

## Research Article

# Parameter Optimization on the Three-Parameter Whitenization Grey Model and Its Application in Simulation and Prediction of Gross Enrollment Rate of Higher Education in China

Jihong Sun <sup>1</sup>, Hui Li <sup>2</sup>, Bo Zeng <sup>2</sup>, Xiaoyun Zhao <sup>3</sup> and Chuanhui Wang <sup>3</sup>

<sup>1</sup>Center of High Education, National Institution of Education Science, Beijing 100088, China

<sup>2</sup>College of Mathematics and Statistics, Chongqing Technology and Business University, Chongqing 400067, China

<sup>3</sup>College of Economics, Qufu Normal University, Rizhao 276800, China

Correspondence should be addressed to Xiaoyun Zhao; niethy@163.com

Received 5 November 2020; Revised 20 November 2020; Accepted 2 December 2020; Published 19 December 2020

Academic Editor: Baogui Xin

Copyright © 2020 Jihong Sun 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.

The gray prediction model, based on the GM(1,1) method, is an important branch of gray theory with the most active research and the most fruitful results, and it is the most widely used because of its small sample size, simple modeling process, and easy to use. Such advantages have been successfully applied in many fields such as transportation, agriculture, energy, medicine, and environment and have been gradually developed into a mainstream predictive modeling method. This study combines the Three-parameter Whitenization Grey Model (TWGM(1,1)), which fits the inhomogeneous exponential law sequence, and the Particle Swarm Algorithm (PSA) to optimize the order and background value coefficients under the condition of the minimum sum of squares of simulation errors, and hence, to solve the problem that the cumulative order is fixed to “1” and the background value coefficient value is fixed to “0.5.” As a result, a parameter-optimized gray system model with flexibility, adaptability, and dynamic adjustment is designed to simulate and predict China’s higher education gross enrollment rate. The application shows that the model has better overall simulation and prediction performance than others. On the one hand, the parametric optimization model significantly improves its own performance, and on the other hand, its intelligent and adjustable adaptivity improves the accuracy and further extends its application.

## 1. Introduction

Since Professor Ju-Long [1] established the gray system theory in 1982, the gray prediction model based on the GM(1,1) method is the most active, fruitful, and widely used branch of gray theory. With feature of the small sample size, simple modeling process, and low cost of learning, it has been successfully applied to many fields, such as transportation, agriculture, energy, medicine, and environment [2–6]. The model gradually developed into a mainstream predictive modeling method. The enrollment rate of higher education is influenced by many factors such as social, economic, cultural, and geographical factors, which are characterized by partly known and partly unknown information; thus, the GM(1,1) model, as a model to study the

uncertainty system problem of “partly known and partly unknown,” has been used to predict the enrollment rate of higher education. For example, Liu et al. [7] used the GM(1, 1) prediction model to make a mathematical analysis of the gross enrollment rate of higher education. It scientifically predicted and eased the employment situation of college students. It provided reference for promoting social stability and economic development. Dong [8] used the GM(1, 1) model to predict the fluctuation of China’s medium- and long-term gross enrollment ratio in higher education, which provides a reference for government assessment and citizens’ choice of higher education.

The GM(1, 1), as the first gray prediction model proposed to be applicable to univariate modeling, has shown favorable simulation and prediction performance for the

sequence of unity exponential growth characteristics. However, according to the Statistical Yearbook, the gross enrollment rate of higher education in China was 51.6% in 2019, which exceeded the 50% mark for the first time, indicating that China's higher education has moved from the mass stage to the generalized stage. The shift in development scale has undoubtedly increased the complexity and uncertainty of the higher education GER system, which makes the system behavior series more similar to nonuniform exponential growth characteristics, and the original system behavior series with similar unity exponential growth characteristics can no longer satisfy the existing system. How to develop a high-performance gray prediction model that is more flexible, adaptable, and suitable for modeling and predicting the gross enrollment ratio in higher education dynamically is the focus of this research.

In recent years, researchers have carried out a lot of research on the modeling ability of approximate nonhomogeneous exponential sequences. These studies can be divided into three categories according to different modeling ideas. The first is to optimize the structure of the GM(1,1) model to ensure that the final reduction formula of the model presents an approximate homogeneous exponential form. The NHGM(1, 1,  $k$ ) [9] model extends the basic form  $x^{(0)}(k) + az^{(1)}(k) = b$  of the GM(1,1) model to  $x^{(0)}(k) + az^{(1)}(k) = 0.5(2k - 1)b + c$ . The SIGM model [10] expands the basic form  $x^{(0)}(k) + az^{(1)}(k) = b$  of the GM(1,1) model to  $x^{(0)}(k) + az^{(1)}(k) = kb + c$ . The KRNGM model [11] introduces the nonlinear function  $f(t)$  into the whitening differential equation of the GM(1,1) model. The second is to modify the adaptability of modeling objects. The conversion from a nonhomogeneous exponential sequence to a homogeneous exponential sequence can be realized by the difference between adjacent elements in the original sequence. The third is the direct modeling method. Wang [12] proposed a direct modeling method of the GM(1,1) model generated from the original data without accumulation, and it had been gradually optimized [13, 14]. Ye and Li [15] established a whitening weight gray prediction model to measure the influence of the probability of interval gray number in the prediction results. Zeng and Liu [16] established a direct modeling method of DGM(1, 1) based on the original sequence by omitting the accumulation and subtraction processes. In addition, the models mentioned above have been applied in variate fields. Xiao et al. [17–20] studied parameter optimization of Grey Riccati model and its application in the prediction on the energy consumption and carbon. A new structure of the Gray Verhulst model is proposed by Xiao et al. [21–24], which improves the ability of the gray model to model saturated S-sequences, and it is applied in China's tight gas production forecasting.

The above optimization models around approximate nonaligned exponential sequences have better properties, modeling capabilities, and a wider range of applications than the GM(1,1) model. The parameters (such as order, background value coefficients, and initial values.) of the gray prediction model are crucial to the performance of the model. However, these performance parameters are often simplified to a specific value in the above optimization

model. Therefore, optimization of performance parameters is a key means to improve the stability, applicability, and flexibility of the model. A large number of studies have been launched for the optimization of performance parameters from many directions, such as initial value [25–27], order [28, 29], and background value [30, 31]. The optimal value of each performance parameter should meet the condition of the minimum sum of squared simulation errors of the model, and the optimization process requires a lot of calculations. The Particle Swarm Algorithm (PSA) provides an optional solution for the optimization of those parameters, which find the global optimum by following the currently searched optimum. A distinctive feature compared to other modern optimization methods is that the PSA requires few parameters to be adjusted, is simple and easy to implement, and converges quickly, which has become a research hotspot in the field of modern optimization methods [32–35]. Zeng and Liu [36] proposed the SAIGM\_FO model and used the PSA to optimize the order of the model. Wang and Li [37] used the PSA to optimize the structural parameters ( $a$ ,  $b$ ) of the model. It has been proved that the PSA in the gray prediction model improves the modeling accuracy and model flexibility.

In this study, it is firstly proposed that the PSA is applied in the Three-parameter Whitenization Grey Model (TWGM(1,1)) [38] to optimize the order and background value coefficients under the condition of minimum simulation error squared. The solution is not only capable to solve the problem that the cumulative order is fixed to "1" and the background value coefficient value is fixed to "0.5" but also displays a feature of flexible, adaptable, and dynamically adjustable in the application in modeling China's higher education gross enrollment ratio (GER). The empirical analysis shows that the model has better overall simulation and prediction performance than others. On the one hand, the parametric optimization model significantly improves its own performance, and on the other hand, its intelligent and adjustable adaptivity improves the accuracy and further extends its application. The rest of the study is organized as below. Section 2 describes the original Three-parameter Whitenization Grey Model. Then, the parameter optimization process of the Three-parameter Whitenization Grey Model is followed in Section 3, and Section 4 gives an application of the optimization model on China's higher education gross enrollment rate. The conclusion of the study is revealed in the last section.

## 2. Original TWGM(1,1) Model

According to the derivation process of the time-response function of the classic GM(1,1) model, the whitenization equation of the Three-parameter Whitenization Grey Model is established, and then the time-response function is derived by solving the differential equation. This model is called the Three-parameter Whitenization Grey Model, or TWGM (1,1) model in short.

*Definition 1.* Suppose sequence  $X^{(0)} = (x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n))$  is the original data sequence, where  $x^{(0)}(k) \geq 0$ ,

$k = 1, 2, \dots, n$ , and  $X^{(1)} = (x^{(1)}(1), x^{(1)}(2), \dots, x^{(1)}(n))$  is the 1-AGO sequence of  $X^{(0)}$ , where

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

$Z^{(1)} = (z^{(1)}(2), z^{(1)}(3), \dots, z^{(1)}(n))$  is the immediate mean-generating sequence of  $X^{(1)}$ , where

$$z^{(1)}(k) = 0.5 \times [x^{(1)}(k) + x^{(1)}(k-1)], \quad k = 2, 3, \dots, n. \quad (2)$$

*Definition 2.* Assuming that the sequences  $X^{(0)}$  and  $X^{(1)}$  are as shown in Definition 1, then

$$\frac{dx^{(1)}}{dt} + ax^{(1)} = bt + c. \quad (3)$$

Equation (3) is the whitening differential equation of the TWGM(1,1) model. The equation

$$x^{(0)}(k) + az^{(1)}(k) = 0.5(2k-1)b + c, \quad (4)$$

is the basic form of the TWGM(1,1) model.

According to the modeling ideas of the GM(1,1), the model parameters  $a$ ,  $b$ , and  $c$  are estimated through the basic form of the TWGM(1,1) model. The time-response function of the TWGM(1,1) model is obtained by solving the differential equation.

**Theorem 1.** Suppose the sequence  $X^{(0)}$ ,  $X^{(1)}$ , and  $Z^{(1)}$  are as shown in Definition 1,  $\hat{p} = (a, b, c)^T$  is the parameter list, and

$$Y = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(n) \end{bmatrix}, \quad (5)$$

$$B = \begin{bmatrix} -z^{(1)}(2) & \frac{3}{2} & 1 \\ -z^{(1)}(3) & \frac{5}{2} & 1 \\ \vdots & \vdots & \vdots \\ -z^{(1)}(n) & \frac{(2n-1)}{2} & 1 \end{bmatrix}.$$

Then, the least-square estimation parameter list of the TWGM(1,1) model is satisfied.

**Theorem 2.** The time-response function of the TWGM(1,1) model is

$$x^{(1)}(t) = \left( x^{(1)}(1) - \frac{b}{a} + \frac{b}{a^2} - \frac{c}{a} \right) e^{-a(t-1)} + \frac{b}{a}t - \frac{b}{a^2} + \frac{c}{a}. \quad (6)$$

The inverse-accumulating reduction formula of the TWGM(1,1) model is

$$\hat{x}^{(0)}(t) = (1 - e^a) \left( x^{(0)}(1) - \frac{b}{a} + \frac{b}{a^2} - \frac{c}{a} \right) e^{-a(t-1)} + \frac{b}{a}. \quad (7)$$

### 3. Parameter Optimization of the Three-Parameter Whitening Gray Model

*3.1. Background Value Optimization.* In the TWGM(1,1) model modeling process, the sequence generated next to the mean is used as the background value, which is a common smoothing method to weaken the influence of extreme values or outliers in the 1-AGO sequence on the magnitude of the gray effect. Among them, the background value coefficient is the weight of adjacent elements in the process of constructing the adjacent mean sequence. The difference of its value will affect the calculation result of the adjacent mean value series and then affect the simulation and prediction effects of the model. In the above modeling process, the background value coefficient is set to 0.5. This processing method lacks flexibility and cannot guarantee that the model achieves the best simulation effect. So, it is necessary to optimize the background value of the original model. The optimization of the background value is mainly based on the optimization of the background value coefficient. Optimize the definition of the sequence immediately adjacent to the mean.

*Definition 3.* Suppose that  $X^{(1)}$  is as shown in Definition 1,  $Z_w^{(1)} = (z_w^{(1)}(2), z_w^{(1)}(3), \dots, z_w^{(1)}(n))$  is the optimized sequence generated by the immediate mean of  $X^{(1)}$ , and  $w$  is the background value coefficient,  $w \in (0, 1)$ , where

$$z_w^{(1)}(k) = wx^{(1)}(k) + (1-w)x^{(1)}(k-1), \quad k = 2, 3, \dots, n. \quad (8)$$

Then,

$$x^{(0)}(k) + a(wx^{(1)}(k) + (1-w)x^{(1)}(k-1)) = kb + c, \quad (9)$$

is the TWGM(1,1) model with the background value coefficient  $w$  ( $0 < w < 1$ ).

Among them, the optimal value of the background value coefficient  $w$  should satisfy the minimum sum of squared simulation errors of the model, namely,

$$\min f(w) = \frac{1}{n-1} \sum_{k=2}^n [x^{(0)}(k) - \hat{x}^{(0)}(k)]^2, \quad 0 < w < 1. \quad (10)$$

In equation (10),  $x^{(0)}(k)$  is the original modeling data and  $\hat{x}^{(0)}(k)$  is the simulated data of the model.

**3.2. Order Optimization.** In the original modeling process, the cumulative order is fixed at “1,” which leads to poor flexibility and adaptability of the model. Therefore, Wu et al. [39, 40] introduced fractional order into gray modeling based on the “in between” idea. This realizes the expansion of the cumulative order of the gray prediction model from integer to fraction. Meng et al. [38, 41] used the Gamma function to give the functional expression of the fractional accumulation operator, which provided the basis for constructing the fractional gray prediction model. This section optimizes the order of the TWGM(1,1) model in order to improve model performance and model adaptability.

**Definition 4.** Suppose  $X^{(0)}$  is as shown in Definition 1 and  $X^{(r)}$  ( $r \in \mathbb{R}^+$ ) is a new sequence; then, sequence  $X^{(r)} = (x^{(r)}(1), x^{(r)}(2), \dots, x^{(r)}(n))$  is called the  $r$ -order cumulative generating sequence of  $X^{(0)}$ , where

$$x^{(r)}(k) = \sum_{i=1}^k \frac{\Gamma(r+k-i)}{\Gamma(k-i+1)\Gamma(r)} x^{(0)}(i), \quad k = 1, 2, \dots, n. \quad (11)$$

Then, the sequence  $Z^{(r)} = (z^{(r)}(2), z^{(r)}(3), \dots, z^{(r)}(n))$  is called the adjacent mean-value generating sequence of  $X^{(r)}$ , where

$$z^{(r)}(k) = \frac{x^{(r)}(k) + x^{(r)}(k-1)}{2}, \quad k = 2, 3, \dots, n. \quad (12)$$

Then,

$$x^{(r-1)}(k) + az^{(r)}(k) = 0.5(2k-1)b + c, \quad (13)$$

is the TWGM(1,1) model of  $r$  order.

Among them, the optimal value of the  $r$  order should satisfy the minimum sum of squared simulation errors of the model, namely,

$$\min f(r) = \frac{1}{n-1} \sum_{k=2}^n [x^{(0)}(k) - \hat{x}^{(0)}(k)]^2, \quad r \in \mathbb{R}^+. \quad (14)$$

**3.3. Simultaneous Optimization of Background Value and Order.** Based on the above optimization analysis, this section derives and constructs a new TWGM(1,1) model under the condition of simultaneous optimization of background value and order. The new model is called the TWGM(1, 1| $w, r$ ) model.

### 3.3.1. Definition

**Definition 5.** Suppose the sequences  $X^{(0)}$  and  $X^{(r)}$  are as shown in Definitions 1 and 4; then,

$$\frac{dx^{(r)}}{dt} + ax^{(r)} = bt + c, \quad (15)$$

is called the whitening differential equation of the TWGM(1, 1| $w, r$ ) model.

The sequence  $Z_w^{(r)} = (z_w^{(r)}(2), z_w^{(r)}(3), \dots, z_w^{(r)}(n))$  is called the optimized sequence generated by the adjacent mean value of  $X^{(r)}$  under the condition of the background value coefficient  $w$ , where

$$z_w^{(r)}(k) = wx^{(r)}(k) + (1-w)x^{(r)}(k-1), \quad k = 2, 3, \dots, n. \quad (16)$$

Integrating both ends of formula (15) in interval  $[k-1, k]$ , we can obtain

$$\int_{k-1}^k dx^{(r)}(t) + a \int_{k-1}^k x^{(r)}(t)dt = \int_{k-1}^k bt dt + \int_{k-1}^k c dt. \quad (17)$$

Because

$$\begin{aligned} \int_{k-1}^k dx^{(r)}(t) &= x^{(r)}(k) - x^{(r)}(k-1) = x^{(r-1)}(k), \\ \int_{k-1}^k bt dt &= 0.5(2k-1)b, \quad \int_{k-1}^k c dt = c, \end{aligned} \quad (18)$$

among them, the size of  $\int_{k-1}^k x^{(r)}(t)dt$  can be approximately replaced by the area represented by  $z_w^{(r)}(k)$ , namely,

$$\int_{k-1}^k x^{(r)}(t)dt \approx wx^{(r)}(k) + (1-w)x^{(r)}(k-1) = z_w^{(r)}(k). \quad (19)$$

So, equation (15) can be converted to

$$x^{(r-1)}(k) + z_w^{(r)}(k) = 0.5(2k-1)b + c. \quad (20)$$

Equation (20) is the basic form of the TWGM(1, 1| $w, r$ ) model.

### 3.3.2. Parameter Estimation

**Theorem 3.** Suppose the sequences  $X^{(0)}$ ,  $X^{(r)}$ , and  $Z_w^{(r)}$  are as shown in Definitions 1, 4, and 5,  $\hat{p} = (a, b, c)^T$  is the parameter list, and

$$Y = \begin{bmatrix} x^{(r-1)}(2) \\ x^{(r-1)}(3) \\ \vdots \\ x^{(r-1)}(n) \end{bmatrix} = \begin{bmatrix} x^{(r)}(2) - x^{(r)}(1) \\ x^{(r)}(3) - x^{(r)}(2) \\ \vdots \\ x^{(r)}(n) - x^{(r)}(n-1) \end{bmatrix},$$

$$B = \begin{bmatrix} -z_w^{(r)}(2) & \frac{3}{2} & 1 \\ -z_w^{(r)}(3) & \frac{5}{2} & 1 \\ \vdots & \vdots & \vdots \\ -z_w^{(r)}(n) & \frac{(2n-1)}{2} & 1 \end{bmatrix} = \begin{bmatrix} -wx^{(r)}(2) - (1-w)x^{(r)}(1) & \frac{3}{2} & 1 \\ -wx^{(r)}(3) - (1-w)x^{(r)}(2) & \frac{5}{2} & 1 \\ \vdots & \vdots & \vdots \\ -wx^{(r)}(n) - (1-w)x^{(r)}(n-1) & \frac{(2n-1)}{2} & 1 \end{bmatrix}. \quad (21)$$

Then, the least-square estimation parameter list of the TWGM(1,1| $w,r$ ) model satisfies  $\hat{p} = (a, b, c)^T = (B^T B)^{-1} B^T Y$ .

**3.3.3. Time-Response Derivation.** According to the whitening differential equation of the TWGM(1,1| $w,r$ ) model to derive its time-response function and according to formula (15), the corresponding homogeneous equation is

$$\frac{dx^{(r)}(t)}{dt} + ax^{(r)}(t) = 0 \Rightarrow \frac{dx^{(r)}(t)}{dt} = -ax^{(r)}(t). \quad (22)$$

Then,

$$\ln|x^{(r)}| = -at + \ln|C_1|. \quad (23)$$

The general solution of homogeneous equation (22) is

$$x^{(r)}(t) = C_1 e^{-at}. \quad (24)$$

Using the constant variation method, replace formula (24)  $C_1$  with  $u(t)$ , and let

$$x^{(r)}(t) = u(t)e^{-at}. \quad (25)$$

Derivation from both ends of equation (25) with respect to  $t$ :

$$\frac{dx^{(r)}(t)}{dt} = u'(t)e^{-at} - au(t)e^{-at}. \quad (26)$$

Substituting equation (26) into formula (15), we obtain

$$u'(t)e^{-at} - au(t)e^{-at} = bt + c - ax^{(r)}. \quad (27)$$

According to equation (27), there is  $u'(t) = (bt + c)e^{at}$ , so

$$u(t) = \int (bt + c)e^{at} dt = b \int te^{at} dt + c \int e^{at} dt = \frac{b}{a}te^{at} - \frac{b}{a^2}e^{at} + \frac{c}{a}e^{at} + C. \quad (28)$$

Substituting equation (28) into (25), we obtain

$$x^{(r)}(t) = \left( \frac{b}{a}te^{at} - \frac{b}{a^2}e^{at} + \frac{c}{a}e^{at} + C \right) e^{-at}. \quad (29)$$

Calculating the above equation,

$$x^{(r)}(t) = \frac{b}{a}t - \frac{b}{a^2} + \frac{c}{a} + Ce^{-at}. \quad (30)$$

Among them,  $x^{(r)}(1) = x^{(0)}(1)$  is a known item. At that time  $t = 1$ , equation (30) can be obtained:



$$x^{(r)}(1) = Ce^{-a} + \frac{b}{a} - \frac{b}{a^2} + \frac{c}{a}. \quad (31)$$

We obtain

$$C = \frac{x^{(r)}(1) - (b/a) + (b/a^2) - (c/a)}{e^{-a}}. \quad (32)$$

Substituting equation (32) into (30), we obtain

$$x^{(r)}(t) = \left( x^{(r)}(1) - \frac{b}{a} + \frac{b}{a^2} - \frac{c}{a} \right) e^{-a(t-1)} + \frac{b}{a}t - \frac{b}{a^2} + \frac{c}{a}. \quad (33)$$

Among them,  $t = 2, 3, \dots, n$  and equation (33) is the time-response formula. At this point, the time-response derivation is over.

### 3.3.4. Derivation of Accumulative Reduction

*Definition 6.* Suppose sequence  $X^{(0)}$  is as shown in Definition 1 and if  $r \in \mathbf{R}^+$ , then sequence  $X^{(-r)} = (x^{(-r)}(1), x^{(-r)}(2), \dots, x^{(-r)}(n))$  is called the  $r$ -order accumulative generating sequence of  $X^{(0)}$ , where

$$x^{(-r)}(k) = \sum_{i=0}^{k-1} (-1)^i \frac{\Gamma(r+1)}{\Gamma(i+1)\Gamma(r-i+1)} x^{(0)}(k-i), \quad k = 1, 2, \dots, n. \quad (34)$$

**Theorem 4.** Suppose sequence  $X^{(0)}$  is as defined in Definition 1,  $p \in \mathbf{R}^+$ ,  $q \in \mathbf{R}^+$ ,  $X^{(p)}$  is the  $p$ th order cumulative generation sequence of  $X^{(0)}$ ,  $X^{(q)}$  is the  $q$ th order cumulative generation sequence of  $X^{(0)}$ ,  $X^{(p+q)}$  is the  $p+q$ -order cumulative generation sequence of  $X^{(0)}$ ,  $(X^{(p)})^{(q)}$  is the  $q$ -order cumulative generation sequence of  $X^{(p)}$ , and  $(X^{(q)})^{(p)}$  is the  $p$ -order cumulative generating sequence of  $X^{(q)}$ ; then, the multiple cumulative generating operator satisfies the commutative law and exponential rate, namely,

$$(X^{(p)})^{(q)} = (X^{(q)})^{(p)} = X^{(p+q)}. \quad (35)$$

**Corollary 1.** According to Theorem 1, it can be derived as

$$X^{(0)} = (X^{(r)})^{(-r)} = (X^{(-r)})^{(r)}. \quad (36)$$

According to Corollary 1, which is equation (36), the final reduction formula of the TWGM(1, 1| $w, r$ ) model can be derived as

$$\hat{x}^{(0)}(k) = (\hat{x}^{(r)}(k))^{(-r)} = \sum_{i=0}^{k-1} (-1)^i \frac{\Gamma(r+1)}{\Gamma(i+1)\Gamma(r-i+1)} \hat{x}^{(r)}(k-i). \quad (37)$$

In equation (37), when  $k = 2, 3, \dots, n$ ,  $\hat{x}^{(0)}(k)$  is called the simulated value and when  $k = n+1, n+2, \dots$ ,  $\hat{x}^{(0)}(k)$  is called the predicted value.

**3.4. TWGM(1,1) Model Parameter Optimization Based on PSO Algorithm.** In the TWGM(1, 1| $w, r$ ) model, there are two undetermined parameters (the background value coefficient and order). In order to obtain the best performance of the TWGM(1, 1| $w, r$ ) model, the optimal values of these two undetermined parameters are required. The combined solution of the two parameters is easy to cause errors due to mutual influence and mutual interference of the parameters, and at the same time, it weakens the optimization effect of the independent parameter on the model. Therefore, in the process of solving the optimal value of the parameter, a step-by-step method is adopted. Specifically, first, in  $0 < w < 1$ , assuming that the order does not change, that is,  $r = 1$ , the optimal background value coefficient is solved. Secondly, on the basis of  $w = w^*$ , in  $r \in \mathbf{R}^+$ , the optimal order is solved. According to Definitions 3 and 4, the optimal background value coefficient  $w$  and the optimal order  $r$  are found within their respective value ranges and should be obtained under

the condition that the sum of squared simulation errors of the model is minimized, that is, to satisfy equations (10) and (14).

Obviously, the optimization process of each parameter to be determined takes a lot of time and takes up limited computer resources. Various group optimization algorithms (such as Particle Swarm Optimization and ant colony algorithm.) provide good solutions to complex distributed optimization problems. The Particle Swarm Optimization (PSO) is a swarm-intelligent global optimization algorithm that simulates the predation behavior of birds. Its basic concept comes from the study of the foraging behavior of birds. The algorithm has the advantages of simple structure, few parameters, and easy programming. At the same time, the Particle Swarm Optimization algorithm based on adaptive mutation of the population fitness variance effectively solves the phenomenon of premature convergence and can significantly improve the global convergence performance. It has been used in function optimization and neural network. It is widely used in training, engineering, and other fields [28–31]. Therefore, this paper uses PSO to optimize the background value coefficient and order step by step. The PSO algorithm solving steps are as follows:

*Step 1.* Initialize randomly the position and velocity of the particles in the particle swarm.

*Step 2.* Set the particle in the current position to the position of the best particle in the initial population.

*Step 3.* Calculate the average relative simulation error of the TWGM(1,1| $w, r$ ) model when  $w = pBest$  (or  $r = pBest$ ):

*Step 3.1.* Calculate the cumulative sequence  $X^{(r)}$  of order  $r$ .

*Step 3.2.* Calculate the sequence  $Z_w^{(r)}$  immediately adjacent to the background value coefficient  $w$ .

*Step 3.3.* Construct matrix B and Y, and solve model parameter  $\hat{p} = (a, b, c)^T$ .

*Step 3.4.* Calculate the simulation value  $\hat{X}^{(0)}$ .

*Step 3.5.* Calculate the average relative simulation error  $f(pBest)$  of  $\hat{X}^{(0)}$ .

*Step 3.6.* Judge whether  $|f(pBest) - f(gBest)|$  is less than the given convergence value  $\delta$ ; if it is satisfied, go to *Step 9*. Otherwise, execute *Step 4*.

*Step 4.* Perform the following operations for all particles in the particle swarm:

*Step 4.1.* Update particle position and velocity:

$$\begin{aligned} V &= \omega \times V + c_1 \times \text{rand} \times (pBest - \text{Present}) \\ &\quad + c_2 \times \text{rand} \times (gBest - \text{Present}), \\ \text{Present} &= \text{Present} + V, \end{aligned} \quad (38)$$

$$\omega = \omega_{\max} - \text{run} \times \frac{(\omega_{\max} - \omega_{\min})}{\text{runMax}}.$$

*Step 4.2.* If the particle fitness is better than the fitness, set it to the new position.

*Step 4.3.* If the particle fitness is better than the fitness, set it to the new position.

*Step 5.* Calculate the variance of the group fitness  $\sigma^2$ , and calculate  $f(pBest)$ :

$$\begin{aligned} \sigma^2 &= \sum_{i=1}^n \left( \frac{f_i - f_{\text{avg}}}{f} \right)^2, \\ f &= \begin{cases} \max\{|f_i - f_{\text{avg}}|\}, & \max\{|f_i - f_{\text{avg}}|\} > 1, \\ 1, & \text{others.} \end{cases} \end{aligned} \quad (39)$$

*Step 6.* Calculate the probability of mutation  $p_m$ .

*Step 7.* Generate random number  $\varepsilon \in [0, 1]$ ; if  $\varepsilon < p_m$ , press to execute mutation operation. Otherwise, go to *Step 8*.

*Step 8.* Determine whether the algorithm convergence criterion is satisfied; if it is satisfied, execute *Step 9*. Otherwise, turn to *Step 3*.

*Step 9.* Output the optimal value  $gBest$  of the background value coefficient  $w$  (or order  $r$ ) and the

simulation and prediction data of the TWGM(1,1| $w, r$ ) model at this time, and the algorithm operation ends.

#### 4. Model Application: Forecast of China's Higher Education Gross Enrollment Rate

To identify the optimized performance on each parameter, the variables (the background value coefficient  $w$  and order  $r$ ) are simulated with four different groups. The original TWGM(1,1) model means that the model is not optimized for order and background value coefficients at all. The first group is the model with,  $w=0.5$  and  $r=1$ , that is, the TWGM(1,1|0.5,1) model. The second one is the model with  $w=0.5$  and  $r=r^*$ , that is, the TWGM(1,1|0.5, $r^*$ ) model, which indicates that the model only optimizes the order. The third one is the model with,  $w=w^*$  and  $r=1$ , that is, the TWGM(1,1| $w^*$ , 1) model, which means that the model only optimizes the background value coefficient. The last group is  $w=w^*$  and  $r=r^*$ , that is, the TWGM(1,1| $w^*$ ,  $r^*$ ) model, which suggests that the model optimizes the order and background value coefficients at the same time. The simulation results with Matlab for four parameter groups are shown in Table 1, where  $x^{(0)}(k)$  is China's high education enrollment data from 1991 to 2019.

In order to intuitively reflect the simulation effects of different models, a comparison chart of the simulation curves of different models is drawn, as shown in Figure 1, where models with different parameter values show different simulation effects. The curve of the TWGM(1,1|0.5, 1) and TWGM(1,1| $w^*$ , 1) models overlap and show a same trend. The TWGM(1,1| $w^*$ ,  $r^*$ ) model initially displays better simulating and predicting effects for the gross enrollment ratio in higher education in China since it produces the most close result with raw data.

In order to measure the optimization performance of different parameters, the average relative simulation percentage error  $\bar{\Delta}_S$ , average relative prediction percentage error  $\bar{\Delta}_F$ , and comprehensive percentage error  $\bar{\Delta}$  of the gray model of each model are calculated according to the results in Table 1. The results are shown in Table 2. The relevant calculation equation is as follows:

$$\begin{aligned} \bar{\Delta}_S &= \frac{1}{t-1} \sum_{k=2}^t \Delta k, \\ \bar{\Delta}_F &= \frac{1}{n-t} \sum_{k=t+1}^n \Delta k, \\ \bar{\Delta} &= \frac{1}{n-1} \sum_{k=2}^n \Delta k. \end{aligned} \quad (40)$$

Among them,  $\Delta k$  represents the simulation/prediction error of  $\hat{x}^{(0)}(k)$ ,  $n$  represents the number of experimental data, and  $t$  represents the number of experimental data used for simulation.

Based on the results of the calculations in Table 2, the relative percentage error comparison figure with different parameter values is plotted, as shown in Figure 2 in

TABLE 1: Simulation and prediction results of different models on China's higher education gross enrollment rate.

Serial number	$x^{(0)}(k)$	TWGM(1,1 0.5,1)		TWGM(1,1 0.5, $r^*$ )		TWGM(1,1  $w^*$ ,1)		TWGM(1,1  $w^*$ , $r^*$ )	
		$\hat{x}^{(0)}(k)$	$\Delta(k)$	$\hat{x}^{(0)}(k)$	$\Delta(k)$	$\hat{x}^{(0)}(k)$	$\Delta(k)$	$\hat{x}^{(0)}(k)$	$\Delta(k)$
Simulation data									
$k=2$	3.5	3.14	10.2188	3.13	10.6725	3.15	10.1019	3.10	11.5502
$k=3$	3.9	4.04	3.5078	4.06	4.1620	4.04	3.6783	4.05	3.7288
$k=4$	5	4.97	0.6278	5.00	0.0311	4.98	0.4414	4.98	0.4818
$k=5$	6	5.94	1.0113	5.96	0.6118	5.95	0.8097	5.93	1.1729
$k=6$	7.2	6.95	3.4645	6.96	3.2806	6.97	3.2561	6.92	3.9439
$k=7$	8.3	8.00	3.5661	8.00	3.5641	8.02	3.3488	7.94	4.3312
$k=8$	9.1	9.10	0.0159	9.09	0.1324	9.12	0.2488	9.01	1.0183
$k=9$	9.8	10.24	4.5378	10.22	4.2692	10.27	4.7877	10.12	3.2693
$k=10$	10.5	11.44	8.9113	11.40	8.5588	11.46	9.1774	11.28	7.4608
$k=11$	12.5	12.68	1.4111	12.63	1.0562	12.71	1.6634	12.50	0.0000
$k=12$	13.3	13.97	5.0293	13.92	4.6743	14.00	5.2945	13.77	3.5632
$k=13$	15	15.32	2.1024	15.27	1.8063	15.35	2.3637	15.11	0.7258
$k=14$	17	16.72	1.6586	16.68	1.8634	16.76	1.4040	16.51	2.8898
$k=15$	19	18.18	4.3196	18.16	4.4102	18.23	4.0694	17.98	5.3818
$k=16$	21	19.70	6.1832	19.71	6.1376	19.75	5.9357	19.52	7.0515
$k=17$	22	21.29	3.2394	21.33	3.0264	21.34	2.9821	21.14	3.9173
$k=18$	23	22.94	0.2639	23.04	0.1533	23.00	0.0032	22.84	0.7002
$k=19$	23.3	24.66	5.8381	24.82	6.5170	24.73	6.1234	24.63	5.6917
Forecast data									
$k=20$	24.2	26.45	9.3103	26.69	10.2816	26.52	9.6068	26.50	9.5236
$k=21$	26.5	28.32	6.8708	28.65	8.1088	28.40	7.1623	28.48	7.4706
$k=22$	26.9	30.27	12.5146	30.71	14.1460	30.35	12.8231	30.56	13.5932
$k=23$	30	32.29	7.6443	32.86	9.5415	32.38	7.9409	32.74	9.1365
$k=24$	34.5	34.40	0.2759	35.13	1.8131	34.50	0.0000	35.04	1.5616
$k=25$	37.5	36.60	2.3881	37.50	0.0000	36.71	2.1169	37.46	0.1167
$k=26$	40	38.90	2.7601	39.99	0.0210	39.00	2.4889	40.00	0.0000
$k=27$	42.7	41.28	3.3183	42.61	0.2195	41.40	3.0476	42.68	0.0541
$k=28$	45.7	43.77	4.2234	45.35	0.7650	43.89	3.9543	45.49	0.4502
$k=29$	48.1	46.36	3.6163	48.23	0.2714	46.49	3.3446	48.46	0.7477
$k=30$	51.6	49.06	4.9237	51.25	0.6712	49.20	4.6548	51.58	0.0365

TABLE 2: Error comparison of different models.

Error type	TWGM(1,1 0.5,1) (%)	TWGM(1,1 0.5, $r^*$ ) (%)	TWGM(1,1  $w^*$ ,1) (%)	TWGM(1,1  $w^*$ , $r^*$ ) (%)
Average relative simulation percentage error $\bar{\Delta}_s$	3.6615	3.6071	3.6494	3.7155
Average relative prediction percentage error $\bar{\Delta}_F$	5.2587	4.1672	5.1946	3.8810
Grey model comprehensive percentage error $\bar{\Delta}$	4.2673	3.8195	4.2355	3.7782

supplementary materials. From Table 2 and Figure 2 in supplementary materials, the following conclusions are drawn:

- (1) The background value optimization effect: comparing the TWGM(1,1|0.5,1) model and TWGM(1,1| $w^*$ ,1) model with a  $w^* = 0.4538$ , the graph shows that the integrated percent error of the gray model after the background value optimization is slightly lower than the original model, but the difference is not remarkable, which suggests that the background value optimization has an optimization effect on the comprehensive performance of China's

higher education gross enrollment rate, but the effect is not significant.

- (2) The order optimization effect: compare the TWGM(1,1|0.5,1) model with the TWGM(1,1|0.5, $r^*$ ) model with a  $r^* = 1.1127$  in Figure 2 in supplementary materials, the average relative percentage prediction error and the comprehensive percentage error after the order optimization are significantly lower than the original model, which shows that the order optimization has a remarkable improvement on the model performance in simulating China's higher education gross enrollment rate.



- (3) The TWGM(1,1| $w^*$ ,  $r^*$ ) model possesses the lowest error of 3.7782%, which indicates that the optimization of both the background value and the order is of the best option for improvement on the model performance. Thus, the TWGM(1,1| $w^*$ ,  $r^*$ ) model provides the best fitting effect and is preferred for simulation and prediction.

## 5. Conclusions

The study proposed a new optimization model by combining the Three-parameter Whitenization Grey Model with the PSA. To demonstrate the performance of this new model, an application analysis is carried out in the predication of China's high education enrollment ratio with Matlab simulation. The findings suggest that the new optimization model proposed in this study improves the performance of the original gray model remarkably. Specifically, compared with the original gray model, the average relative percentage prediction error and comprehensive percentage error after the optimization are significantly reduced, which indicates that the simultaneous optimization of the background value and order has a notable improvement on the overall performance in simulating and predicting China's higher education gross enrollment ratio. Besides, it is indicated that the intelligent and adjustable adaptivity improves the accuracy of optimization and, furthermore, extends its application in the empirical study in the future.

## Data Availability

The data (High Education Gross Enrollment Ratio of China) can be found in the supplementary information files or in "The Education Statistic Yearbook of China," Chapter 1, Page 21: Gross Enrollment Ratio of Education by Level (<https://data.cnki.net/trade/Yearbook/Single/N2020070382?z=Z017>).

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

The study is supported by Fundamental Research Funds for the Central Scientific Research Institutes of NIES, "Research on the evaluation of first-class undergraduate education," No. GYB2019002.

## Supplementary Materials

Figure 1: Comparison of simulation curves of different models. Figure 2: Comparison of relative percentage errors of different models. (*Supplementary Materials*)

## References

- [1] D. Ju-Long, "Control problems of grey systems," *Systems & Control Letters*, vol. 1, no. 5, pp. 288–294, 1982.
- [2] N. Xu, S. Ding, Y. Gong, and J. Bai, "Forecasting Chinese greenhouse gas emissions from energy consumption using a novel grey rolling model," *Energy*, vol. 175, pp. 218–227, 2019.
- [3] N.-T. Nguyen and T.-T. Tran, "Optimizing mathematical parameters of Grey system theory: an empirical forecasting case of Vietnamese tourism," *Neural Computing and Applications*, vol. 31, no. 2, pp. 1075–1089, 2019.
- [4] B. Zeng, S. Liu, B. Yun, and M. Zhou, "Research on early prediction and warning of human diseases based on grey system modeling technology," *Chinese Management Science*, vol. 28, no. 1, pp. 144–152, 2020.
- [5] Z. X. Wang, Z. W. Wang, and Q. Li, "Forecasting the industrial solar energy consumption using a novel seasonal GM(1,1) model with dynamic seasonal adjustment factors," *Energy*, vol. 200, 2020.
- [6] H. M. Duan, X. P. Xiao, J. Long et al., "Tensor alternating least squares grey model and its application to short-term traffic flows," *Applied Soft Computing Journal*, vol. 89, 2020.
- [7] L. Liu, Y. Feng, and C. Yang, "My country's higher education gross enrollment rate GM(1,1) prediction model and its application research," *Journal of Yunnan Normal University (Philosophy and Social Sciences Edition)*, vol. 41, no. 3, pp. 123–127, 2009.
- [8] X. Dong, "Research on the grey prediction model of national medium and long-term higher education gross enrollment rate," *Journal of National Academy of Education Administration*, vol. 5, pp. 43–47, 2010.
- [9] L. Zhan and H. Shi, "Grey modeling method and model for approximate inhomogeneous index data," *System Engineering Theory and Practice*, vol. 33, no. 3, pp. 689–694, 2013.
- [10] B. Zeng and C. Li, "Forecasting the natural gas demand in China using a self-adapting intelligent grey model," *Energy*, vol. 112, pp. 810–825, 2016.
- [11] X. Ma, Y.-S. Hu, and Z.-B. Liu, "A novel kernel regularized nonhomogeneous grey model and its applications," *Communications in Nonlinear Science and Numerical Simulation*, vol. 48, pp. 51–62, 2017.
- [12] Y. Wang, "GM(1,1) Gradually optimizing the promotion of direct modeling method," *System Engineering Theory and Practice*, vol. 02, pp. 120–124, 2003.
- [13] Y. Wang, G. Liu, and K. Liu, "A gradual optimization and direct modeling method of GM(1,1)," *Systems Engineering Theory and Practice*, vol. 09, pp. 99–104+144, 2000.
- [14] Y. Wang, "The direct modeling method and properties of GM(1,1)," *System Engineering Theory and Practice*, vol. 1, pp. 27–31, 1988.
- [15] L. Ye and J. Li, "Research on interval grey number prediction model considering whitening weight function," *Fuzzy Systems and Mathematics*, vol. 33, no. 4, pp. 165–174, 2020.
- [16] B. Zeng and S. Liu, "DGM(1,1) direct modeling method of approximate non-homogeneous exponential sequence," *System Engineering Theory and Practice*, vol. 31, no. 2, pp. 297–301, 2011.
- [17] Q. Xiao, M. Gao, X. Xiao et al., "A novel grey Riccati-Bernoulli model and its application for the clean energy consumption prediction," *Engineering Applications of Artificial Intelligence*, vol. 95, Article ID 103863, 2020.
- [18] Q. Xiao, M. Shan, M. Gao, X. Xiao, and M. Goh, "Parameter optimization for nonlinear grey Bernoulli model on biomass energy consumption prediction," *Applied Soft Computing*, vol. 95, Article ID 106538, 2020.
- [19] M. Gao, H. Yang, Q. Xiao et al., "A novel fractional grey Riccati model for carbon emission prediction," *Journal of Cleaner Production*, vol. 10, Article ID 124471, 2020.

- [20] Q. Xiao, M. Shan, M. Gao et al., "Grey information coverage interaction relational decision making and its application," *Journal of Systems Engineering and Electronics*, vol. 31, no. 2, pp. 359–369, 2020.
- [21] B. Xiao, H. Li, and X. Ma, "A novel multi-variable grey forecasting model and its application in forecasting the grain production in China," *Computers & Industrial Engineering*, vol. 150, Article ID 106915, 2020.
- [22] B. Zeng, X. Ma, and M. Zhou, "A new-structure grey Verhulst model for China's tight gas production forecasting," *Applied Soft Computing*, vol. 96, Article ID 106600, 2020.
- [23] B. Zeng, M. Zhou, X. Liu et al., "Application of a new grey prediction model and grey average weakening buffer operator to forecast China's shale gas output," *Energy Reports*, vol. 6, pp. 1608–1618, 2020.
- [24] B. Zhang, M. Tong, and X. Ma, "A new-structure grey Verhulst model: development and performance comparison," *Applied Mathematical Modelling*, vol. 81, pp. 522–537, 2020.
- [25] J. J. Wang, Y. G. Dang, J. Ye et al., "An improved grey prediction model based on matrix representations of the optimized initial value," *Journal of Grey System*, vol. 30, no. 3, pp. 143–156, 2018.
- [26] S. Ding, K. W. Hipel, and Y.-G. Dang, "Forecasting China's electricity consumption using a new grey prediction model," *Energy*, vol. 149, pp. 314–328, 2018.
- [27] S. Ding, R. Li, and Y. Dang, "GM(1,1) power model based on optimization of initial conditions and its application," *China Management Science*, vol. 28, no. 1, pp. 153–161, 2020.
- [28] S. H. Mao, M. Y. Gao, X. P. Xiao et al., "A novel fractional grey system model and its application," *Applied Mathematical Modelling*, vol. 40, no. 7-8, pp. 5063–5076, 2016.
- [29] 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.
- [30] X. Zeng, S. Yan, F. He, and Y. Shi, "Multi-variable grey model based on dynamic background algorithm for forecasting the interval sequence," *Applied Mathematical Modelling*, vol. 80, pp. 99–114, 2020.
- [31] W. Zhang, R. Xiao, B. Shi et al., "Forecasting slope deformation field using correlated grey model updated with time correction factor and background value optimization," *Engineering Geology*, vol. 260, 2019.
- [32] K. Li, L. Liu, J. Zhai, T. M. Khoshgoftaar, and T. Li, "The improved grey model based on particle swarm optimization algorithm for time series prediction," *Engineering Applications of Artificial Intelligence*, vol. 55, pp. 285–291, 2016.
- [33] N. Xu, Y. Dang, and Y. Gong, "Novel grey prediction model with nonlinear optimized time response method for forecasting of electricity consumption in China," *Energy*, vol. 118, pp. 473–480, 2017.
- [34] M. Nouri, A. Bekrar, A. Jemai, S. Niar, and A. C. Ammari, "An effective and distributed particle swarm optimization algorithm for flexible job-shop scheduling problem," *Journal Of Intelligent Manufacturing*, vol. 29, no. 3, pp. 603–615, 2018.
- [35] M. Navabi, A. Davoodi, and M. Reyhanoglu, "Optimum fuzzy sliding mode control of fuel sloshing in a spacecraft using PSO algorithm," *Acta Astronautica*, vol. 167, pp. 331–342, 2019.
- [36] B. Zeng and S. Liu, "A self-adaptive intelligence gray prediction model with the optimal fractional order accumulating operator and its application," *Mathematical Methods in the Applied Sciences*, vol. 40, no. 18, pp. 7843–7857, 2017.
- [37] 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.
- [38] W. Meng, B. Zeng, and S. Li, *Grey Forecasting Theory and its Application*, Science Press, Beijing, China, 2020.
- [39] L. F. Wu, S. F. Liu, and L. G. Yao, "Discrete gray model based on fractional order accumulation," *Systems Engineering Theory and Practice*, vol. 34, no. 7, pp. 1822–1827, 2014.
- [40] 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.
- [41] W. Meng and B. Zeng, "Discrete gray model and order optimization based on reciprocal fractional order operator," *Control and Decision Making*, vol. 31, no. 10, pp. 1903–1907, 2016.