Qin Shen,1,2 Quan Shi,2 Tian-Pei Tang,2,3 and Lin-Quan Yao1

1School of Rail Transportation, Soochow University, Suzhou 215006, China
2School of Transportation and Civil Engineering, Nantong University, Nantong 226019, China
3Business School, University of Shanghai for Science and Technology, Shanghai 200093, China

Correspondence should be addressed to Quan Shi; shiquannt@sina.com and Lin-Quan Yao; lqyao@suda.edu.cn

Received 5 June 2019; Revised 16 October 2019; Accepted 25 October 2019; Published 20 January 2020

Academic Editor: Hassan Zargarzadeh

Copyright © 2020 Qin-Qin Shen et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Based on the ideas of the new information priority principle and the fractional-accumulation generating operator, in this paper we propose a novel weighted fractional GM(1,1) (WFGM(1,1)) prediction model. In the new model, the original sequence is first transformed by using the weighted fractional-accumulation generating operator, which involves two parameters. With special choices of these parameters, the proposed WFGM(1,1) model reduces to the classical GM(1,1) model and the fractional GM(1,1) (FGM(1,1)) model, as well as the new information priority GM(1,1) (NIPGM(1,1)) model studied recently. Stability property of the WFGM(1,1) model is studied in detail. In practice, the quantum particle swarm optimization algorithm is adopted to choose the quasi-optimal parameters for the new model so as to get the best fitting accuracy. Finally, four numerical examples from different practical applications are present. Numerical results show that the new proposed prediction model is very efficient and has both the best fitting accuracy and the best prediction accuracy compared with the GM(1,1) and the FGM(1,1) as well as the NIPGM(1,1) prediction models.

1. Introduction

Given some historic data and the current data, fitting the data trend and predicting or forecasting what will happen in the near future is of great interest and has been an ongoing hot topic in many engineering managements and applications [1]. A great deal of effective prediction models has been proposed in the past few decades. Among numerous existing prediction models, the grey theory-based models, which were first proposed by Deng in 1982 [2], have attracted many researchers’ attention due to its simple form, high computing efficiency, and high prediction accuracy especially for limited data and insufficient information [3–6].

The classical GM(1,1) model, in which the first one denotes the first order and the second one denotes the single variable, is the foundation of grey theory. Denote by \( X^{(0)} \) the original sequence. The basic idea of the GM(1,1) model is first to use the first order-accumulated generating operator (1-AGO) to transform the original sequence \( X^{(0)} \) into a monotonicity increased sequence \( X^{(1)} \). And then use the solution \( \hat{X}^{(1)} \) of a first order-ordinary differential equation, which is called the whitenization differential equation, to approximate the sequence \( X^{(1)} \). Finally, applying the first order-inverse accumulated generating operator (1-IAGO), one can get the fitted value \( \hat{X}^{(0)} \) of \( X^{(0)} \) and obtain the future trend. In the past few decades, it has been widely utilized in a great deal of engineering applications, such as the global integrated circuit industry prediction [7], the vehicle fatality risk estimation [8], the short-term freeway traffic parameter prediction [9], the fashion retailing prediction problem [10], the electricity consumption forecasting [11], and the natural gas consumption forecasting [12, 13]. To improve the prediction accuracy of the GM(1,1) model, many researchers have carried out a lot of works from different aspects, such as finding new accumulation generating operators [14–19], constructing more accurate background value formula [20, 21], choosing parameter optimization methods [22], improving initial guess [23], and reducing residuals based on Fourier analysis and Markov chain [9, 20]. Recently,
some nonhomogeneous, nonlinear, hybrid, and multivariable grey models are proposed, see [6, 24–26] for examples. Modeling mechanism analysis can be found in [27–29]. For detailed descriptions of the GM(1,1) model and its improvements, we refer the readers to [4] and a very good survey published recently [5] and also references therein.

According to the modeling process of the GM(1,1) prediction model, we can see that the 1-AGO, which generally reduces stochastic of the original data, is the first and the most important step [4, 5]. So, in this paper we mainly focus on the accumulation generating operator. To improve the precision of the GM(1,1) prediction model, many works have been carried out in the past few decades to propose new accumulation generating operators according to data sequences’ different characteristics so as to meet with the grey exponential law. For instance, for decreasing sequences, a reverse-AGO and a reciprocal-AGO were proposed, respectively, in [14, 15]. Based on random fluctuation characteristics of the data, Li and Xie designed a varied weighted-AGO with the arctangent function, which not only takes the overall growth trend but also the local fluctuation characteristics into account [19]. For fashion retailing prediction problem, in which the data sequence exhibits obvious periodicity, a cycle truncation accumulation generating operator was put forward to obtain a discrete GM(1,1) model [10]. Then, this idea was further applied to the short-term traffic flow prediction problem [30, 31]. According to the grey system theory, the effect of new information is greater than that of old information. This is called the principle of new information priority. However, the above accumulation generating operators do not actually reflect this principle. To remedy this, many researchers have proposed many new accumulation generating operators based on the idea of new information priority. These operators can be divided into two categories. On one hand, by setting different weights on different elements, a weighted-AGO (W-AGO) [16] and a new information priority-AGO (NIP-AGO) [18] were proposed to construct new GM(1,1) prediction models. In fact, the NIP-AGO can be regarded as a special case of the W-AGO and has been used in constructing multivariable prediction model [32] and discrete grey model [33]. On the other hand, by analyzing the modeling mechanism of $r$–AGO and setting the integer $r$ to be a fractional number, a fractional-AGO (F-AGO) and the corresponding fractional GM(1,1) (FGM(1,1)) model were proposed by Wu et al. [17]. Based on the idea of F-AGO, several other grey models such as fractional GM(q,1) model [34], discrete fractional grey model [35, 36], multivariable fractional grey model [37], and self-adaptive intelligence fractional grey prediction model [38] have been put forward and successfully applied in many practical applications, including the fuel production [35], the CO$_2$ emission [36], and the electricity consumption [37]. Numerical results show that the fractional grey prediction models have much higher precision than the traditional GM(1,1) model. Modeling mechanism of the fractional grey models has been studied, see [17, 29, 35, 39–41].

In this paper, based on the principle of the new information priority and the F-AGO, we first try to construct a new accumulation generating operator and call it the weighted fractional-accumulation generating operator (WF-AGO). The new WF-AGO contains two parameters, which not only can reflect the new information priority but also adjust the order of the summation according to different data sequences. Then, a new weighted fractional GM(1,1) (WFGM(1,1)) model, which involves the classical GM(1,1), the NIPGM(1,1), and the FGM(1,1) prediction models as special cases, is constructed. In addition, by defining a nonlinear constrained optimization problem for fitted values, the quantum particle swarm optimization (QPSO) method is adopted to find the best parameters [42–44]. Theoretically, stability property of the WFGM(1,1) model is analyzed in detail. With four numerical examples from different practical problems, we show that both the fitting accuracy and the prediction accuracy of the proposed model are better than those of the original GM(1,1) model [3], the FGM(1,1) model [17], and the NIPGM(1,1) model [18] proposed recently.

The remainder of this paper is organized as follows. In Section 2, a new weighted fractional-accumulation generating operator and the weighted fractional GM(1,1) model are established. Parameter estimation methods are studied. In addition, the quantum particle swarm optimization algorithm is adopted to choose the optimal parameters. Computational steps of the new model are also presented. Stability property of the new WFGM(1,1) prediction model is analyzed. Theoretical results show that the WFGM(1,1) model is also suitable for small sample data. In Section 3, four cases from different applications are studied to verify the efficacy of the new proposed WFGM(1,1) model. In addition, comparisons of the new WFGM(1,1) model with the classical GM(1,1) model, the FGM(1,1) model, and the NIPGM(1,1) model are discussed in detail. Finally, we end this paper with some conclusions and outlooks in Section 4.

2. The Weighted Fractional GM(1,1) Model

2.1. The Weighted Fractional-Accumulation Generating Operator. In this section, a new weighted fractional-accumulation generating operator (WF-AGO) and the corresponding inverse accumulation generating operator (WF-IAGO) are defined.

Definition 1 (see [34]). Let $N$ and $N^+$ denote the set of natural numbers and the set of positive integers, respectively. Let $r$ be a real number. Then, for $n \in N$, the general rising factorial function of a rational number is

$$
\left[\begin{array}{c}
r
\end{array}\right] = \begin{cases} 
1, & \text{if } n = 0, \\
\frac{r(r+1)\cdots(r+n-1)}{n!}, & \text{if } n \in N^+.
\end{cases}
$$

(1)
It has been studied in Theorem 1 [41] that \( \begin{bmatrix} r \\ n \end{bmatrix} \) is a monotonically decreasing function and a monotonically increasing function with respect to \( n \) if \( r \in (0, 1) \) and \( r \in (1, +\infty) \), respectively. In particular, if \( r = 1 \), then 

\[
\begin{bmatrix} r \\ n \end{bmatrix} \equiv 1. \]

By letting \( r \in (0, 1] \) and introducing another weighted parameter \( \lambda \in (0, 1] \) used in the NIPGM(1,1) model [18], we define a new accumulation generating operator as follows.

**Definition 2.** Let

\[
X^{(0)} = \begin{bmatrix} x^{(0)}(1), x^{(0)}(2), \ldots, x^{(0)}(n) \end{bmatrix}^T,
\]

where \( x^{(0)}(i) \) be the original nonnegative sequence, \( r, \lambda \in (0, 1] \) be two nonnegative parameters, and the weighted fractional-accumulation generating sequence of \( X^{(0)} \) is defined as

\[
X^{(r)} = \begin{bmatrix} x^{(r)}(1), x^{(r)}(2), \ldots, x^{(r)}(n) \end{bmatrix}^T,
\]

where

\[
x^{(r)}(k) = \sum_{i=1}^{k} \frac{r^{k-i} x^{(0)}(i)}{k-i}, \quad k = 1, \ldots, n.
\]

According to Definition 2, the weighted fractional-accumulation generating operator (WF-AGO) can be simply expressed by using the following matrix-vector form:

\[
X^{(r)} = A_{\lambda} X^{(0)},
\]

where \( A_{\lambda} \) is a unit lower triangular matrix and takes the form

\[
A_{\lambda} = \begin{pmatrix}
1 & 0 & 0 & \cdots & 0 \\
\lambda & 1 & 0 & \cdots & 0 \\
\lambda^2 & \lambda & 1 & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
\lambda^{n-1} & \lambda^{n-2} & \lambda^{n-3} & \cdots & 1
\end{pmatrix}
\]

Evidently, if \( r = \lambda = 1 \), then \( A_{\lambda} \) is a lower triangular matrix with all nonzero elements being one, which reduces to the well-known first order-accumulation generating operator (1-AGO) matrix. If \( r = 1 \) and \( \lambda = 1 \), then the weighted fractional-accumulation generating operator matrix \( A_{r, \lambda} \) reduces to

\[
A_{r, \lambda} = \begin{pmatrix}
1 & 0 & 0 & \cdots & 0 \\
\lambda & 1 & 0 & \cdots & 0 \\
\lambda^2 & \lambda & 1 & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
\lambda^{n-1} & \lambda^{n-2} & \lambda^{n-3} & \cdots & 1
\end{pmatrix}
\]

which are called the new information priority-accumulation generating operator (NIP-AGO) matrix [18] and the fractional-accumulation generating operator (F-AGO) matrix [17], respectively.

**Definition 3.** Let \( X^{(0)} \) and \( X^{(r)} \) be the original sequence and the weighted fractional-accumulation generating sequence defined as in (2) and (3), respectively. The weighted fractional-inverse accumulation generating operator (WF-IAGO) is

\[
X^{(0)} = A_{r, \lambda}^{-1} X^{(r)} \equiv D_{\lambda} X^{(r)}.
\]

With the notations given above, we can use \( D_r \) and \( D_{\lambda} \) to denote the fractional-inverse accumulation generating operator (F-IAGO) [17] and the new information priority-inverse accumulation generating operator (NIP-IAGO) [18] matrices, which are the inverse matrices of \( A_r \) and \( A_{r, \lambda} \), respectively. For the specific expression of \( D_{r, \lambda} \), we have the following result.

**Theorem 1.** The WF-IAGO matrix \( D_{r, \lambda} \) satisfies

\[
D_{r, \lambda} = \begin{pmatrix}
-r & 0 & 0 & \cdots & 0 \\
r \lambda & -r & 0 & \cdots & 0 \\
r^2 \lambda & r \lambda & -r & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
r^{n-1} \lambda & r^{n-2} \lambda & r^{n-3} \lambda & \cdots & -r
\end{pmatrix}
\]
Proof. According to Definition 1 and (6), we can see that the matrix $A_{x\lambda}$ is a unit lower triangular matrix. Therefore, $\det(A_{x\lambda}) = 1$ and its inverse matrix $D_{x\lambda}$ exists.

$$(A_{x\lambda}D_{x\lambda})_{ij} = \lambda^{i-j} \left[ \begin{array}{c} r \\ -r \\ \vdots \\ 0 \end{array} \right]_{i-j} = \lambda^{i-j}(A_{x\lambda}D_{x\lambda})_{ij}$$

From Theorem 2 [41], we can see that $A_{x\lambda}D_{x\lambda}$ is an identity matrix. Thus, for $i > j$ we have $(A_{x\lambda}D_{x\lambda})_{ij} = 0$, which means that the matrix $D_{x\lambda}$ defined as in (9) is the inverse matrix of $A_{x\lambda}$. Therefore, the proof is completed. \hfill \Box

2.2. The Representation of the WFGM(1,1) Model. Based on the idea of the grey theory, the new WFGM(1,1) model is constructed in this section.

Definition 4. Let $X^{(0)} = \{x^{(0)}(1), x^{(0)}(2), \ldots, x^{(0)}(n)\}^T$ and $X^{(r)} = \{x^{(r)}(1), x^{(r)}(2), \ldots, x^{(r)}(n)\}^T$ be the original sequence (2) and the WF-AGO sequence (3), respectively. We call the following linear differential equation

$$\frac{dx^{(r)}(t)}{dr} + ax^{(r)}(t) = b,$$

the whiteningization equation of the WFGM(1,1) model, and the difference equation

$$x^{(r)}(k) - x^{(r)}(k-1) + az^{(r)}(k) = b,$$

the basic model of the WFGM(1,1), where

$$z^{(r)}(k) = \frac{1}{2} (x^{(r)}(k) + x^{(r)}(k-1)), \quad k = 2, 3, \ldots, n,$$

is called the background value, and $a$ and $b$ are called the development coefficient and the grey input action, respectively.

From the definition of the WFGM(1,1) model and given the initial guess $x^{(r)}(t)_{t=1} = x^{(0)}(1)$, we can easily obtain the solution of the whiteningization equation (11):

$$\hat{x}^{(r)}(k) = (x^{(0)}(1) - \frac{b}{a}) e^{-a(k-1)} + \frac{b}{a} \quad k = 1, 2, \ldots, n,$$

(14)

which is called the time response function of the WFGM(1,1) model. Note that the time response function (14) is often used to approximate the values of the AGO sequence $X^{(r)} = \{x^{(r)}(1), \ldots, x^{(r)}(n)\}$. By using the WF-AGO defined as in (8), the predicted values $\hat{X}^{(0)} = \{\hat{x}^{(0)}(1), \ldots, \hat{x}^{(0)}(n)\}$ of the original sequence can be obtained:

$$\hat{X}^{(0)}(k) = \sum_{i=1}^{k} \left[ -r \atop k-i \right] x^{(r)}(i), \quad k = 1, 2, \ldots, n,$$

(15)

or in matrix form, it satisfies

$$\hat{X}^{(0)} = D_{x\lambda} \hat{X}^{(r)}.$$  

(16)

To prove that the matrix $D_{x\lambda}$ has form (9), we only need to show that $A_{x\lambda}D_{x\lambda}$ is an identity matrix. Since both $A_{x\lambda}$ and $D_{x\lambda}$ are unit lower triangular matrices, so is $A_{x\lambda}D_{x\lambda}$. In addition, for $i > j$, we have

$$(A_{x\lambda}D_{x\lambda})_{ij} = \lambda^{i-j} \left( \begin{array}{c} r \\ -r \\ \vdots \\ 0 \end{array} \right)_{i-j} = \lambda^{i-j}(A_{x\lambda}D_{x\lambda})_{ij}$$

Complexity
Next, we turn to study how to determine the parameters $r$ and $\lambda$, which play very important roles to make the WF-AGO sequence $X^{(\lambda)}$ better satisfy the grey exponential law and the principle of new information priority, and thus get better prediction accuracy. That is to say, the parameters must be chosen to get both the best fitting accuracy and the best prediction accuracy. There are many ways to define the model prediction accuracy, such as the mean absolute percentage error (MAPE), the mean absolute error (MAE), and the root mean square error (RMSE). Among these, the MAPE is one of the most common indexes to judge the merits of prediction models. Therefore, we use the MAPE as an objective function in this paper:

\[
\text{MAPE}(r, \lambda) = \frac{1}{n} \sum_{k=1}^{n} \left| \frac{x^{(0)}(k) - x^{(\lambda)}(k)}{x^{(0)}(k)} \right| \times 100%.
\] (20)

Theoretically, the MAPE is constrained by several conditions, such as the constraints on parameters $r$ and $\lambda$ and the time response function $\tilde{x}^{(\lambda)}(k)$ and its inverse sequence $\tilde{x}^{(0)}(k)$, as well as the parameters $a$ and $b$. In summary, the following constrained optimization problem can be defined to obtain the optimal parameters $r$ and $\lambda$:

\[
\begin{align*}
\min & \quad \text{MAPE}(r, \lambda), \\
\text{s.t.} & \quad 0 < \lambda \leq 1, \\
& \quad 0 < r \leq 1, \\
& \quad \begin{pmatrix} a \\ b \end{pmatrix} = (B^T B)^{-1} B^T Y, \\
& \quad \tilde{x}^{(\lambda)}(k) = \left( x^{(0)}(1) - \frac{b}{a} \right) e^{-a(k-1)} + \frac{b}{a}, \\
& \quad \tilde{x}^{(0)}(k) = \sum_{i=1}^{k} \left[ \begin{array}{c} -r \\ k-i \end{array} \right] \lambda^{k-i} \tilde{x}^{(\lambda)}(i), \\
& \quad k = 1, 2, \ldots, n.
\end{align*}
\] (21)

In general, the above optimization problem is non-differentiable, nonlinear, and is very difficult to solve. To find its optimal solution, we need to determine the parameters $a$ and $b$ in advance. This is a contradiction. Fortunately, we can use the quantum particle swarm optimization (QPSO) algorithm to fast get a quasi-optimal solution of (21). Note that the QPSO algorithm is a valuable development and quality improvement of the classical PSO algorithm. The greatest advantage of the QPSO algorithm over the PSO algorithm is that the QPSO ameliorates the bug of PSO preferably and improves the search efficiency [42]. Now, it has been used to solve such nonlinear constrained optimization problems [42–44]. In the following, we present a brief introduction to the QPSO algorithm to solve the nonlinear constrained optimization problem (21). Matlab code for the QPSO algorithm can be found in Appendix A.

Since there are two variables in (21), only a two-dimensional problem needs to be considered. Assume that there are $m$ discrete particles distributed in the two-dimensional space. At the $\ell$th iteration, they have the form

\[
E(\ell) = (E_1(\ell), E_2(\ell), \ldots, E_m(\ell)) = \left( \begin{array}{cccc} E_{11}(\ell) & E_{12}(\ell) & \cdots & E_{1m}(\ell) \\ E_{21}(\ell) & E_{22}(\ell) & \cdots & E_{2m}(\ell) \\ \vdots & \vdots & \ddots & \vdots \\ E_{m1}(\ell) & E_{m2}(\ell) & \cdots & E_{mm}(\ell) \end{array} \right).
\] (22)

where $E_i(\ell)$ $(i = 1, \ldots, m)$ denotes the position of the $i$th particle and $(E_{11}(\ell), E_{12}(\ell))^T$ denotes its coordinate. Here, $E_{11}(\ell)$ and $E_{12}(\ell)$ can be regarded as the choice for the parameters $r$ and $\lambda$, respectively. At the initial iteration, i.e., $\ell = 1$, let $P_i(1) = E_i(1)$ $(i = 1, \ldots, m)$. Then, at the $\ell$th iteration, the optimal position for the $i$th particle (often called “pbest”) is obtained by computing
and the global optimal position, which is called "gbest" and can be regarded as the global quasi-optimal solution for the nonlinear constrained optimization problem (21), is obtained:

$$\text{gbest} = P_d(l) = \begin{cases} E_i(l), & \text{if } \text{MAPE}(E_i(l)) < \text{MAPE}(P_i(l-1)), \\ P_i(l-1), & \text{if } \text{MAPE}(E_i(l)) \geq \text{MAPE}(P_i(l-1)), \end{cases} \quad i = 1, \ldots, m, \quad (23)$$

where

$$g = \arg \min_{1 \leq i \leq m} \text{MAPE}(P_i(l)). \quad (25)$$

At the $(l+1)$th iteration, the position $(E_{i1}$ $(l+1), E_{i2}(l+1))^T$ for each particle is updated by the following rule:

$$E_{id}(l+1) = Q_{id}(l) \pm \epsilon C_{id}(l) \cdot \ln \frac{1}{u_{id}(l)},$$

$$u_{id}(l) \sim U(0,1), \quad d = 1, 2,$$

where

$$Q_{id}(l) = \phi_{id}(l)P_{id}(l) + (1 - \phi_{id}(l))P_{gd}(l), \quad \phi_{id} \sim U(0,1),$$

$$C_{id}(l) = \frac{1}{m} \sum_{i=1}^{m} P_{id}(l),$$

$$\alpha = 0.5 - 0.5 \frac{l_{\text{max}} - l}{l_{\text{max}}},$$

and $l_{\text{max}}$ is the given largest iteration number.

According to the above discussion, the computational steps of the new WFGM(1,1) model can be summarized as follows. The corresponding Matlab code can be found in Appendix B and Appendix C.

Step 1: determine the original data sequence $X^{(0)}$ according to the actual application background. Set the initial parameters $r$ and $\lambda$, as well as the largest iteration number $l_{\text{max}}$ for the QPSO algorithm.

Step 2: compute the WF-AGO sequence $X^{(r)} = A_{r1}X^{(0)}$ and the background value $e^{(r)}(k)$ according to (3) and (13), respectively.

Step 3: solve the unconstrained optimization problem (17) or its equivalent form, i.e., the least squares problem (18), to obtain the parameters $a$ and $b$.

Step 4: substitute the parameters $a$ and $b$ into (14) to compute the approximated values $\hat{X}^{(r)}(k)$ of the WF-AGO sequence $X^{(r)}$ and compute the predicted values $\hat{X}^{(0)}$ of the WFGM(1,1) model according to (16).

Step 5: determine the local optimal parameters $r$ and $\lambda$ to minimize the MAPE (21) by using the QPSO algorithm.

Step 6: repeat Steps 2–5 to obtain the global quasi-optimal ("gbest") parameters $r^\ast$, $\lambda^\ast$, and the corresponding optimal fitted and predicted values $\hat{X}^{(0)}$.

2.4. Stability Analysis. Stability is a very important property for a prediction model since some data may be not accurate in many engineering applications due to several reasons. For the original GM(1,1) model, Wu et al. first studied its stability by using the tools of matrix perturbation analysis. They defined a perturbation bound to show how the solution is affected by the change of the original data $X^{(0)}$. In this section, we extend this idea to study the stability property for the new WFGM(1,1) model.

To this end, we first present a useful lemma and give the definition of the perturbation bound, which can be found in [27] and some recent works. Here and in the sequel, $\|\|$ denotes the 2-norm of either a vector or a matrix. $\text{rank}(\cdot)$ stands for the rank of the corresponding matrix.

Lemma 1 (see [27, 45]). Suppose that $x$ and $\bar{x}$ satisfy

$$\|Bx - y\| = \min,$$

$$\|(B + \Delta B)x - (Y + \Delta Y)\| = \min,$$

respectively, where $B, \Delta B \in C^{m \times n}$ with $m \leq n$, and $Y \neq 0$, $\Delta Y \in C^{m}$. Let $\kappa_{t} = \|B^{t}\| \cdot \|\Delta B\|$, $\gamma_{t} = 1 - \|B^{t}\| \cdot \|\Delta B\|$, and $r_{x} = Y - Bx$, $B^{t}$ be the pseudoinverse of the matrix $B$. If $\text{rank}(B) = \text{rank}(B + \Delta B)$ and $\|B^{t}\| \cdot \|\Delta B\| < 1$, then

$$\|\bar{x} - x\| \leq \frac{\kappa_{t}}{\gamma_{t}} \left( \frac{\|\Delta B\|}{\|B\|} \|x\| + \frac{\|\Delta Y\|}{\|B\|} \|y\| + \frac{\|\Delta B\|}{\|B\|} \|r_{x}\| \right).$$

Definition 5 (see [27]). Assume that $X^{(0)} = \{x^{(0)}(1), x^{(0)}(2), \ldots, x^{(0)}(n)\}$ is the original nonnegative sequence. Let $B$ and $Y$ be defined as in (19) and $B^{t} \cdot \Delta B < 1$. We call

$$L[x^{(0)}(k)] = \frac{\kappa_{t}}{\gamma_{t}} \left( \frac{\|\Delta B\|}{\|B\|} \|x^{(0)}\| + \frac{\|\Delta Y\|}{\|B\|} \|y^{(0)}\| + \frac{\|\Delta B\|}{\|B\|} \|r_{x}\| \right),$$

$$k = 1, 2, \ldots, n,$$

(30)

the perturbation bound when $\epsilon$ is a disturbance of $x^{(0)}(k)$ ($k = 1, 2, \ldots, n$).

Analogous to the analysis in [27], we have the following results for the new WFGM(1,1) model.

Theorem 2. Assume that $X^{(0)} = \{x^{(0)}(1), x^{(0)}(2), \ldots, x^{(0)}(n)\}^{T}$ is the original nonnegative sequence and $\epsilon$ is a disturbance of $x^{(0)}(k)$ ($k = 1, 2, \ldots, n$). Let $B$ and $Y$ be defined as in (19). Denote by
\[\theta_i = \left( \begin{bmatrix} r \\ i - 1 \end{bmatrix} x_i^{-1} + \begin{bmatrix} r \\ i \end{bmatrix} x_i \right)^2, \]
\[\varphi_i = \left( \begin{bmatrix} r \\ i \end{bmatrix} x_i^i - \begin{bmatrix} r \\ i - 1 \end{bmatrix} x_i^{i-1} \right)^2.\] (31)

Then, we have

\[
L[x^{(0)}(k)] = \begin{cases} 
\frac{\kappa_i}{y^*_i} \left( \frac{\sum_{i=1}^{n-1} \theta_i}{2 \| B \|} \| x \| + \frac{\sum_{i=1}^{n-1} \varphi_i}{\| B \|} + \frac{\sum_{i=1}^{n-1} \psi_i}{2 \| B \|} \| r_x \| \right), & k = 1, \\
\frac{\kappa_i}{y^*_i} \left( \frac{\sum_{i=1}^{n-k} \theta_i}{2 \| B \|} \| x \| + \frac{\sum_{i=1}^{n-k} \varphi_i}{\| B \|} + \frac{\sum_{i=1}^{n-k} \psi_i}{2 \| B \|} \| r_x \| \right), & k = 2, \ldots, n - 1, \\
\frac{\kappa_i}{y^*_i} \left( \frac{\| x \|}{2 \| B \|} + \frac{1}{2 \| B \|} \| r_x \| \right), & k = n,
\end{cases}
\]

and \(L[x^{(0)}(k)](k = 1, 2, \ldots, n - 1)\) is an increasing function with respect to \(n\).

Proof. To deduce the perturbation bound for each data, we first simplify the expression of the matrix \(B\) and the vector \(Y\). Let \(e = (1, 1, \ldots, 1)^T \in \mathbb{R}^n\). Since \(x^{(r)} = A_{x^{(r)}}x^{(0)}\) and \(e = A_{x^{(r)}}D_{x^{(r)}}e\), we have

\[
B = \begin{pmatrix}
-2^{(r)}(2) & 1 \\
-2^{(r)}(3) & 1 \\
\vdots & \vdots \\
-2^{(r)}(n) & 1
\end{pmatrix}_{(n-1) \times 2} = \begin{pmatrix}
-\frac{1}{2} & 1 & 0 & \ldots & 0 & 0 \\
0 & \frac{1}{2} & 1 & \ldots & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & \ldots & \frac{1}{2} & 1 \\
\end{pmatrix}_{(n-1) \times n}
\]

\[
Y = \begin{pmatrix}
x^{(r)}(2) - x^{(r)}(1) \\
x^{(r)}(3) - x^{(r)}(2) \\
\vdots \\
x^{(r)}(n) - x^{(r)}(n - 1)
\end{pmatrix}_{(n-1) \times 1} = \begin{pmatrix}
-1 & 1 & 0 & \ldots & 0 & 0 \\
0 & -1 & 1 & \ldots & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & \ldots & -1 & 1 \\
\end{pmatrix}_{(n-1) \times n}
\]

\[
B_{1} A_{x^{(r)}} \begin{pmatrix} X^{(0)} - D_{x^{(r)}}e \end{pmatrix},
\]

\[
B_{2} A_{x^{(r)}} X^{(0)}.
\]
(1) If $\varepsilon$ is regarded as a disturbance of $x^{(0)}(1)$, i.e.,
$X_1^{(0)}(\varepsilon) = \{x^{(0)}(1) + \varepsilon, x^{(0)}(2), \ldots, x^{(0)}(n)\}$. Then, we have

$$B + \Delta B = B_1 A_r \lambda \left( X_1^{(0)}(\varepsilon) - D_r A \varepsilon \right) = B + B_1 A_r \lambda \begin{pmatrix}
\varepsilon & 0 \\
0 & 0 \\
\vdots & \vdots \\
0 & 0
\end{pmatrix},$$

$$\Delta B = \begin{pmatrix}
-\frac{1}{2} \left( \begin{bmatrix} r \\ 0 \end{bmatrix} + \begin{bmatrix} r \\ 1 \end{bmatrix} \right) \varepsilon & 0 \\
-\frac{1}{2} \left( \begin{bmatrix} r \\ 0 \end{bmatrix} + \begin{bmatrix} r \\ 1 \end{bmatrix} \right) \varepsilon & 0 \\
\vdots & \vdots \\
-\frac{1}{2} \left( \begin{bmatrix} r \\ n-2 \end{bmatrix} \lambda^{n-2} + \begin{bmatrix} r \\ n-1 \end{bmatrix} \lambda^{n-1} \right) \varepsilon & 0
\end{pmatrix},$$

$$Y + \Delta Y = B_2 A_r X_1^{(0)}(\varepsilon) = Y + B_2 A_r \begin{pmatrix}
\varepsilon \\
0 \\
\vdots \\
0
\end{pmatrix},$$

$$\Delta Y = \begin{pmatrix}
\left( \begin{bmatrix} r \\ 1 \end{bmatrix} - \begin{bmatrix} r \\ 0 \end{bmatrix} \right) \varepsilon \\
\left( \begin{bmatrix} r \\ 2 \end{bmatrix} - \begin{bmatrix} r \\ 1 \end{bmatrix} \right) \varepsilon \\
\vdots \\
\left( \begin{bmatrix} r \\ n-1 \end{bmatrix} \lambda^{n-1} - \begin{bmatrix} r \\ n-2 \end{bmatrix} \lambda^{n-2} \right) \varepsilon
\end{pmatrix}.$$

Thus, it satisfies

$$\Delta B^T \Delta B = \begin{pmatrix}
\frac{\varepsilon^2}{4} \sum_{i=1}^{n-1} \left( \begin{bmatrix} r \\ i-1 \end{bmatrix} \lambda^{i-1} + \begin{bmatrix} r \\ i \end{bmatrix} \lambda^i \right)^2 \\
0 \\
0 \\
0
\end{pmatrix},$$

(35)

$$\Delta Y^T \Delta Y = \varepsilon^2 \sum_{i=1}^{n-1} \left( \begin{bmatrix} r \\ i \end{bmatrix} \lambda^i - \begin{bmatrix} r \\ i-1 \end{bmatrix} \lambda^{i-1} \right)^2.$$

Based on the definition of 2-norm, from (35) we have

$$\|\Delta B\| = \frac{\varepsilon}{2} \sqrt{\sum_{i=1}^{n-1} \left( \begin{bmatrix} r \\ i-1 \end{bmatrix} \lambda^{i-1} + \begin{bmatrix} r \\ i \end{bmatrix} \lambda^i \right)^2} \leq \frac{\varepsilon}{2} \sqrt{\sum_{i=1}^{n-1} \theta_i},$$

(36)

$$\|\Delta Y\| = \frac{\varepsilon}{2} \sqrt{\sum_{i=1}^{n-1} \left( \begin{bmatrix} r \\ i-1 \end{bmatrix} \lambda^{i-1} + \begin{bmatrix} r \\ i \end{bmatrix} \lambda^i \right)^2} \leq \frac{\varepsilon}{2} \sqrt{\sum_{i=1}^{n-1} \theta_i},$$

(37)

where the notations in (31) are used. Substituting (36) and (37) into Lemma 1 (30), we immediately have the disturbance bound

$$L [x^{(0)}(1)] = \frac{\kappa_x}{\gamma_x} \left[ \varepsilon \left( \frac{\sqrt{\sum_{i=1}^{n-1} \theta_i}}{2 \|B\|} + \frac{\sqrt{\sum_{i=1}^{n-1} \phi_i}}{\|B\|} + \frac{\kappa_\phi}{\gamma_\phi} \frac{\sqrt{\sum_{i=1}^{n-1} \theta_i}}{2 \|B\|} \frac{\|r_2\|}{\|B\|} \right) \right],$$

(38)

when $\varepsilon$ is regarded as a disturbance of $x^{(0)}(1)$.

(2) If $\varepsilon$ is regarded as a disturbance of $x^{(0)}(k)$ ($k = 2, 3, \ldots, n-1$), then we have
Similar to the discussion of the first case, we obtain

\[
\| \Delta B \| = \frac{|\epsilon|}{2} \sqrt{1 + \sum_{i=1}^{n-k} \left( \left[ \begin{array}{c} r \\ i-1 \end{array} \right] \lambda^{i-1} + \left[ \begin{array}{c} r \\ i \end{array} \right] \lambda^{i} \right)^2} \leq \frac{|\epsilon|}{2} \sqrt{1 + \sum_{i=1}^{n-k} \theta_i}, \quad \text{(40)}
\]

\[
\| \Delta Y \| \leq |\epsilon| \sqrt{1 + \sum_{i=1}^{n-k} \left( \left[ \begin{array}{c} r \\ i \end{array} \right] \lambda^{-i} - \left[ \begin{array}{c} r \\ i-1 \end{array} \right] \lambda^{-i-1} \right)^2} = |\epsilon| \sqrt{1 + \sum_{i=1}^{n-k} \varphi_i}, \quad \text{(41)}
\]

Substituting (40) and (41) again into Lemma 1 (30), we have the following disturbance bound:

\[
L[\chi^{(0)}(k)] = \frac{\kappa^*}{y^*} |\epsilon| \left( \frac{\sqrt{1 + \sum_{i=1}^{n-k} \theta_i}}{2\|B\|} \|x\| + \frac{\sqrt{1 + \sum_{i=1}^{n-k} \varphi_i}}{\|B\|} \frac{\sqrt{1 + \sum_{i=1}^{n-k} \theta_i}}{2\|B\|} \|r_x\| \right), \quad \text{(42)}
\]
when \( \varepsilon \) is regarded as a disturbance of \( x^{(0)}(k) \) \((k = 2, \ldots, n - 1)\).

(3) Now, we consider the last case, i.e., \( \varepsilon \) is regarded as a disturbance of \( x^{(0)}(n) \). Similar to the above two cases, we have

\[
B + \Delta B = B + B_t A_{k_t} = \begin{pmatrix}
0 & 0 \\
\vdots & \vdots \\
0 & 0
\end{pmatrix},
\]

\[
\Delta B = \frac{1}{2} \begin{pmatrix}
r \\
0
\end{pmatrix} \varepsilon \quad \Delta Y = G A_{k_t} \begin{pmatrix}
0 \\
\vdots \\
0
\end{pmatrix} \varepsilon,
\]

\[
\| \Delta B \| = |\varepsilon| \left( \frac{1}{2} \begin{pmatrix}
r \\
0
\end{pmatrix} \right)^2 = |\varepsilon|, \quad \| \Delta Y \| = |\varepsilon| \left( \frac{1}{2} \begin{pmatrix}
r \\
0
\end{pmatrix} \right)^2 = |\varepsilon|. \tag{43}
\]

Substituting (44) into Lemma 1 (30) again, we immediately obtain

\[
L = \frac{\kappa_1}{\gamma_1} |\varepsilon| \left( \frac{1}{2} \| B \| \| x \| + \frac{1}{\gamma_1} \frac{1}{\gamma} \frac{1}{\| B \|} \| r_x \| \right), \tag{45}
\]

when \( \varepsilon \) is regarded as a disturbance of \( x^{(0)}(n) \).

As can be seen from (32), the disturbance bounds \( L \) \((k = 1, 2, \ldots, n - 1)\) are increasing functions with respect to \( n \) since \( \theta_i > 0 \) and \( \varphi_j > 0 \). That is to say the larger the sample size of the original sequence is, the larger the perturbation bound of the solution will be. Therefore, similar to the traditional GM(1,1), the FGM(1,1) model, and the NIPGM(1,1) model, the new WFGM(1,1) model is also suitable for prediction problem with small samples.

### 3. Numerical Experiments

In this section, four numerical examples, which come from the annual per capita electricity consumption prediction problem [26], the energy consumption prediction problem [18], the natural gas consumption prediction problem [13, 33], and the total output value of construction industry prediction problem, respectively, are adopted to show the high precision of the new WFGM(1,1) model. To better show its advantages, numerical results of the WFGM(1,1) model are compared with those of the original GM(1,1) model, the fractional GM(1,1) model (FGM(1,1)) [17], and the new information priority GM(1,1) model (NIPGM(1,1)) [18] studied recently. MAPE (20) is used as the evaluation index. In addition, we also compare the relative percentage error of each data for the four discussed grey models:

\[
\text{MAPE} = \frac{\sum_{i=1}^{n} \left| \frac{x_i^{(0)}(k) - \hat{x}_i^{(0)}(k)}{x_i^{(0)}(k)} \right|}{n} \times 100\%.
\]

Note that all the FGM(1,1) model, the NIPGM(1,1) model, and the new WFGM(1,1) model can be regarded as the accumulation generating operator improvements of the traditional GM(1,1) model. So, it is fair to compare their performance. Besides, the QPSO algorithm, which is derived for the WFGM(1,1) model in Section 2.3, is also suitable for the FGM(1,1) model and the NIPGM(1,1) model. In actual implementations of the QPSO algorithm, we take \( m = 5 \) and \( l_{\max} = 500 \), i.e., 5 discrete particles are distributed in the problem domain and the largest iteration number is 500. The detailed Matlab codes for these test problems are listed in appendices.

**Example 1.** Consider the annual per capita electricity consumption prediction problem in China [26]. The raw data (unit: kilowatt hour) are collected from the official website (http://data.stats.gov.cn/english/easyquery.htm?cn=C01) of the National Bureau of Statistics of China.

For the first example, we list the raw data in Table 1. Note that the official website only offers the data from 2000 to 2015 \((k = 1, \ldots, 16)\), which have been studied in [26]. Here, the data from 2000 to 2012 \((k = 1, \ldots, 13)\) are used to build the grey models and the data from 2013 to 2015 \((k = 14, 15, 16)\) are used to verify the prediction accuracy. By using the QPSO algorithm studied in Section 2.3, the quasi-optimal parameters for the FGM(1,1), the NIPGM(1,1), and the new WFGM(1,1) prediction models can be obtained. We list them in Table 2. The fitted data and the predicted data for each model are listed in Table 1. The relative error for each data and the error index MAPE are also listed in Table 1. To better show their comparison, we plot the raw data, the fitted data, and the predicted data for each prediction model in Figure 1. Their relative errors are also plotted.

From Table 1 and Figure 1, we can see that the accumulation generating operator has great effects on both the fitting accuracy and the prediction accuracy of the grey model. The fitting accuracy of the FGM(1,1) model is slightly better than that of the NIPGM(1,1) model, while the NIPGM(1,1) model has slightly better prediction accuracy than the FGM(1,1) model. Although the fitting error of the
Table 1: Fitted and predicted values of annual per capita electricity consumption of China (Example 1).

<table>
<thead>
<tr>
<th>Year</th>
<th>( k )</th>
<th>Actual values</th>
<th>GM(1,1)</th>
<th>FGM(1,1)</th>
<th>NIPGM(1,1)</th>
<th>WFGM(1,1)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>( x(0)(k) )</td>
<td>( \hat{x}(0)(k) )</td>
<td>( \Delta k(%) )</td>
<td>( \hat{x}(0)(k) )</td>
<td>( \Delta k(%) )</td>
</tr>
<tr>
<td>2000</td>
<td>1</td>
<td>1066.9</td>
<td>1066.9</td>
<td>0.00</td>
<td>1066.9</td>
<td>0.00</td>
</tr>
<tr>
<td>2001</td>
<td>2</td>
<td>1157.6</td>
<td>1266.0</td>
<td>9.36</td>
<td>1143.0</td>
<td>1.26</td>
</tr>
<tr>
<td>2002</td>
<td>3</td>
<td>1286.0</td>
<td>1399.9</td>
<td>8.85</td>
<td>1314.2</td>
<td>2.19</td>
</tr>
<tr>
<td>2003</td>
<td>4</td>
<td>1477.1</td>
<td>1548.0</td>
<td>4.80</td>
<td>1508.2</td>
<td>2.10</td>
</tr>
<tr>
<td>2004</td>
<td>5</td>
<td>1695.2</td>
<td>1711.7</td>
<td>0.97</td>
<td>1714.1</td>
<td>1.12</td>
</tr>
<tr>
<td>2005</td>
<td>6</td>
<td>1913.0</td>
<td>1892.8</td>
<td>1.06</td>
<td>1929.4</td>
<td>0.86</td>
</tr>
<tr>
<td>2006</td>
<td>7</td>
<td>2180.6</td>
<td>2093.0</td>
<td>4.02</td>
<td>2153.4</td>
<td>1.25</td>
</tr>
<tr>
<td>2007</td>
<td>8</td>
<td>2482.2</td>
<td>2314.4</td>
<td>6.76</td>
<td>2386.1</td>
<td>3.87</td>
</tr>
<tr>
<td>2008</td>
<td>9</td>
<td>2607.6</td>
<td>2559.2</td>
<td>1.86</td>
<td>2627.9</td>
<td>0.78</td>
</tr>
<tr>
<td>2009</td>
<td>10</td>
<td>2781.7</td>
<td>2829.9</td>
<td>1.73</td>
<td>2879.3</td>
<td>3.51</td>
</tr>
<tr>
<td>2010</td>
<td>11</td>
<td>3134.8</td>
<td>3129.3</td>
<td>0.18</td>
<td>3140.8</td>
<td>0.19</td>
</tr>
<tr>
<td>2011</td>
<td>12</td>
<td>3497.0</td>
<td>3460.3</td>
<td>1.05</td>
<td>3413.0</td>
<td>2.40</td>
</tr>
<tr>
<td>2012</td>
<td>13</td>
<td>3684.2</td>
<td>3826.3</td>
<td>3.86</td>
<td>3696.5</td>
<td>0.33</td>
</tr>
<tr>
<td>MAPE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2013</td>
<td>14</td>
<td>3993.0</td>
<td>4231.1</td>
<td>5.96</td>
<td>3992.1</td>
<td>0.20</td>
</tr>
<tr>
<td>2014</td>
<td>15</td>
<td>4132.9</td>
<td>4678.7</td>
<td>13.21</td>
<td>4300.3</td>
<td>4.05</td>
</tr>
<tr>
<td>2015</td>
<td>16</td>
<td>4321.0</td>
<td>5173.6</td>
<td>22.28</td>
<td>4621.8</td>
<td>9.24</td>
</tr>
<tr>
<td>MAPE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Parameters \( r \) and \( \lambda \) used in grey prediction models for Example 1.

<table>
<thead>
<tr>
<th></th>
<th>GM(1,1)</th>
<th>FGM(1,1)</th>
<th>NIPGM(1,1)</th>
<th>WFGM(1,1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( r )</td>
<td>1</td>
<td>0.2401</td>
<td>1</td>
<td>0.3337</td>
</tr>
<tr>
<td>( \lambda )</td>
<td>1</td>
<td>1</td>
<td>0.2074</td>
<td>0.7307</td>
</tr>
</tbody>
</table>

Figure 1: Continued.
classical GM(1,1) model is nice, its prediction error is over 10% and seems unacceptable. Both the fitting accuracy and the prediction accuracy show that the new WFGM(1,1) model is the best one. Most important, the fitting error of the WFGM(1,1) model for each data is less than 5% and the corresponding MAPE value is less than 4%, which show very high precision of the new proposed WFGM(1,1) model.

Example 2. Consider the energy consumption prediction problem for Jiangsu Province, China. The raw data (unit: ten thousand tons standard coal) come from [18], Table 2) and have been studied to show high precision of the NIPGM(1,1) model. For the second example, the raw data, which show the energy consumption of Jiangsu Province, China from 2001 to 2012, are listed in Table 3. Here, in order to verify the accuracy of the WFGM(1,1), the data of energy consumption from 2001 to 2008 (k = 1, ..., 8) are used to fit and the data from 2009 to 2012 (k = 9, ..., 12) are used to predict. The quasi-optimal parameters of the FGM(1,1), the NIPGM(1,1) and the new WFGM(1,1) prediction models for the second example are listed in Table 4. The fitted data and the predicted data of the four discussed prediction models are listed in Table 3 and are plotted in Figure 2. In Figure 2, we also plot the relative error of each data for the four discussed prediction models. Both the fitting accuracy and the prediction accuracy show that all the FGM(1,1) model, the NIPGM(1,1) model, and the WFGM(1,1) model greatly improve the classical GM(1,1) model. From Table 3 and Figure 2, we can see again that the new WFGM(1,1) model has both the best fitting accuracy and the best prediction accuracy, which confirms our original intention of this work. More specifically, from the aspect of the fitting accuracy, the NIPGM(1,1) model and the proposed WFGM(1,1) model have almost the same result and both are better than the FGM(1,1) model. From the aspect of the prediction accuracy, the proposed WFGM(1,1) model is better than the NIPGM(1,1) model and the later one is better than the FGM(1,1) model.

Example 3. Consider the China’s natural gas consumption prediction problem (unit: billion cubic meters). This problem has been deeply studied in [13] and was used to be a test problem in [33] to show the efficiency of the new information priority discrete grey model. The raw data of the third example are listed in Table 5, which can be found in Table 1 [33]. These data show China’s natural gas consumption from 2003 to 2013 (k = 1, ..., 11). Here, we use the data from 2003 to 2009 (k = 1, ..., 7) as the sample and the data from 2010 to 2013 (k = 8, ..., 11) to test the performance of different prediction models. In Table 6, we list the quasi-optimal parameters of the FGM(1,1), the NIPGM(1,1), and the new WFGM(1,1) prediction models computed by the QPSO algorithm for the third example. In Table 5, we list the numerical results for the GM(1,1), the FGM(1,1), the NIPGM(1,1), and the new WFGM(1,1) prediction models. To better show their difference, we plot the fitted data and the predicted data for each prediction model in Figure 3. The relative errors for the fitted data and the predicted data are also plotted. From these numerical results, we can see again that the new WFGM(1,1) model has both the best fitting accuracy and the best prediction accuracy compared with the classical GM(1,1) model, the FGM(1,1) model, and the recent developed NIPGM(1,1) model. Different from the second example, the FGM(1,1) model has better prediction accuracy than the NIPGM(1,1) model. This shows that for some problems the FGM(1,1) model has advantages over the NIPGM(1,1) model. Our new model can take full use of advantages of both models. In addition, the fitting error and prediction error for each data are less than 5%, which shows that the new proposed WFGM(1,1) model is a very powerful prediction model.

Example 4. Consider the China’s total output value of construction industry prediction problem. The raw data (unit: ¥100 million) are collected from 2008 to 2018 and can be found in
Table 3: Fitted and predicted values of energy consumption of Jiangsu Province, China (Example 2).

<table>
<thead>
<tr>
<th>Year</th>
<th>( k )</th>
<th>Actual values</th>
<th>( \hat{x}^{(0)}(k) )</th>
<th>( \hat{\Delta} ) (%)</th>
<th>Fitted values</th>
<th>Predicted values</th>
<th>( \hat{\Delta} ) (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2001</td>
<td>1</td>
<td>8881</td>
<td>8881</td>
<td>0.00</td>
<td>8881</td>
<td>8881</td>
<td>0.00</td>
</tr>
<tr>
<td>2002</td>
<td>2</td>
<td>9609</td>
<td>10483</td>
<td>9.09</td>
<td>9609</td>
<td>9621</td>
<td>1.26</td>
</tr>
<tr>
<td>2003</td>
<td>3</td>
<td>11061</td>
<td>11988</td>
<td>8.38</td>
<td>11990</td>
<td>11826</td>
<td>6.92</td>
</tr>
<tr>
<td>2004</td>
<td>4</td>
<td>13652</td>
<td>13710</td>
<td>0.42</td>
<td>14048</td>
<td>14031</td>
<td>2.78</td>
</tr>
<tr>
<td>2005</td>
<td>5</td>
<td>17167</td>
<td>15678</td>
<td>8.67</td>
<td>16129</td>
<td>16224</td>
<td>5.50</td>
</tr>
<tr>
<td>2006</td>
<td>6</td>
<td>18742</td>
<td>17930</td>
<td>4.33</td>
<td>18239</td>
<td>18404</td>
<td>1.80</td>
</tr>
<tr>
<td>2007</td>
<td>7</td>
<td>20948</td>
<td>20505</td>
<td>2.12</td>
<td>20391</td>
<td>20572</td>
<td>1.80</td>
</tr>
<tr>
<td>2008</td>
<td>8</td>
<td>22232</td>
<td>23449</td>
<td>5.47</td>
<td>22599</td>
<td>22728</td>
<td>2.23</td>
</tr>
<tr>
<td>MAPE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2.63</td>
<td>2.61</td>
</tr>
<tr>
<td>2009</td>
<td>9</td>
<td>23709</td>
<td>26816</td>
<td>13.12</td>
<td>24875</td>
<td>24872</td>
<td>4.90</td>
</tr>
<tr>
<td>2010</td>
<td>10</td>
<td>25774</td>
<td>30667</td>
<td>18.99</td>
<td>27228</td>
<td>27004</td>
<td>4.77</td>
</tr>
<tr>
<td>2011</td>
<td>11</td>
<td>27589</td>
<td>35071</td>
<td>27.12</td>
<td>29669</td>
<td>29124</td>
<td>5.56</td>
</tr>
<tr>
<td>2012</td>
<td>12</td>
<td>28850</td>
<td>40108</td>
<td>39.02</td>
<td>32207</td>
<td>31232</td>
<td>8.26</td>
</tr>
<tr>
<td>MAPE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2.63</td>
<td>2.61</td>
</tr>
</tbody>
</table>

Table 4: Parameters \( r \) and \( \lambda \) used in grey prediction models for Example 2.

<table>
<thead>
<tr>
<th></th>
<th>GM(1,1)</th>
<th>FGM(1,1)</th>
<th>NIPGM(1,1)</th>
<th>WFGM(1,1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( r )</td>
<td>1</td>
<td>0.4481</td>
<td>1</td>
<td>0.7499</td>
</tr>
<tr>
<td>( \lambda )</td>
<td>1</td>
<td>1</td>
<td>0.6562</td>
<td>0.5325</td>
</tr>
</tbody>
</table>
the official website (http://data.stats.gov.cn/easyquery.htm?cn=C01) of the National Bureau of Statistics of China.

For the fourth example, the raw data are listed in Table 7. The data from 2009 to 2014 \((k=1, \ldots, 6)\) are used as the sample and the data from 2015 to 2019 \((k=7, \ldots, 10)\) are used to test the performance of four different prediction models. The quasi-optimal parameters of the FGM(1,1), the NIPGM(1,1) model, and the WFGM(1,1) models are computed by the QPSO algorithm and listed in Table 8. In Table 7, the fitted data, the predicted data, and the corresponding values of error indexes for each model are listed. Figure 4 clearly shows these numerical results. From Table 7 and Figure 4, we can see that all prediction models can effectively simulate the sample data. Although the GM(1,1) model has the worst fitness MAPE, its value is only 1.97%. The fitness MAPE values of other three models are almost the same and only a quarter of the GM(1,1) model. However, the predicted data again show that the original GM(1,1) model is unacceptable. All other improved GM(1,1) models can effectively simulate the near further trend and greatly reduce the prediction errors. Among them, the WFGM(1,1) model performs the best. This further confirms that the new proposed WFGM(1,1) model is a valuable improvement of the grey models.

### 4. Concluding Remarks and Outlooks

In this paper, we have proposed a novel weighted fractional GM(1,1) model, which is a valuable development and a quality improvement of the classical GM(1,1) model from

---

**Table 5: Fitted and predicted values of natural gas consumption of China (Example 3).**

<table>
<thead>
<tr>
<th>Year</th>
<th>(k)</th>
<th>Actual values</th>
<th>GM(1,1) (x^{(0)}(k))</th>
<th>FGM(1,1) (\hat{x}^{(0)}(k))</th>
<th>NIPGM(1,1) (\hat{x}^{(0)}(k))</th>
<th>WFGM(1,1) (\hat{x}^{(0)}(k))</th>
<th>(\hat{A}_k(%))</th>
<th>(\hat{A}_k(%))</th>
<th>(\hat{A}_k(%))</th>
<th>(\hat{A}_k(%))</th>
</tr>
</thead>
<tbody>
<tr>
<td>2003</td>
<td>1</td>
<td>35</td>
<td>35.0</td>
<td>0.00</td>
<td>35.0</td>
<td>0.00</td>
<td>35.0</td>
<td>0.00</td>
<td>35.0</td>
<td>0.00</td>
</tr>
<tr>
<td>2004</td>
<td>2</td>
<td>41.5</td>
<td>43.7</td>
<td>5.22</td>
<td>41.6</td>
<td>0.31</td>
<td>41.5</td>
<td>0.00</td>
<td>41.5</td>
<td>0.00</td>
</tr>
<tr>
<td>2005</td>
<td>3</td>
<td>49.3</td>
<td>50.3</td>
<td>2.03</td>
<td>50.6</td>
<td>2.70</td>
<td>50.3</td>
<td>2.05</td>
<td>50.6</td>
<td>2.70</td>
</tr>
<tr>
<td>2006</td>
<td>4</td>
<td>58.6</td>
<td>57.9</td>
<td>1.12</td>
<td>59.6</td>
<td>1.71</td>
<td>59.3</td>
<td>1.14</td>
<td>59.8</td>
<td>2.00</td>
</tr>
<tr>
<td>2007</td>
<td>5</td>
<td>69.2</td>
<td>66.7</td>
<td>3.55</td>
<td>68.5</td>
<td>1.06</td>
<td>68.4</td>
<td>1.19</td>
<td>68.7</td>
<td>0.74</td>
</tr>
<tr>
<td>2008</td>
<td>6</td>
<td>80.3</td>
<td>76.9</td>
<td>4.25</td>
<td>77.3</td>
<td>3.74</td>
<td>77.6</td>
<td>3.12</td>
<td>77.3</td>
<td>3.69</td>
</tr>
<tr>
<td>2009</td>
<td>7</td>
<td>85.2</td>
<td>88.6</td>
<td>3.95</td>
<td>86.2</td>
<td>1.13</td>
<td>87.1</td>
<td>2.17</td>
<td>85.8</td>
<td>0.60</td>
</tr>
<tr>
<td>MAPE</td>
<td></td>
<td></td>
<td>2.87</td>
<td>1.52</td>
<td>1.41</td>
<td>1.39</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010</td>
<td>8</td>
<td>94.8</td>
<td>102.2</td>
<td>7.61</td>
<td>95.1</td>
<td>0.35</td>
<td>96.6</td>
<td>1.92</td>
<td>93.8</td>
<td>1.03</td>
</tr>
<tr>
<td>2011</td>
<td>9</td>
<td>103.1</td>
<td>117.5</td>
<td>13.98</td>
<td>104.2</td>
<td>1.10</td>
<td>106.3</td>
<td>3.15</td>
<td>101.7</td>
<td>1.38</td>
</tr>
<tr>
<td>2012</td>
<td>10</td>
<td>107.2</td>
<td>135.4</td>
<td>26.28</td>
<td>113.5</td>
<td>5.90</td>
<td>116.2</td>
<td>8.43</td>
<td>109.2</td>
<td>1.94</td>
</tr>
<tr>
<td>2013</td>
<td>11</td>
<td>119.3</td>
<td>155.9</td>
<td>30.71</td>
<td>123.0</td>
<td>3.12</td>
<td>126.3</td>
<td>5.86</td>
<td>116.6</td>
<td>2.23</td>
</tr>
<tr>
<td>MAPE</td>
<td></td>
<td></td>
<td>19.65</td>
<td>2.62</td>
<td>4.84</td>
<td>1.65</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table 6: Parameters \(r\) and \(\lambda\) used in grey prediction models for Example 3.**

<table>
<thead>
<tr>
<th></th>
<th>GM(1,1)</th>
<th>FGM(1,1)</th>
<th>NIPGM(1,1)</th>
<th>WFGM(1,1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(r)</td>
<td>1</td>
<td>0.4318</td>
<td>1</td>
<td>0.3842</td>
</tr>
<tr>
<td>(\lambda)</td>
<td>1</td>
<td>0.7464</td>
<td>0.4446</td>
<td></td>
</tr>
</tbody>
</table>
Figure 3: Numerical results for Example 3: fitted and predicted values (a) and relative errors (b).

Table 7: Fitted and predicted values of China’s total output value of construction industry (Example 4).

<table>
<thead>
<tr>
<th>Year</th>
<th>$k$</th>
<th>Actual values</th>
<th>GM(1,1)</th>
<th>NIPGM(1,1)</th>
<th>WFGM(1,1)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$x^{(0)}(k)$</td>
<td>$\hat{x}^{(0)}(k)$</td>
<td>$\hat{\Delta}_x(%)$</td>
<td>$\hat{x}^{(0)}(k)$</td>
</tr>
<tr>
<td>2009</td>
<td>1</td>
<td>76807.70</td>
<td>76807.70</td>
<td>0.00</td>
<td>76807.70</td>
</tr>
<tr>
<td>2010</td>
<td>2</td>
<td>96031.10</td>
<td>99920.93</td>
<td>4.05</td>
<td>96031.28</td>
</tr>
<tr>
<td>2011</td>
<td>3</td>
<td>116463.30</td>
<td>115815.32</td>
<td>0.56</td>
<td>117392.32</td>
</tr>
<tr>
<td>2012</td>
<td>4</td>
<td>137217.86</td>
<td>134238.02</td>
<td>2.17</td>
<td>138102.74</td>
</tr>
<tr>
<td>2013</td>
<td>5</td>
<td>160366.06</td>
<td>155591.22</td>
<td>2.98</td>
<td>157967.27</td>
</tr>
<tr>
<td>2014</td>
<td>6</td>
<td>176713.40</td>
<td>180341.07</td>
<td>2.05</td>
<td>177039.71</td>
</tr>
<tr>
<td>MAPE</td>
<td></td>
<td>1.97</td>
<td>0.52</td>
<td>0.18</td>
<td>0.51</td>
</tr>
<tr>
<td>2015</td>
<td>7</td>
<td>180757.47</td>
<td>209027.86</td>
<td>15.64</td>
<td>195402.49</td>
</tr>
<tr>
<td>2016</td>
<td>8</td>
<td>193566.78</td>
<td>242277.86</td>
<td>25.16</td>
<td>213131.48</td>
</tr>
<tr>
<td>2017</td>
<td>9</td>
<td>213943.56</td>
<td>280816.92</td>
<td>31.26</td>
<td>230290.99</td>
</tr>
<tr>
<td>2018</td>
<td>10</td>
<td>235085.53</td>
<td>325486.37</td>
<td>38.45</td>
<td>246934.64</td>
</tr>
<tr>
<td>MAPE</td>
<td></td>
<td>27.63</td>
<td>7.72</td>
<td>9.71</td>
<td>7.00</td>
</tr>
</tbody>
</table>
the aspect of the accumulation generating process. The new
model involves some existing improved GM(1,1) models as
special cases. We have derived the modeling process of the new
model and studied its stability property. In addition, the QPSO
algorithm is used to obtain two optimal parameters involved in
the new accumulation generating operator. With four nu-
merical examples from different applications, we have shown
that the new WFGM(1,1) prediction model has much better
fitting accuracy and prediction accuracy than the classical
GM(1,1) model and two improved GM(1,1) models studied
recently, i.e., the FGM(1,1) model [17] and the NIPGM(1,1)
model [18]. Future works should focus on testing the
WFGM(1,1) model on other applications where data exhibit
strong nonlinear or periodicity characteristics, constructing
new weighted fractional self-adaptive nonlinear and non-
homogeneous grey models, error estimates, and so on.

Appendix

A. Matlab code for the QPSO algorithm

1) function [gbestfit,lambda, r]=

Table 8: Parameters $r$ and $\lambda$ used in grey prediction models for Example 4.

<table>
<thead>
<tr>
<th></th>
<th>GM(1,1)</th>
<th>FGM(1,1)</th>
<th>NIPGM(1,1)</th>
<th>WFGM(1,1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r$</td>
<td>1</td>
<td>0.2282</td>
<td>1</td>
<td>0.2061</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>1</td>
<td>1</td>
<td>0.6888</td>
<td>0.4505</td>
</tr>
</tbody>
</table>

Figure 4: Numerical results for Example 4: fitted and predicted values (a) and relative errors (b).
% Initializing parameters
(33) ibestpos = popul; % Initializing individual best position
(34) ibestfit = fitness; % Fitness of each individual
(35) [fbestpart, g] = min(ibestfit); % The best global fit
(36) gbestfit = fbestpart;
(37) gbestpos = ibestpos(:, g); % The optimal parameters correspond to the best global fitness

(38) for count = 1:maxgen
(39) for m = 1:popsize
(40) m_NFGM = WFGM(x0(1:t), n, t, popul(1, m), popul(2, m));
(41) fitness(m) = sum((abs(m_NFGM(1:t)-x0(1:t))./x0(1:t)) * 100)/t;
(42) if fitness(m) < ibestfit(m);
(43) ibestfit(m) = fitness(m);
(44) ibestpos(:, m) = popul(:, m);
(45) end
(46) end
(47) end
(48) m_best = sum(ibestpos)/popsize;
(49) [fbestpart, g] = min(fitness); % Update global history best position
(50) if fbestpart < gbestfit
(51) gbestfit = fbestpart;
(52) gbestpos = popul(:, g);
(53) end
(54) bestfitness(count) = gbestfit;
(55) end
(56) % update the position
(57) W(count) = 0.5–0.5 * (maxgen-count)/maxgen;
(58) for j = 1:popsize
(59) for i = 1:dimensions
(60) f = rand;
(61) pp = f * ibestpos(i, j) + (1-f) * gbestpos(i)';
(62) u = rand;
(63) if u > 0.5
(64) popul(i, j) = pp - W(count) * abs(m_best(1, j) - popul(i, j)) * log(1/u);
(65) else
(66) popul(i, j) = pp + W(count) * abs(m_best(1, j) - popul(i, j)) * log(1/u);
(67) end
(68) if popul(i, j) > populmax(i)
(69) popul(i, j) = rand * (populmax(i) - populmin(i)) + populmin(i); %rand * 1.5; popul(max
(70) elseif popul(i, j) < populmin(i)
(71) popul(i, j) = rand * (populmax(i) - populmin(i)) + populmin(i); %rand * (-1.5); popul(min
(72) end
(73) end
(74) end
(75) end
(76) end
B. Matlab code for the WFGM(1,1) prediction model

(1) function \texttt{x0\_WFGM} = WFGM(x0, n, t, lambda, r)
(2) % Input: x0-data sequence
(3) % n-total length of the input data
(4) % t-data length used for fitting
(5) % lambda-weight
(6) % r-fractional order
(7) % Output: x0\_WFGM - predicted data
(8) % Construct weighted fractional accumulation matrix A
(9) d(1) = 1;
(10) for i = 2 : n
(11) \quad d(i) = d(i-1) \times \lambda \times \frac{r+i-2}{i-1};
(12) end
(13) for i = 1 : n
(14) \quad A(i:n, i) = d(1:n-i+1);
(15) end
(16) % Data preprocess
(17) x1 = A(1:t,1:t) * x0';
(18) % Compute background values
(19) z1 = 0.5 * (x1(2:t)+x1(1:t-1));
(20) % Compute the development coefficient and the grey input action
(21) Z = \exp(z1) \times \frac{1}{a} + \frac{b}{a};
(22) % Compute the constant for whitening equation based on initial condition
(23) c = \exp(a) \times (x0(1)-b/a);
(24) % Compute the predicted value
(25) \hat{x}_1 = c \times \exp(-a \times (1:n)+b/a);
(26) x0\_WFGM = A \hat{x}_1';
(27) x0\_WFGM = x0\_WFGM';
(28) return

C. Matlab code for test problems

(1) clc;
(2) clear;
(3) format short
(4) % Example = input('Choose Example: ');
(5) if Example == 1
(6) \quad x0 = [1066.9 1157.6 1286 1477.1 1695.2 1913 2180.6 2482.2
(7) \quad 2607.6 2781.7 3134.8 3497.0 3684.2 3993 4132.9 4231];
(8) \quad t = 13;
(9) \quad % Annual per capita electricity consumption prediction problem. Data comes from Table 3 of
(10) % "B.-L. Wei, N.-M. Xie and A. Hu, Optimal solution for novel grey polynomial prediction model,
(12) elseif Example == 2
(13) \quad x0 = [8881 9609 11061 13652 17167 18742 20948 22232 23709 25774 27589 28850];
(14) \quad t = 8;
(15) \quad % Energy consumption prediction problem. Data comes from Table 1 of "W.-J. Zhou, H.-R.
(16) % Dang and Z.-X. Wang, New information priority accumulated grey discrete model and
(18) elseif Example == 3
(19) \quad x0 = [35 41.5 49.3 58.6 69.2 80.3 85.2 94.8 103.1 107.2 119.3];
(20) \quad t = 7;
(21) \quad % Natural gas output prediction problem. Data comes from Table 4 of "B. Zeng,
(22) % Forecasting the relation of supply and demand of natural gas in China during 2015–2020
(24) elseif Example == 4
(25) \quad x0 = [76807.70 96031.10 116463.30 137217.86 160366.06 176713.40
(26) \quad 180757.47 193566.78 213943.56 235085.53];
(27) \quad t = 6;
Complexity 19

(29) % Total output value of construction industry prediction problem.
(30) % Data comes from the official website http://data.stats.gov.cn/easyquery.htm?cn=C01
(31) End
(32) n = length(x0); % Compute the length of the data sequence
(33) fprintf('Fitting and predicting by the WFGM(1,1) model[n]
(34) populmax1 = 1; populmin1 = 0;
(35) fprintf('Quasi-optimal parameters for WFGM(1,1) are: lambda %1.4f, r %1.4f/n',
(36) xWFGM = WFGM(x0(1 : t), n, t, lambda_WFGM,r_WFGM);
(37) epsilon_Fitness_xWFGM = (abs(xWFGM(1 : t) - x0(1 : t)))./x0(1 : t)) * 100;
(38) MAPE_Fitness_xWFGM = sum(epsilon_Fitness_xWFGM)/t;
(39) epsilon_Predict_xWFGM = (abs(xWFGM(t + 1 : n)-x0(t + 1 : n)))./x0(t + 1 : n)) * 100;
(40) fprintf('Fitting MAPE for WFGM(1,1) is %1.4f/n', MPAE_Fitness_xWFGM)
(41) fprintf('Predict MAPE for WFGM(1,1) is %1.4f/n', MPAE_Predict_xWFGM)

Data Availability
The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest
The authors declare that they have no conflicts of interest.

Acknowledgments
This work was supported by the National Natural Science Foundation of China (nos. 61771265 and 11771225), the Natural Science Foundation of the Jiangsu Higher Education Institutions of China (nos. 18KJB580012 and 19KJB580003), and the Science and Technology Project of Nantong City (nos. CP12017001, GY12017006, and JC2018142).

Supplementary Materials
Matlab code for four test problems are illustrated in Section 3; WFGM: Matlab code for the weighted fractional GM(1,1) prediction model; WFGM_QPSO: Matlab code for the QPSO algorithm used in the WFGM(1,1) model. (Supplementary Materials)

References


Submit your manuscripts at
www.hindawi.com