

Research Article

Grey Forecast Model with Aging Fractional Accumulation and Its Properties

Leping Tu , Yan Chen , and Lifeng Wu 

School of Management Engineering and Business, Hebei University of Engineering, Handan 056038, China

Correspondence should be addressed to Lifeng Wu; wulifeng@hebeu.edu.cn

Received 9 April 2021; Revised 2 July 2021; Accepted 7 July 2021; Published 20 July 2021

Academic Editor: Ali Jaballah

Copyright © 2021 Leping Tu 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.

A novel aging fractional accumulation operator is proposed. The aging accumulation operator can dynamically update the accumulation weight of data and flexibly change the forecast trend by adjusting the aging parameter. In addition, a new aging accumulated grey model is obtained by using the aging accumulation operator to improve the traditional grey model. In the analysis of four examples, the existing grey accumulation operator and prediction method are compared. The results show that the proposed aging accumulation operator and aging accumulation grey model have excellent performance.

1. Introduction

The grey model is a very effective forecasting method to deal with the problem with poor information and little data [1]. Other existing prediction methods such as neural network algorithm [2, 3], exponential smoothing [4], support vector regression [5], and autoregression [6] often depend on the amount of data. The grey prediction model only needs at least 4 data to make a prediction. This advantage makes the grey prediction model achieve good results even when the amount of data is small or data collection is difficult [7]. However, the traditional grey model still has some shortcomings. The improvements in recent years mainly focus on the following four aspects:

- (1) Optimization of model background value: traditional background values $z^{(1)}(k) = 0.5(x^{(1)}(k) + x^{(1)}(k+1))$ are suitable for smooth sequences, which can be optimized to adapt to other situations. The model background value is reconstructed by the Simpson formula, and the unbiased GM(SD) (1, 1) model is obtained [8]. By increasing the number of parameters in the background value, the smoothness of the background value is improved and the influence of the extremum in the original sequence is

weakened [9]. The NNGM (1, 1) model is constructed by a neural network algorithm, so there is no need to determine the background value [10].

- (2) The extension of the modeling equation: the DGM model is proposed by using the discrete modeling method, which avoids the jumping error of GM (1, 1) from discrete equation to continuous equation [11]. An unbiased nonlinear grey Bernoulli model is constructed to achieve better performance by adjusting nonlinear parameters [12]. The GMCO (1, N) model with optimized parameters is proposed, which can accurately describe any linear dynamic grey system [13].
- (3) Improvement of grey buffer operator: the original sequence is usually irregular, but its potential law can be revealed by appropriate grey buffer operators. The fractional buffer operator obtained by extending the integer buffer operator can adjust the buffer effect more accurately [14]. Three new fractional weakening buffer operators are proposed, which can effectively weaken the interference of disturbance factors on time series [15]. An optimized grey buffer operator is proposed by introducing accumulation and translation transformation [16].

- (4) Error correction: the prediction accuracy can be further improved by error analysis of prediction results combined with correction technology. The Fourier error correction method is used to improve the existing grey forecasting model [17]. The triangle residual error correction method is used to eliminate the inherent error of the original grey model, and a new grey prediction model with error correction is proposed [18].

In addition to the optimization mentioned above, there are many effective improvement methods. These improvements have effectively improved the prediction accuracy of the grey model. Therefore, the grey model is widely used in energy [19], economic [20], environmental governance [21], and other related research studies. It is worth emphasizing that the advantage of the grey forecasting model compared with other forecasting methods lies in dealing with small sample problems. In fact, the grey accumulation generation plays an important role. Therefore, this paper proposes a novel aging accumulation operator to improve the traditional GM (1, 1) model. As a data preprocessing method, the aging accumulation operator can dynamically update the accumulation weight of data according to the time development. In addition, it can flexibly change the forecast trend by adjusting the aging parameter. Compared with the existing cumulative generation operator, it is an important innovation.

The other parts of this article are arranged as follows. Section 2 introduces the definition and properties of the aging accumulation operator. The aging accumulation grey model and its properties are proposed in Section 3. Section 4 introduces the optimization algorithm of the aging parameter and evaluates the performance evaluation of the proposed model by four cases. The conclusions are given in Section 5.

2. Definition and Properties of Aging Accumulation Operator

By analyzing the advantages and disadvantages of existing accumulation operators in the grey model, a new aging accumulation operator is defined. Besides, the operation details and related properties of the aging accumulation operator are introduced in detail.

2.1. Existing Accumulation Generation Operators. In the modeling process of the grey prediction model, cumulative generation is an important operation. By accumulating operation, scattered data can show certain regularity. The traditional grey prediction model uses 1-AGO to accumulate the original data. For example, the accumulation generated

sequence of the original sequence $\{x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(m)\}$ by 1-AGO is

$$x^{(1)}(i) = \sum_{k=1}^i x^{(0)}(k), \quad i = 1, 2, \dots, m. \quad (1)$$

It can be seen from equation (1) that 1-AGO treats all data indiscriminately. According to the new information priority principle, new data are more important than old data. When accumulating the original sequence, we should give full consideration to the new and old data. In other words, new information should be given more weight, and old information should be given less weight. Based on this consideration, many new cumulative generation operators have been proposed. Among them, the fractional accumulation operator is an important innovation [22]. Assuming that the original sequence is $\{x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(m)\}$, the fractional accumulation generated sequence is

$$x^{(r)}(i) = \sum_{k=1}^i \binom{i-k+r-1}{i-k} x^{(0)}(k), \quad (i = 1, 2, \dots, m), \quad (2)$$

where $\binom{r-1}{0} = 1$, $\binom{k-1}{k} = 0$, $\binom{k-i+r-1}{k-i} = ((r+k-i-1)(r+k-i-2) \cdots (r+1)r)/(k-i)!$.

Fractional accumulation operator can effectively allocate the weight of new and old data, thus describing the development trend of series more accurately. In addition, there are some other effective grey accumulation operators [23–25]. However, most operators can only increase the weight of new data. We hope that the cumulative generation operator can dynamically update the weights of all data according to time changes. Therefore, the aging accumulation operator is proposed.

2.2. The Aging Accumulation Operator. When accumulating data, the time value of data must be fully considered. Generally speaking, the timeliness of data is decreasing. Therefore, we propose a novel aging accumulation operator. It is defined as follows.

Definition 1. Assuming that $X^{(0)} = \{x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(m)\}$ is the original non-negative sequence, the aging cumulative sequence $X^{(\gamma)} = \{x^{(\gamma)}(1), x^{(\gamma)}(2), \dots, x^{(\gamma)}(m)\}$ of $X^{(0)}$ can be obtained by using the aging decreasing function $g(i)$ as the aging weighting. This transformation is called the aging accumulation operator, and its calculation formula is

$$x^{(\gamma)}(i) = \sum_{k=1}^i x^{(0)}(k)g(i-k), \quad i = 1, 2, \dots, m. \quad (3)$$

The matrix form of equation (3) is

$$X^{(\gamma)} = [x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(m)] \begin{bmatrix} g(0) & g(1) & \cdots & g(m-1) \\ 0 & g(0) & \cdots & g(m-2) \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & g(1) \\ 0 & 0 & \cdots & g(0) \end{bmatrix}, \quad (4)$$

where $g(i) = (\gamma/(i + \gamma))$ ($\gamma > 0$) and γ is called the aging parameter. It is used to adjust the aging change of data. Obviously, no matter what the value γ is, the aging of the latest data is 1. The smaller the value γ , the more time-sensitive the new information. On the contrary, the greater the value of γ , the more consistent the timeliness of new and old information. When γ tends to infinity, aging accumulation degenerates into traditional first-order accumulation.

Property 1. Adding new data to accumulation will dynamically update the accumulation weight of existing data.

Proof. According to the definition of the aging decreasing function $g(i) = \gamma/(i + \gamma)$, we have

- (1) $g(0) = \gamma/(0 + \gamma) = 1$.
- (2) $g(i) = \gamma/(i + \gamma)$ is the decreasing function of i .

That is to say, the aging value of the latest data is always 1, and the aging values of other data decrease with time. As shown in Figure 1, when calculating $x^{(\gamma)}(10)$, the aging

value corresponding to $x^{(0)}(10)$ is $g(0) = 1$, the aging value corresponding to $x^{(0)}(9)$ is $g(1) = \gamma/(1 + \gamma)$, and so on. Then, $x^{(\gamma)}(10) = \sum_{k=1}^{10} x^{(0)}(k)g(10 - k)$ can be obtained. To put it simply, the addition of new data will replace the aging value of the latest data, thus pushing down the aging value of all data. This method can dynamically update the weight of all data, and the latest data always keep a higher weight. The weight generated by this accumulation method is more following the law of the development of objective things. Besides, the data metabolism can be realized by flexibly adjusting the value of the aging parameter γ .

Property 2. Assuming $X^{(0)} = \{x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(m)\}$ is the original non-negative sequence and its aging accumulation sequence is $X^{(\gamma)} = \{x^{(\gamma)}(1), x^{(\gamma)}(2), \dots, x^{(\gamma)}(m)\}$, then $x^{(\gamma)}(i)$ ($i = 1, 2, \dots, m$) is the increasing function of the aging parameter γ .

Proof. According to Definition 1, $\forall i = 1, 2, \dots, m$, there is

$$\begin{aligned} x^{(\gamma)}(i) &= [x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(i)] [g(i-1) \ g(i-2) \ \cdots \ g(0)]^T \\ &= [x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(i)] \left[\frac{\gamma}{i+\gamma-1} \ \frac{\gamma}{i+\gamma-2} \ \cdots \ 1 \right]^T \\ &= [x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(i)] \left[1 - \frac{i-1}{i+\gamma-1} \ 1 - \frac{i-1}{i+\gamma-2} \ \cdots \ 1 \right]^T. \end{aligned} \quad (5)$$

Because $x^{(0)}(i)$ is non-negative and $1 - (k/(\gamma + k))$ ($k = i-1, i-2, \dots, 1$) is the increasing function of γ , $x^{(\gamma)}(i)$ ($i = 1, 2, \dots, m$) is the increasing function of γ .

Proof completed.

Lemma 1. Let $Y = \{y_1, y_2, \dots, y_n\}$ be a non-negative equidistant time series; then, $\Delta(k) = |y_{k+1} - y_k|$ represents the information difference between the data [26].

Property 3. Assuming that the original non-negative sequence $X^{(0)} = \{x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(m)\}$ increases monotonously and its aging accumulation sequence is $X^{(\gamma)} = \{x^{(\gamma)}(1), x^{(\gamma)}(2), \dots, x^{(\gamma)}(m)\}$, then we have

- (1) The aging accumulation sequence $x^{(\gamma)}(i)$ ($i = 1, 2, \dots, m$) is the increasing function of i .
- (2) The information difference $\Delta(i) = |x^{(\gamma)}(i+1) - x^{(\gamma)}(i)|$ between $x^{(\gamma)}(i+1)$ and $x^{(\gamma)}(i)$ is the increasing function of γ .

Proof. (1) $X^{(0)} = \{x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(m)\}$ is a monotonically increasing non-negative sequence; then, we have

$$x^{(0)}(i) - x^{(0)}(i-1) > 0, \quad (i = 2, 3, \dots, m). \quad (6)$$

Also, for $1 \leq i < m$, we have $g(x) = (\gamma/(i + \gamma)) > 0$. Then, we can obtain

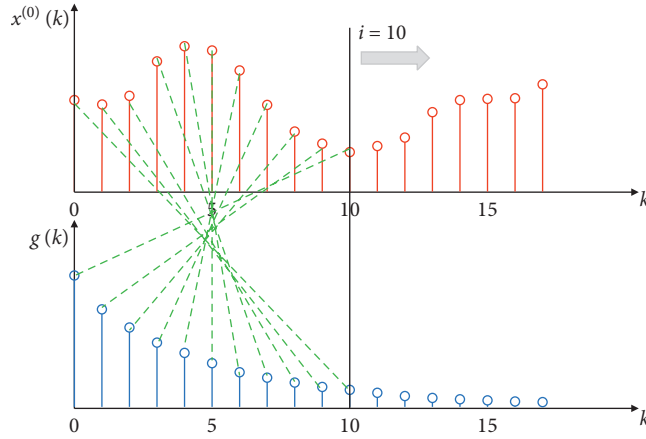


FIGURE 1: Dynamic update process of data aging.

$$\begin{aligned}
 & x^{(\gamma)}(i+1) - x^{(\gamma)}(i) \\
 &= \sum_{k=1}^{k=i+1} x^{(0)}(k)g(i+1-k) - \sum_{k=1}^{k=i} x^{(0)}(k)g(i-k) \\
 &= x^{(0)}(1)g(i) + \sum_{k=2}^{k=i+1} [(x^{(0)}(k) - x^{(0)}(k-1))g(i+1-k)] > 0.
 \end{aligned} \tag{7}$$

(2) From (1), we know that $\forall i = 2, 3, \dots, m$, we have

$$\begin{aligned}
 \Delta(i) &= |x^{(\gamma)}(i+1) - x^{(\gamma)}(i)| = x^{(\gamma)}(i+1) - x^{(\gamma)}(i) \\
 &= [x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(i+1)] [g(i) - g(i-1), g(i-1) - g(i-2), \dots, g(0)]^T \\
 &= [x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(i+1)] \left[\frac{-\gamma}{(i+\gamma)(i+\gamma-1)}, \frac{-\gamma}{(i+\gamma-1)(i+\gamma-2)}, \dots, 1 \right]^T.
 \end{aligned} \tag{8}$$

Obviously, $(-\gamma/(k+\gamma)(k+\gamma-1))$ ($k = 1, 2, \dots, i$) is the increasing function of γ and $x^{(0)}(k)$ ($k = 1, 2, \dots, m$) > 0 . So, we can conclude that the information difference $\Delta(i)$ is the increasing function of γ .

Proof completed.

According to the principle of difference information, information comes from the difference [27]. Fully mining the information difference of sequence can maximize the value of data. However, in practical application, data fluctuation may lead to deviation of information difference. Based on the modeling mechanism of the grey model, the fluctuation of the older data will cause greater deviation. As an improvement, the introduction of the aging parameter γ can weaken extreme interference while retaining important difference information.

3. The Aging Accumulation Grey Model

Based on the aging accumulation operator, a new aging accumulation grey model AGM (1, 1) is proposed in this section. In order to highlight the advantages of the AGM

(1, 1) model and understand its applicable scope, the validity of initial value, monotonicity, prediction trend, and reduction error of this model are analyzed and discussed.

3.1. The Definition of the Aging Accumulation Grey Model.

The traditional GM (1, 1) model uses the first-order cumulative generation operation to reduce the random disturbance, which can improve the model effect to a certain extent. However, traditional 1-AGO ignores the difference of timeliness between old and new data. Therefore, in this section, the GM (1, 1) model is optimized by using the aging accumulation operator, and the aging accumulation grey model is obtained. It is defined as follows.

Definition 2. Assuming that $\{x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(m)\}$ is the original sequence, then the aging accumulation sequence of the original sequence can be obtained as $\{x^{(\gamma)}(1), x^{(\gamma)}(2), \dots, x^{(\gamma)}(m)\}$ by Definition 1. Then, AGM (1, 1) can be written as

$$\frac{dx^{(\gamma)}}{dt} + ax^{(\gamma)} = b, \quad (9)$$

where a is the development coefficient and b is the grey action quantity. The solution of whitening differential equation (9) is $x^{(\gamma)}(i) = (x^{(0)}(1) - (b/a))e^{-a(i-1)} + (b/a)$. Its parameters generally use the least-squares solution of the AGM (1, 1) model. The least-squares estimation of the AGM (1, 1) model satisfies

$$\begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix} = (B^T B)^{-1} B^T Y, \quad (10)$$

where

$$B = \begin{bmatrix} -0.5(x^{(\gamma)}(1) + x^{(\gamma)}(2)) & 1 \\ -0.5(x^{(\gamma)}(2) + x^{(\gamma)}(3)) & 1 \\ \vdots & \vdots \\ -0.5(x^{(\gamma)}(m-1) + x^{(\gamma)}(m)) & 1 \end{bmatrix}, \quad (11)$$

$$Y = \begin{bmatrix} x^{(\gamma)}(2) - x^{(\gamma)}(1) \\ x^{(\gamma)}(3) - x^{(\gamma)}(2) \\ \vdots \\ x^{(\gamma)}(m) - x^{(\gamma)}(m-1) \end{bmatrix}.$$

Inputting \hat{a} and \hat{b} into the solution of the whitening differential equation, the time response equation can be obtained as

$$\hat{x}^{(\gamma)}(i) = \left(x^{(0)}(1) - \frac{\hat{b}}{\hat{a}} \right) e^{-\hat{a}(i-1)} + \frac{\hat{b}}{\hat{a}}, \quad i = 1, 2, \dots, m, \dots, m + mf, \quad (12)$$

where mf represents the number to be predicted. Then, the fitted and predicted values $\hat{X}^{(0)} = \{\hat{x}^{(0)}(1), \hat{x}^{(0)}(2), \dots, \hat{x}^{(0)}(m + mf)\}$ can be obtained as

$$\hat{X}^{(0)} = [\hat{x}^{(\gamma)}(1), \hat{x}^{(\gamma)}(2), \dots, \hat{x}^{(\gamma)}(m + mf)] \begin{bmatrix} g(0) & g(1) & \dots & g(m + mf - 1) \\ 0 & g(0) & \dots & g(m + mf - 2) \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & g(1) \\ 0 & 0 & \dots & g(0) \end{bmatrix}^{-1}. \quad (13)$$

3.2. The Effectiveness of the Initial Value by AGM (1, 1). The principle of minimum information is one of the six axioms of the grey theory, which holds that the existing information must be fully utilized [27]. Therefore, it is necessary to study the utilization degree of data by the grey prediction model. It has been proved that the initial value of the traditional grey prediction model is invalid [28]. This paper will prove that the initial value of the AGM(1, 1) model is valid by Property 4.

Property 4. Assuming that the fitting value $\hat{X}^{(0)} = \{\hat{x}^{(0)}(1), \hat{x}^{(0)}(2), \dots, \hat{x}^{(0)}(m)\}$ of $X^{(0)} = \{x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(m)\}$ is obtained by the AGM (1, 1) model, then $\hat{X}^{(0)} = \{\hat{x}^{(0)}(1), \hat{x}^{(0)}(2), \dots, \hat{x}^{(0)}(m)\}$ will change with the change of $x^{(0)}(1)$.

Proof. A case from [29] was used for empirical analysis. The world's renewable energy is taken as the raw data:

$$X^{(0)} = \{124.1, 144, 170.6, 203.6, 238.8, 282.5, 319.5, 368.5, 416.8, 490.2, 561.3\}. \quad (14)$$

Set $\gamma = 10$, and we have

$$g(x) = \{1.00, 0.91, 0.83, 0.77, 0.71, 0.67, 0.63, 0.59, 0.56, 0.53, 0.50\}. \quad (15)$$

With equation (4), the aging accumulation sequence of $X^{(0)}$ can be calculated as

$$X^{(10)} = \{124.10, 256.82, 404.93, 574.15, 765.47, 986.08, 1227.35, 1500.23, 1801.94, 2155.87, 2555.91\}. \quad (16)$$

Then, we have

$$\begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix} = (B^T B)^{-1} B^T Y = \begin{pmatrix} -0.12 \\ 108.58 \end{pmatrix}. \quad (17)$$

The time response function $\hat{x}^{(10)}(i) = (124.10 - (108.58/-0.12))e^{0.12(i-1)} + (108.58/-0.12)$ can be obtained by substituting \hat{a} and \hat{b} into equation (12). Then, the time response sequence can be calculated as

$$\hat{X}^{(10)} = \{124.10, 255.73, 404.47, 572.55, 762.48, 977.10, 1219.62, 1493.67, 1803.35, 2153.29, 2548.72\}. \quad (18)$$

Finally, the fitted values are obtained by inverse accumulation as

$$\hat{X}^{(0)} = \{124.10, 142.91, 171.14, 202.42, 237.28, 276.26, 319.97, 369.07, 424.29, 486.48, 556.56\}. \quad (19)$$

However, if the original sequence is

$$X^{(0)} = \{80.1, 144, 170.6, 203.6, 238.8, 282.5, 319.5, 368.5, 416.8, 490.2, 561.3\}, \quad (20)$$

the fitted values are obtained by the same method as

$$\hat{X}^{(0)} = \{80.10, 141.67, 170.48, 202.22, 237.42, 276.63, 320.44, 369.51, 424.53, 486.33, 555.78\}. \quad (21)$$

Proof completed.

3.3. The Monotonicity and Forecast Trend of AGM (1, 1). From equation (13), we conclude that the fitted and predicted values $\{\hat{x}^{(0)}(1), \hat{x}^{(0)}(2), \dots, \hat{x}^{(0)}(m + mf)\}$ are related to model parameters. Therefore, the monotonicity of the predictive value is uncertain and data-driven.

We consider an example from [22]. The data from 2001 to 2007 are used for fitting, and the data from 2008 to 2009 are used for testing. The original data and model results are shown in Table 1. As for the original data, the data increased from 2001 to 2006 but decreased from 2006 to 2007, that is to say, the latest data showed a downward trend. However, the results of the traditional GM (1, 1) model increased monotonously from 2001 to 2009, which did not conform to the objective law. On the contrary, the AGM (1, 1) model perceived the trend change from 2006 to 2007, and its model results showed a downward trend from 2006 to 2009. It

shows that the AGM (1, 1) model has better performance and pays more attention to new information.

In addition, by adjusting the aging parameter γ , the prediction trend of the AGM (1, 1) model can be adjusted flexibly. Figure 2 shows the results of the AGM (1, 1) model when aging parameters are 1, 3, 5, and 10, respectively. With the increase of the aging parameter, the prediction trend tends to be flat. On the contrary, the smaller the aging parameter, the steeper the prediction trend. The advantage of this flexible adjustment mechanism is that it can be analyzed by combining subjective experience with objective data, which is very suitable for forecasting uncertain systems.

3.4. The Relationship between the Error and Aging Parameter γ . Accumulation operation can make scattered data show a certain trend, but it inevitably leads to reductive error. In this section, we will further study the relationship between reductive errors and the aging parameter γ of the AGM (1, 1)

TABLE 1: Model results with different aging parameters.

Year	Actual value	GM (1, 1)	AGM ⁽¹⁾ (1, 1)	AGM ⁽³⁾ (1, 1)	AGM ⁽⁵⁾ (1, 1)	AGM ⁽¹⁰⁾ (1, 1)
2001	247.84	247.84	247.84	247.84	247.84	247.84
2002	273.02	278.58	267.66	270.72	273.1685	275.9409
2003	289.01	282.20	286.13	284.76	283.7412	282.7841
2004	285.21	285.87	294.04	291.63	289.8058	287.7632
2005	288.82	289.59	295.59	294.02	292.7876	291.275
2006	297.08	293.36	293.39	293.49	293.5915	293.6141
2007	293.66	297.17	289.05	291.05	292.8283	295.005
2008	290.40	301.04	283.57	287.34	290.9281	295.6225
2009	279.14	304.95	277.55	282.82	288.2036	295.6053

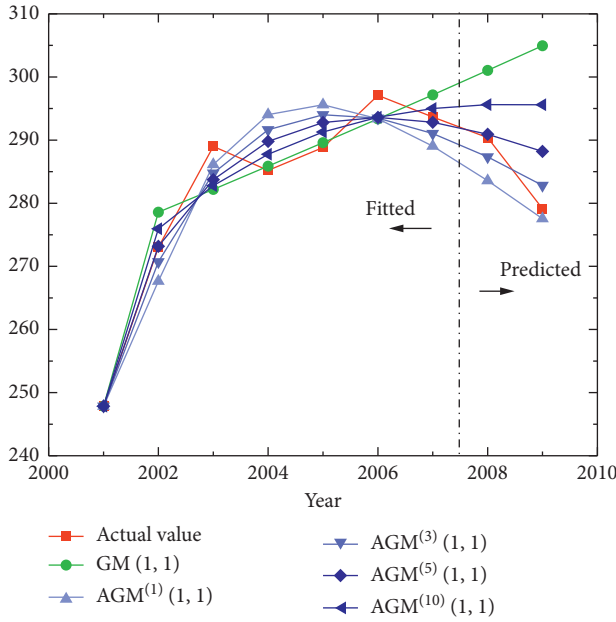


FIGURE 2: The results of different grey models.

model, so as to effectively control the errors caused by accumulative reduction operations.

Theorem 1. Assume that $X^{(\gamma)} = \{x^{(\gamma)}(1), x^{(\gamma)}(2), \dots, x^{(\gamma)}(m)\}$ is the accumulated sequence of the original sequence $X^{(0)} = \{x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(m)\}$ by AGM (1, 1) model with aging parameter γ . $\hat{X}^{(0)} = \{\hat{x}^{(0)}(1), \hat{x}^{(0)}(2), \dots, \hat{x}^{(0)}(m)\}$ is the cumulative reduction sequence of time response sequence $\{\hat{x}^{(\gamma)}(1), \hat{x}^{(\gamma)}(2), \dots, \hat{x}^{(\gamma)}(m)\}$.

If $|\hat{x}^{(\gamma)}(i) - \hat{x}^{(0)}(i)| < \varepsilon$ ($1 < i \leq m$), then $|\hat{x}^{(0)}(i) - x^{(0)}(i)| < \varepsilon * \sum_{j=1}^i |R(j, i)|$ ($1 < i \leq m$), where

$$R = \begin{bmatrix} g(0) & g(1) & \cdots & g(i-1) \\ 0 & g(0) & \cdots & g(i-2) \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & g(1) \\ 0 & 0 & \cdots & g(0) \end{bmatrix}^{-1}. \quad (22)$$

Proof. According to the definition of AGM (1, 1), we have

$$\begin{aligned} & [\hat{x}^{(0)}(1) - x^{(0)}(1), \hat{x}^{(0)}(2) - x^{(0)}(2), \dots, \hat{x}^{(0)}(m) - x^{(0)}(m)] \\ &= [\hat{x}^{(\gamma)}(1) - \hat{x}^{(\gamma)}(1), \hat{x}^{(\gamma)}(2) - \hat{x}^{(\gamma)}(2), \dots, \hat{x}^{(\gamma)}(m) - \hat{x}^{(\gamma)}(m)] \begin{bmatrix} g(0) & g(1) & \cdots & g(m-1) \\ 0 & g(0) & \cdots & g(m-2) \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & g(1) \\ 0 & 0 & \cdots & g(0) \end{bmatrix}^{-1}. \end{aligned} \quad (23)$$

Let

$$R = \begin{bmatrix} g(0) & g(1) & \cdots & g(i-1) \\ 0 & g(0) & \cdots & g(i-2) \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & g(1) \\ 0 & 0 & \cdots & g(0) \end{bmatrix}^{-1}.$$

For $\forall i = 1, 2, \dots, m$, we have

$$\begin{aligned}
& |\hat{x}^{(0)}(i) - x^{(0)}(i)| \\
&= \left| [\hat{x}^{(\gamma)}(1) - \hat{x}^{(\gamma)}(1), \hat{x}^{(\gamma)}(2) - \hat{x}^{(\gamma)}(2), \dots, \hat{x}^{(\gamma)}(i) - \hat{x}^{(\gamma)}(i)] * R(\sim, i) \right| \\
&< \varepsilon * \sum_{j=1}^i |R(j, i)|,
\end{aligned} \tag{24}$$

where $R(\sim, i)$ represents the column i of the matrix R and $R(j, i)$ represents the element in row j and column i of the matrix R .

Proof completed.

When determining the value of the aging parameter γ , we should minimize $\|R\|_1$ as much as possible to avoid large reductive errors.

4. Performance Evaluation of the Proposed Model

The definition and related properties of the AGM (1, 1) model have been introduced above. In fact, the performance of AGM (1, 1) depends on the value of the aging parameter γ . In this section, an optimization algorithm is introduced to determine the optimal aging parameter, and four examples are used to prove the effectiveness of the proposed AGM (1, 1) model.

4.1. Optimization Algorithm of the Optimal Aging Parameter γ . Taking the average absolute percentage error (MAPE = $(1/m) \sum_{k=1}^m |(\hat{x}_1^{(0)}(k) - x_1^{(0)}(k))/x_1^{(0)}(k)| \times 100\%$) as the optimization objective and the main formula of the AGM (1, 1) model as the constraint condition, the following nonlinear programming is constructed.

$$\begin{aligned}
\min Z(\gamma) &= \frac{1}{m} \sum_{k=1}^m \left| \frac{\hat{x}_1^{(0)}(k) - x_1^{(0)}(k)}{x_1^{(0)}(k)} \right| \times 100\%, \\
\text{s.t. } \left\{ \begin{aligned} B &= \begin{bmatrix} -0.5(x^{(\gamma)}(1) + x^{(\gamma)}(2)) & 1 \\ -0.5(x^{(\gamma)}(2) + x^{(\gamma)}(3)) & 1 \\ \vdots & \vdots \\ -0.5(x^{(\gamma)}(m-1) + x^{(\gamma)}(m)) & 1 \end{bmatrix}, \\ Y &= \begin{bmatrix} x^{(\gamma)}(2) - x^{(\gamma)}(1) \\ x^{(\gamma)}(3) - x^{(\gamma)}(2) \\ \vdots \\ x^{(\gamma)}(m) - x^{(\gamma)}(m-1) \end{bmatrix}, \\ [\hat{a}, \hat{b}]^T &= (B^T B)^{-1} B^T Y, \\ \hat{x}^{(\gamma)}(i) &= \left(x^{(0)}(1) - \frac{\hat{b}}{\hat{a}} \right) e^{-\hat{a}(i-1)} + \frac{\hat{b}}{\hat{a}} \quad i = 1, 2, \dots, m. \end{aligned} \right. \tag{25}
\end{aligned}$$

Particle swarm optimization (PSO) is a mature intelligent optimization algorithm, which is derived from simulating the group behavior of bird foraging. Particle swarm optimization algorithm has the characteristics of simple operation and rapid convergence, so it has been widely used [30–32]. Figure 3 shows the flowchart for determining the optimal aging parameter γ . The specific steps of the algorithm are shown in Algorithm 1.

4.2. Application and Analysis. In addition to the MAPE, we applied the mean absolute error (MAE) and the root mean square error (RMSE) to measure the predictive performance of the AGM (1, 1) model. They are defined as follows:

$$\begin{aligned}
\text{MAE} &= \frac{1}{m} \sum_{k=1}^m |\hat{x}_1^{(0)}(k) - x_1^{(0)}(k)|, \\
\text{RMSE} &= \sqrt{\frac{1}{m} \sum_{k=1}^m (\hat{x}_1^{(0)}(k) - x_1^{(0)}(k))^2}. \tag{26}
\end{aligned}$$

Case 1. Forecasting logistics demand in Jiangsu province.

This example comes from [7]. This is a case with small sample size. Similar to [7], the data from 2000 to 2005 are used for fitting, and the data of 2006 are used for the test. The smaller the sample size, the higher the prediction accuracy of the traditional GM (1, 1). Therefore, this paper compares AGM (1, 1) with GM (1, 1) models with different sample sizes, and Table 2 shows the comparison results. In the stage of fitting and testing, AGM (1, 1) with six sample sizes gets better results than GM (1, 1) with four sample sizes. The results show that with the increase of the number of samples, the special metabolic function of AGM(1, 1) can reduce the interference of old data and improve the prediction performance of the model.

Case 2. The example for the waste volume sequence of TV in China.

This example is from [33]. The data are the waste volume of TV (10000 units) in China. It is a steady growth sequence. The in-sample data and out-of-sample data are the same as [33]. To prove the performance of AGM(1,1) model, eight commonly used forecasting methods are used for comparison. The errors of the nine models are shown in Table 3. In the fitting and testing stage, MAPE, MAE, and RMSE of AGM (1, 1) are the lowest, which shows that the proposed AGM (1, 1) model has excellent performance in dealing with medium and long-term stationary sequences.

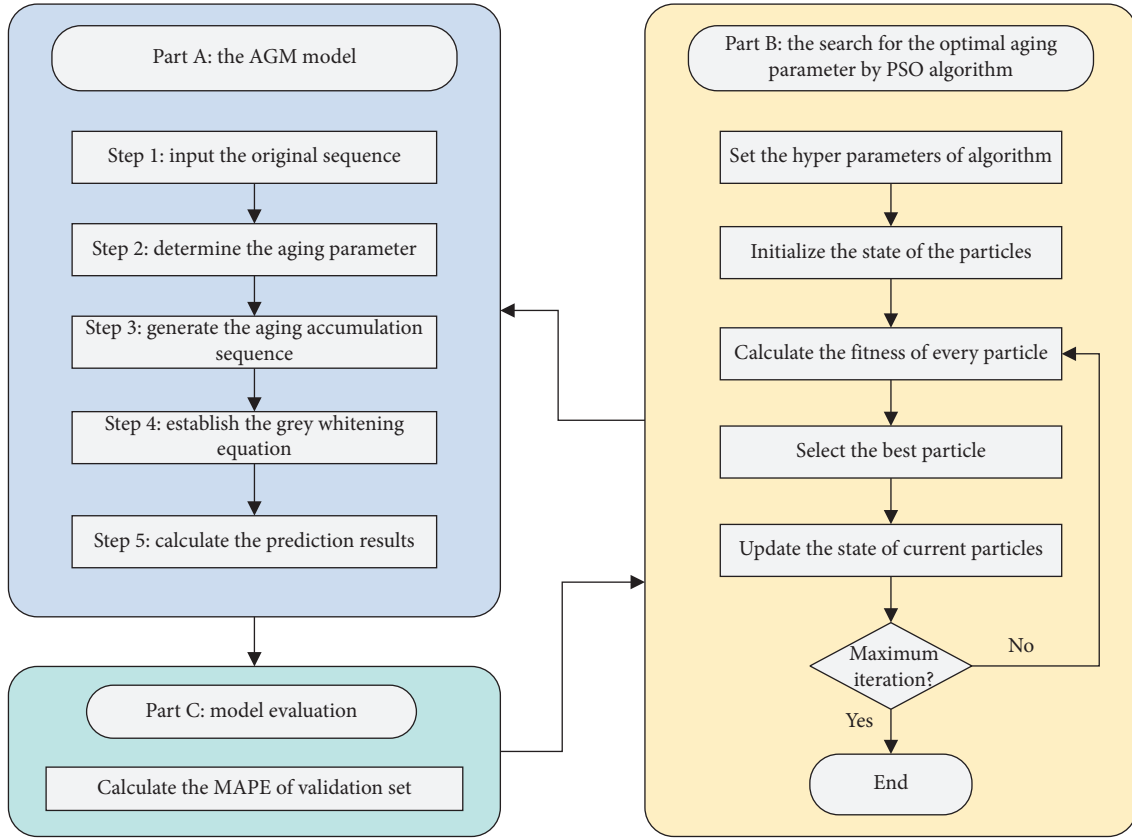


FIGURE 3: Flowchart for determining the optimal aging parameter.

Input: the sample set $X^{(0)} = \{x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(m)\}$

Output: the optimal value of γ

(1) Initialize parameters in the PSO algorithm:

Particle number N , dimension D , maximum generation T , learning factor c_1, c_2 , inertia weight η .

(2) Initialize the position w_i and velocity v_i

(3) for $j = 1: T$ do

(4) for $i = 1: N$ do

(5) Calculate $X^{(\gamma)} = \{x^{(\gamma)}(1), x^{(\gamma)}(2), \dots, x^{(\gamma)}(m)\}$ by Definition 1;

(6) Calculate $\{\hat{b}_1, \hat{b}_2, \dots, \hat{b}_n, \hat{u}\}$ by equation (10);

(7) Compute $\hat{X}_1^{(0)}$ using equation (13);

(8) Compute the fitness function $Z(\gamma)$;

(9) Update the position and velocity of particles

$$v_i = \eta v_i + c_1 r_1 (Q_p - w_i) + c_2 r_2 (Q_g - w_i);$$

$$w_i = w_i + v_i.$$

where γ are random vectors and belong to $[0, 10]$; Q_p and Q_g represent the individual optimal position and the global optimal position, respectively.

(10) end for

(11) end for

(12) return optimal value of γ

ALGORITHM 1: Optimization algorithm of the optimal aging parameter γ (solution to optimize the optimal aging parameter γ).

Case 3. Predicting foreign tourists to China.

This example is from [34]. The data are the annual historical data of tourists from Russia and Singapore from 2003 to 2017 in China. They are long and fluctuating sequences. Similar to [34], the data from 2003 to 2015 were

used for fitting, and the data from 2016 and 2017 were used for the test. Table 4 summarizes the test results of 12 different prediction methods. In Russia, the MAPE of AGM (1, 1) is only lower than that of F-OGMp (1, 1), and MAE and RMSE are the lowest. For Singapore, the overall performance of

TABLE 2: The fitting and predicted values of grey models on different samples.

Serial number	Actual value	GM ₆ (1, 1)	GM ₅ (1, 1)	GM ₄ (1, 1)	AGM (1, 1)
2000	132.4	132.4			132.40
2001	144.6	142	144.6		142.75
2002	156.3	157.3	155.3	156.3	155.57
2003	173.7	174.2	173	172.1	171.45
2004	190.2	193	192.8	192.5	191.10
2005	216.7	213.8	214.8	215.3	215.41
MAPE		1.12	0.81	0.93	0.69
MAE		1.63	1.24	1.33	1.17
RMSE		2.01	1.54	1.57	1.38
2006	249.4	236.8	239.4	240.8	245.50
MAPE		5.04	4.03	3.46	1.57
MAE		12.6	10	8.6	3.90

The smallest values of these model errors are in bold.

TABLE 3: The fitting and predicted values of nine models.

Model	Fitting value			Testing value	
	MAPE	MAE	RMSE	MAPE	MAE
AGM (1, 1)	0.4	35.92	47.95	0.11	13.03
EFGM	0.58	50.81	69.04	1.13	134.43
LSSVR	1.06	91.36	131.08	7.96	949.89
ANN	1.61	130.34	147.58	0.63	75.49
ARIMA	0.63	48.08	54.82	6.88	821.07
DGM (1, 1)	1.65	135.68	168.63	3.64	434.92
Verhulst	22.35	1597.68	2161.39	12.38	1477.09
GM (1, 1)	1.64	135.21	168.55	3.57	426.49
FGM (1, 1)	0.85	72.26	87.48	1.46	174.66

The smallest values of these model errors are in bold.

TABLE 4: The predicted values of twelve models.

Model	Russia			Singapore		
	MAPE	MAE	RMSE	MAPE	MAE	RMSE
TGM (1, 1)	9.23	20.83	23.64	1.39	1.28	1.78
Original GM (1, 1)	12.04	49.77	52.17	15.48	14.44	14.45
Optimized GM (1, 1)	11.95	49.8	51.18	15.48	14.44	14.45
Original NGBM (1, 1)	57.21	250.43	256.94	14.85	13.9	14.98
Optimized NGBM (1, 1)	10.07	44.86	48.14	1.63	1.51	1.88
ARIMA	25.29	104.66	109.39	0.3	0.28	0.38
BPN	9.46	43.62	53.28	2.35	8.72	8.83
Optimized GMp (1, 1)	10.09	45.43	50.64	1.58	1.48	1.48
Original F-GM (1, 1)	17.32	76.29	79.39	5.65	5.31	7.15
Optimized F-GM (1, 1)	17.34	76.32	79.3	5.67	5.34	7.47
Optimized F-GMp (1, 1)	13.19	52.21	65.27	2.89	2.72	3.85
F-OGMp (1, 1)	6.12	27.44	30.05	1.35	1.26	1.39

AGM (1, 1) is only lower than that of ARIMA and equivalent to F-OGMp (1, 1). This shows that the AGM (1, 1) model can well predict the development trend of medium and long-term wave series.

Case 4. Comparison of aging accumulation operator and other existing operators.

Accumulation operation is an important operation of the grey prediction model. By accumulating the data, the interference of random disturbance can be effectively reduced, and the scattered data show a certain rule. To prove the effectiveness of the proposed aging accumulation operator, seven existing grey accumulation operators such as

the traditional first-order accumulation generation operation, the damping accumulation generation operation [25], the adjacent accumulation generation operation [35], the first-order new information priority accumulation generation operation [23], the conformable fractional accumulation generation operation [36], the fractional order accumulation generation operation [22], and fractional Hausdorff accumulation generation operation [37] are compared and analyzed. The data of forecasting competition are often used to verify the performance of forecasting methods [38, 39]. Take the first nine data of the *N7* series in the *M3* prediction contest as an example. The first seven data are used for fitting, and the last two data are used for testing.

TABLE 5: Grey models with different grey accumulation operators.

	GM (1, 1)	AGM (1, 1)	DAGM (1, 1)	AGM ₁ (1, 1)	NGM (1, 1)	CFGM (1, 1)	FGM (1, 1)	FHGM (1, 1)
1	2399.26	2399.26	2399.26	2399.26	2399.26	2399.26	2399.26	2399.26
2	2910.52	2989.36	2910.52	2989.36	2846.92	2910.52	2895.97	2895.97
3	3126.62	3178.09	3192.14	3178.09	3197.53	3184.44	3202.81	3202.81
4	3475.14	3378.73	3433.73	3378.73	3472.12	3427.27	3440.37	3440.37
5	3750.96	3592.04	3645.29	3592.04	3687.18	3642.53	3644.95	3644.95
6	3752.72	3818.81	3833.15	3818.81	3855.62	3833.37	3830.51	3828.03
7	4004.02	4059.91	4001.66	4059.91	3987.53	4002.54	4004.02	4004.02
MAPE		2.07	1.19	2.07	1.34	1.19	1.26	1.26
MAE		72.52	42.20	72.52	45.81	42.32	44.19	43.31
RMSE		85.28	58.12	85.28	58.56	58.43	59.18	59.66
8	3737.38	4316.22	4153.93	4316.22	4090.85	4152.51	4169.51	4172.66
9	4263.98	4588.72	4292.30	4588.72	4171.77	4285.46	4329.50	4336.39
MAPE		11.55	5.9	11.55	5.81	5.81	6.55	6.67
MAE		451.79	222.43	451.79	222.84	218.3	248.83	253.84
RMSE		469.31	295.23	469.31	258.31	293.93	309.05	312.02

Note. AGM₁ (1, 1) is the adjacent accumulation grey model. The smallest values of these model errors are in bold.

Table 5 summarizes the grey model results of different cumulative generation operators. The MAPE, MAE, and RMSE of the AGM (1, 1) model are 1.19%, 42.20, and 58.12, respectively. They are all the lowest in the fitting stage. In the testing stage, AGM (1, 1) is also superior to most models. Actually, MAPE, MAE, and RMSE are only worse than those of NGM (1, 1) and AGM₁ (1, 1) in the prediction stage, which shows that the proposed aging accumulation operator is effective.

5. Conclusions

In this paper, a novel aging accumulation operator is proposed. Different from the existing grey accumulation operator, this operator determines the accumulation weight of data at different times from back to front. The addition of new data will push the old data to roll back so that the timeliness of data can be updated dynamically with the change of the system. The aging accumulation operator is introduced into the grey model, and a new aging accumulation grey model AGM (1, 1) is obtained. Compared with the traditional grey model, AGM (1, 1) can reduce the interference of old data and improve the prediction accuracy of the model by adjusting the aging parameter γ . In addition, the prediction trend of the AGM (1, 1) model is adjustable. The effectiveness of AGM (1, 1) is proved by four case studies, and the following conclusions are obtained:

- (1) The introduction of the aging parameter γ overcomes the problem that the prediction accuracy of the traditional GM (1, 1) model decreases with the increase of sample size. The AGM (1, 1) model can effectively adjust the aging weight of new and old information and get more accurate fitting and prediction results.
- (2) The AGM (1, 1) model not only effectively improves the short-term forecasting ability of the grey model but also outperforms most existing forecasting methods when dealing with medium and long-term smooth and fluctuating series.

- (3) As an improvement of the traditional GM (1, 1) model, the proposed aging accumulation operator is superior to most existing grey accumulation operators.

In a word, the proposed aging accumulation operator and aging accumulation grey model are very effective. Because of the excellent performance of the aging accumulation operator, it can also be used to improve other grey models and forecasting methods. Besides, when defining the aging accumulation operator, the inverse proportional function is selected as the aging decreasing function, and the better aging decreasing function can be mined.

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 study was supported by the Excellent Young Scientist Foundation of Hebei Education Department (SLRC2019001).

References

- [1] N. Xie and R. Wang, "A historic review of grey forecasting models," *Journal of Grey System*, vol. 29, no. 4, pp. 1–29, 2017.
- [2] H. Hewamalage, C. Bergmeir, and K. Bandara, "Recurrent neural networks for time series forecasting: current status and future directions," *International Journal of Forecasting*, vol. 37, no. 1, pp. 388–427, 2021.
- [3] S. Ryu, J. Noh, and H. Kim, "Deep neural network based demand side short term load forecasting," *Energies*, vol. 10, no. 1, p. 3, 2016.
- [4] E. M. de Oliveira and F. L. Cyrino Oliveira, "Forecasting mid-long term electric energy consumption through bagging

- ARIMA and exponential smoothing methods,” *Energy*, vol. 144, pp. 776–788, 2018.
- [5] R. Moazen-zadeh, B. Mohammadi, S. Shamshirband, and K.-w. Chau, “Coupling a firefly algorithm with support vector regression to predict evaporation in northern Iran,” *Engineering Applications of Computational Fluid Mechanics*, vol. 12, no. 1, pp. 584–597, 2018.
 - [6] Z. Zeng and M. Li, “Bayesian median autoregression for robust time series forecasting,” *International Journal of Forecasting*, vol. 37, no. 2, pp. 1000–1010, 2020.
 - [7] L. Wu, S. Liu, L. Yao, and S. Yan, “The effect of sample size on the grey system model,” *Applied Mathematical Modelling*, vol. 37, no. 9, pp. 6577–6583, 2013.
 - [8] X. Ma, W. Wu, and Y. Zhang, “Improved GM(1,1) model based on Simpson formula and its applications,” *Journal of Grey System*, vol. 31, no. 4, pp. 33–46, 2019.
 - [9] S. Li, B. Zeng, X. Mai et al., “A novel grey model with A three-parameter background value and its application in forecasting average annual water consumption per capita in urban areas along the Yangtze river basin,” *Journal of Grey System*, vol. 32, no. 1, pp. 118–132, 2020.
 - [10] Y.-C. Hu, P. Jiang, and P.-C. Lee, “Forecasting tourism demand by incorporating neural networks into Grey-Markov models,” *Journal of the Operational Research Society*, vol. 70, no. 1, pp. 12–20, 2019.
 - [11] N.-m. Xie and S.-f. Liu, “Discrete grey forecasting model and its optimization,” *Applied Mathematical Modelling*, vol. 33, no. 2, pp. 1173–1186, 2009.
 - [12] C. Zheng, W.-Z. Wu, W. Xie, Q. Li, and T. Zhang, “Forecasting the hydroelectricity consumption of China by using a novel unbiased nonlinear grey Bernoulli model,” *Journal of Cleaner Production*, vol. 278, Article ID 123903, 2021.
 - [13] X. Ma and Z. Liu, “The GMC(1, n) model with optimized parameters and its application,” *Journal of Grey System*, vol. 29, no. 4, pp. 122–138, 2017.
 - [14] L. Wu, S. Liu, and Y. Yang, “Using the fractional order method to generalize strengthening buffer operator and weakening buffer operator,” *IEEE/CAA Journal of Automatica Sinica*, vol. 5, no. 6, pp. 1074–1078, 2018.
 - [15] C. Li, Y. Yang, and S. Liu, “Comparative analysis of properties of weakening buffer operators in time series prediction models,” *Communications in Nonlinear Science and Numerical Simulation*, vol. 68, pp. 257–285, 2019.
 - [16] L.-Y. He, L.-L. Pei, and Y.-H. Yang, “An optimised grey buffer operator for forecasting the production and sales of new energy vehicles in China,” *Science of the Total Environment*, vol. 704, Article ID 135321, 2020.
 - [17] A. Bezuglov and G. Comert, “Short-term freeway traffic parameter prediction: application of grey system theory models,” *Expert Systems with Applications*, vol. 62, pp. 284–292, 2016.
 - [18] D. Zhou, A. Al-Durra, K. Zhang, A. Ravey, and F. Gao, “A robust prognostic indicator for renewable energy technologies: a novel error correction grey prediction model,” *IEEE Transactions on Industrial Electronics*, vol. 66, no. 12, pp. 9312–9325, 2019.
 - [19] S. Ding, R. Li, and Z. Tao, “A novel adaptive discrete grey model with time-varying parameters for long-term photovoltaic power generation forecasting,” *Energy Conversion and Management*, vol. 227, Article ID 113644, 2021.
 - [20] E. Javanmardi and S. Liu, “Exploring grey systems theory-based methods and applications in analyzing socio-economic systems,” *Sustainability*, vol. 11, no. 15, p. 4192, 2019.
 - [21] W. Zhou, X. Wu, S. Ding, and Y. Cheng, “Predictive analysis of the air quality indicators in the Yangtze River Delta in China: an application of a novel seasonal grey model,” *Science of The Total Environment*, vol. 748, Article ID 141428, 2020.
 - [22] L. Wu, S. Liu, L. Yao, S. Yan, and D. Liu, “Grey system model with the fractional order accumulation,” *Communications in Nonlinear Science and Numerical Simulation*, vol. 18, no. 7, pp. 1775–1785, 2013.
 - [23] L. Wu and Z. Zhang, “Grey multivariable convolution model with new information priority accumulation,” *Applied Mathematical Modelling*, vol. 62, pp. 595–604, 2018.
 - [24] W. Wu, X. Ma, Y. Zhang, W. Li, and Y. Wang, “A novel conformable fractional non-homogeneous grey model for forecasting carbon dioxide emissions of BRICS countries,” *Science of The Total Environment*, vol. 707, Article ID 135447, 2020.
 - [25] L. Liu, Y. Chen, and L. Wu, “The damping accumulated grey model and its application,” *Communications in Nonlinear Science and Numerical Simulation*, vol. 95, Article ID 105665, 2021.
 - [26] S. Liu and Y. Lin, *Grey Systems: Theory and Applications*, Springer, London, UK, 2010.
 - [27] J. Deng, “Introduction to grey system theory,” *Journal of Grey System*, vol. 1, no. 1, pp. 1–24, 1989.
 - [28] L. Wu, S. Liu, Z. Fang, and H. Xu, “Properties of the GM(1, 1) with fractional order accumulation,” *Applied Mathematics and Computation*, vol. 252, pp. 287–293, 2015.
 - [29] L. Liu and L. Wu, “Forecasting the renewable energy consumption of the European countries by an adjacent non-homogeneous grey model,” *Applied Mathematical Modelling*, vol. 89, pp. 1932–1948, 2021.
 - [30] N. Chouikhi, B. Ammar, N. Rokbani, and A. M. Alimi, “PSO-based analysis of Echo State Network parameters for time series forecasting,” *Applied Soft Computing*, vol. 55, pp. 211–225, 2017.
 - [31] Z.-X. Wang and Q. Li, “Modelling the nonlinear relationship between CO₂ emissions and economic growth using a PSO algorithm-based grey Verhulst model,” *Journal of Cleaner Production*, vol. 207, pp. 214–224, 2019.
 - [32] N. Zeng, H. Zhang, W. Liu, J. Liang, and F. E. Alsaadi, “A switching delayed PSO optimized extreme learning machine for short-term load forecasting,” *Neurocomputing*, vol. 240, pp. 175–182, 2017.
 - [33] S. Mao, Y. Kang, Y. Zhang, X. Xiao, and H. Zhu, “Fractional grey model based on non-singular exponential kernel and its application in the prediction of electronic waste precious metal content,” *ISA Transactions*, vol. 107, pp. 12–26, 2020.
 - [34] Y.-C. Hu, “Forecasting tourism demand using fractional grey prediction models with Fourier series,” *Annals of Operations Research*, vol. 300, no. 2, pp. 467–491, 2020.
 - [35] H. Zhao and W. Lifeng, “Forecasting the non-renewable energy consumption by an adjacent accumulation grey model,” *Journal of Cleaner Production*, vol. 275, Article ID 124113, 2020.
 - [36] X. Ma, W. Wu, B. Zeng, Y. Wang, and X. Wu, “The conformable fractional grey system model,” *ISA Transactions*, vol. 96, pp. 255–271, 2020.
 - [37] Y. Chen, W. Lifeng, L. Lianyi, and Z. Kai, “Fractional Hausdorff grey model and its properties,” *Chaos, Solitons & Fractals*, vol. 138, Article ID 109915, 2020.
 - [38] A. Ingel, N. Shahroudi, M. Kängsepp, A. Tättar, V. Komisarenko, and M. Kull, “Correlated daily time series and forecasting in the M4 competition,” *International Journal of Forecasting*, vol. 36, no. 1, pp. 121–128, 2020.
 - [39] F. Petropoulos, X. Wang, and S. M. Disney, “The inventory performance of forecasting methods: evidence from the M3 competition data,” *International Journal of Forecasting*, vol. 35, no. 1, pp. 251–265, 2019.