Method Article

Adaptive Variable Weight Accumulation AVWA-DGM(1,1) Model Based on Particle Swarm Optimization

Abstract: The development of higher education is an extremely important issue. It is the source of the country's technological innovation and the realization of innovation and development, especially in China, where higher education is still at an exploratory stage. Aiming at the shortcoming that the classical DGM (1,1) model accumulates the raw data series with the weight of constant 1, this paper proposes an adaptive variable weight accumulation optimization DGM (1,1) model, abbreviated as AVWA-DGM (1,1) model. Taking the enrollment numbers of postgraduate, master degree, undergraduate and junior college student and undergraduates students in China as numerical examples, the DGM (1,1) model and AVWA-DGM (1,1) model are established to simulate and predict respectively, and the weighted coefficients of AVWA-DGM (1,1) model are optimized and solved by particle swarm algorithm. The results show that the AVWA-DGM(1,1) model has higher simulation and prediction accuracy than the classical DGM(1,1) model in the four numerical examples provided in this paper. It can be seen that the adaptive accumulation of the original data sequence by the particle swarm optimization algorithm can make the first order accumulation sequence more in line with the requirements of the DGM (1,1) model on the data features, thereby improving the simulation and prediction accuracy.

Key words: Chinese higher education; DGM(1,1) model; AVWA-DGM(1,1) model; Particle swarm optimization; Adaptive variable weight accumulation

1 Introduction

The development of higher education is a concentrated expression of national talent competition and scientific and technological competition, and is the core element for implementing innovation-driven development and building an innovative country. Since China's higher education resumed college entrance examination enrollment and postgraduate education enrollment in 1978, China's higher education has experienced a series of extraordinary developments, and at the same time has harvested many achievements and made significant contributions to the development of all aspects of China. According to the data of the Ministry of Education of China, The enrollment scale of undergraduate and junior college students has reached 7.909 million in 2018. According to the National Graduate Enrollment Survey Report of 2019, the number of master degree students in the national masters reached 2.9 million in 2019, an increase of a record high. Facing the rapid development of higher education in China, scientifically and reasonably predicting the enrollment scale of higher education in the future will further benefit the formulation of higher education system and resource allocation in China, and provide enlightenment for the future development of the country.

The impact of changes in the scale of education on the development of national education is of universal significance. Therefore, many scholars at home and abroad have studied and discussed this and proposed many prediction models. Such as support vector machine [1, 2, 3], neural network [4, 5], time series analysis [6, 7, 8], gray prediction model [9,10,11,12]. Among these prediction models, the gray model has received extensive attention because of its simple calculation and less sample data.

The grey system theory was first proposed by Professor Deng in 1982 [13], which plays a crucial role in dealing with the "small sample" and "poor information" issues. Among them, the grey prediction model is the core part of the grey theory. In the predictive model, the GM (1,1) model is the most classic. At present, the grey prediction model and its improved model have been widely used in various aspects of society, such as energy [14,15], agriculture [16], technology [17], environment [18] and medical [19].In view of this, the majority of experts and scholars are constantly improving and optimizing it. For example, Wu et al. [20] proposed a fractional-order grey prediction model, which optimizes the defect that the first-order accumulation of the grey model can only be an integer. Cui et al. [21] proposed a new grey prediction model and applied it to predict the yield of the concave soil and the CSI 300 index. Luo et al. [22] and Wei et al. [23] studied the GMP (1, 1, N) model with polynomial, where the grey action quantity is $\beta_0+\beta_1 t+\cdots+\beta_N t^N$. Chen and Yu [24] proposed a method to improve the grey action quantity in the NGM (1,1, k, c) model with $bt+c$. Next, Qian et al. [25] proposed a new GM (1,1, t^α) model with a gray action quantity of $bt^\alpha+c$ and used it to predict ground settlement. In recent years, the GM (1, N) model and its promotion model have also received extensive attention. For example, Tien [26, 27], Zeng et al. [28, 29], Wang et al. [30], Ma et al. [31, 32]. However, when the above model performs first-order accumulation processing on the raw data, the weight coefficient of the raw data is constant 1. In response to this problem, some scholars [33, 34, 35, 36] improve the prediction accuracy of the model by establishing different buffer operators to process the raw data. Some scholars [37, 38, 39] make the raw data smoother based on different data transformation techniques. The effect is also significant.

Based on the above literature review, this paper proposes a discrete grey prediction model with adaptive variable weight accumulation, which is abbreviated as AVWA-DGM (1,1) model, and applies the enrollment numbers of postgraduate students, master degree students, undergraduate and junior college students and undergraduate students in China as example data to make simulation and prediction. The calculation results show that the AVWA-DGM(1,1) model is superior to the classical DGM(1,1) model.

2 Traditional DGM (1,1) model

Let $X^{(0)}=\left(x^{(0)}(1),x^{(0)}(2),\cdots,x^{(0)}(n)\right)$ as a non-negative raw sequence. For satisfying a smooth conditional sequence, a grey differential equation can be established. After a *1th-order* accumulation, $X^{(1)}=\left(x^{(1)}(1),x^{(1)}(2),\cdots,x^{(1)}(n)\right)$ is generated. Call $X^{(1)}$ the *1th-order* accumulation generating sequence($1-AGO$) of $X^{(0)}$, where

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

Let non-negative sequence $X^{(0)}$ and *1th-order* accumulation generating sequence $X^{(1)}$ are described above, and call

$$\hat{x}^{(1)}(k+1)=\beta_1\hat{x}^{(1)}(k)+\beta_2, \tag{2}$$

the DGM(1,1) model, or the discrete form of GM(1,1) model [40]. Where the first "1" also represents the first order differential equation, and the second "1" also indicates that there is a variable.

If $\hat{\beta}=[\beta_1,\beta_2]^T$ are parameters, and

$$\Theta_1=\begin{bmatrix} x^{(1)}(2) \\ x^{(1)}(3) \\ \vdots \\ x^{(1)}(n) \end{bmatrix},\ B_1=\begin{bmatrix} x^{(1)}(1) & 1 \\ x^{(1)}(2) & 1 \\ \vdots & \vdots \\ x^{(1)}(n-1) & 1 \end{bmatrix}. \tag{3}$$

Then the least squares estimation parameters $\hat{\beta}=[\beta_1,\beta_2]^T$ of the discrete grey prediction model $\hat{x}^{(1)}(k+1)=\beta_1\hat{x}^{(1)}(k)+\beta_2$ satisfies

$$\hat{\beta}=\left(B_1^T B_1\right)^{-1}B_1^T\Theta_1. \tag{4}$$

Let $\hat{x}^{(1)}(1)=x^{(0)}(1)$ be the recursive function

$$\hat{x}^{(1)}(k+1)=\beta_1^k\left(x^{(0)}(1)-\frac{\beta_2}{1-\beta_1}\right)+\frac{\beta_2}{1-\beta_1},k=1,2,\cdots n-1,\cdots. \tag{5}$$

Restore value is

$$\begin{aligned}\hat{x}^{(0)}(k+1)&=\hat{x}^{(1)}(k+1)-\hat{x}^{(1)}(k)\\&=(\beta_1-1)\left(x^{(0)}(1)-\frac{\beta_2}{1-\beta_1}\right)\beta_1^k,k=1,2,\cdots n-1,\cdots.\end{aligned} \tag{6}$$

3 Adaptive Variable Weight Accumulation Optimized AVWA-DGM(1,1) Model

3.1 Transformation of the raw data sequence

Since a developing system is often disturbed by the impact of changes in the external environment, this leads to the volatility of a certain characteristic data sequence describing the development of the system.The accuracy of prediction will be greatly affected when grey modeling is carried out on such data to predict the future change trend.Common grey prediction models are the GM (1, 1) model and the DGM (1, 1) model.These traditional grey prediction models generally use equal weight accumulation when performing *1th-order* accumulation to generate *1th-order* accumulation sequence (1-AGO), namely, the weight coefficients of each raw data are fixed constants 1.This accumulation method cannot fully exploit the potential information of the raw data sequence, so that the prediction result of the model is not good.Based on this, this paper proposes a variable weight accumulation method, which uses this accumulation method to generate a variable weight accumulation generation sequence (1-AVWAGO).When using this sequence for grey modeling, the variation trend of the raw data sequence is adjusted by adding a weight coefficient to each modeling data, so as to weaken the randomness of the raw data and improve the fitting and prediction accuracy of the model.

Definition 1. Let the raw observation data sequence be $\Upsilon^{(0)}=\left(\gamma^{(0)}(1),\gamma^{(0)}(2),\cdots,\gamma^{(0)}(n)\right)$ and the adjustment weight coefficient be

$$\mu=\left(\mu_1,\mu_2,\cdots,\mu_n\right),\mu_k>0,k=1,2,\cdots,n. \tag{7}$$

Performing a linear weighted transform process on the raw data sequence, and obtaining a weighted new data sequence of $\psi^{(0)}=\left(\varphi^{(0)}(1),\varphi^{(0)}(2),\cdots,\varphi^{(0)}(n)\right)$, where

$$\varphi^{(0)}(k)=\mu_k\gamma^{(0)}(k),(k=1,2,\cdots,n). \tag{8}$$

3.2 Establish an optimized AVWA-DGM (1,1) model

Let $X^{(0)}=\left(x^{(0)}(1),x^{(0)}(2),\cdots,x^{(0)}(n)\right)$ be the raw observation data sequence, $\omega=\left(\omega_1,\omega_2,\cdots,\omega_n\right),\omega_k>0,k=1,2,\cdots,n$ be the weight coefficient, and perform linear weighted transformation on $X^{(0)}$ according to the above formula (8) to obtain $Y^{(0)}=\left(y^{(0)}(1),y^{(0)}(2),\cdots,y^{(0)}(n)\right)$, where

$$y^{(0)}(k)=\omega_k x^{(0)}(k),(k=1,2,\cdots,n). \tag{9}$$

Performing a *1th-order* accumulation on the data sequence $Y^{(0)}$ after the weighted transformation to obtain a weighted *1th-order* accumulation sequence $Y^{(1)}=\left(y^{(1)}(1),y^{(1)}(2),\cdots,y^{(1)}(n)\right)$, where

$$y^{(1)}(k)=\sum_{i=1}^{k}\omega_i x^{(0)}(i),(k=1,2,\cdots,n). \tag{10}$$

The data sequence $Y^{(1)}$ after the weighted transformation process is used to establish the DGM(1,1) model as described above.

$$\hat{y}^{(1)}(k+1)=\alpha_1\hat{y}^{(1)}(k)+\alpha_2. \tag{11}$$

Let $\hat{\alpha}=[\alpha_1,\alpha_2]^T$ be the parameters, if

$$\Theta_2=\begin{bmatrix} y^{(1)}(2) \\ y^{(1)}(3) \\ \vdots \\ y^{(1)}(n) \end{bmatrix},\ B_2=\begin{bmatrix} y^{(1)}(1) & 1 \\ y^{(1)}(2) & 1 \\ \vdots & \vdots \\ y^{(1)}(n-1) & 1 \end{bmatrix}. \tag{12}$$

Then the least squares estimation parameters $\hat{\alpha}=[\alpha_1,\alpha_2]^T$ of the discrete grey prediction model $\hat{y}^{(1)}(k+1)=\alpha_1\hat{y}^{(1)}(k)+\alpha_2$ satisfies

$$\hat{\alpha}=\left(B_2{}^T B_2\right)^{-1} B_2{}^T\Theta_2. \tag{13}$$

Let $\hat{y}^{(1)}(1)=y^{(0)}(1)$ be the recursive function

$$\hat{y}^{(1)}(k+1)=\alpha_1^k\left(y^{(0)}(1)-\frac{\alpha_2}{1-\alpha_1}\right)+\frac{\alpha_2}{1-\alpha_1},k=1,2,\cdots n-1,\cdots. \tag{14}$$

Obtained after subtraction

$$\begin{aligned}\hat{y}^{(0)}(k+1)&=\hat{y}^{(1)}(k+1)-\hat{y}^{(1)}(k)\\ &=(\alpha_1-1)\left(y^{(0)}(1)-\frac{\alpha_2}{1-\alpha_1}\right)\alpha_1^k,k=1,2,\cdots n-1,\cdots.\end{aligned} \tag{15}$$

After the reduction, $\hat{y}^{(0)}(k)$ is obtained, and then the predicted value of the model can be calculated.

$$\begin{aligned}&\hat{x}^{(0)}(k)=\frac{1}{\omega_k}\hat{y}^{(0)}(k),k=1,2,\cdots,n\\ &\hat{x}^{(0)}(k)=\hat{y}^{(0)}(k)=(\alpha_1-1)\left(y^{(0)}(1)-\frac{\alpha_2}{1-\alpha_1}\right)\alpha_1^{k-1},k=n+1,n+2,\cdots.\end{aligned} \tag{16}$$

Where n represents the number of data used for modeling.

3.3 Determination of the optimal weighting coefficient

In order to verify the accuracy of the model and determine the weight coefficients of the weighted transformed AVWA-DGM(1,1) model, absolute percentage error (APE) and mean absolute percentage error (MAPE) are defined. The specific expression are as follows

$$\text{MAPE}=\frac{1}{m-l+1}\sum_{k=l}^{m}\left|\frac{x^{(0)}(k)-\hat{x}^{(0)}(k)}{x^{(0)}(k)}\right|\times 100\%, l\le m\le n, \tag{17}$$

$$\text{APE}(k)=\left|\frac{x^{(0)}(k)-\hat{x}^{(0)}(k)}{x^{(0)}(k)}\right|\times 100\%, k=1,2,\cdots,n. \tag{18}$$

N represents the number of sample data used for modeling.As can be seen from the above formula(17) and (18), when $k=1,2,\cdots,N$, $APE(k)$ is the absolute percentage error of the fitted data. When $k=N+1,N+2,\cdots,n$, $APE(k)$ is the absolute percentage error of the test data. When $l=1,m=N$, $MAPE$ represents the mean absolute percentage error of the simulated data. When $l=N+1,m=n$, $MAPE$ represents the mean absolute percentage error of the test data. When $l=1,m=n$, $MAPE$ represents the mean absolute percentage error of the overall data.

From the modeling process, the unknown parameters existing in the AVWA-DGM(1,1) model are $\omega=(\omega_1,\omega_2,\cdots,\omega_n),\omega_k>0,k=1,2,\cdots,n$ .When the weight coefficients $\omega=(\omega_1,\omega_2,\cdots,\omega_n),\omega_k>0,k=1,2,\cdots,n$ are determined, the parameters $\hat{\alpha}=[\alpha_1,\alpha_2]^T$ can be solved by the least squares method.Therefore, according to the principle of minimum error, choose $\omega=(\omega_1,\omega_2,\cdots,\omega_n),\omega_k>0,k=1,2,\cdots,n$ as the parameters of the optimized MAPE, and establish the following mathematical optimization model.

$$\min MAPE(\omega_1,\omega_2,\cdots,\omega_n)=\frac{1}{n}\sum_{k=1}^{n}\left|\frac{x^{(0)}(k)-\hat{x}^{(0)}(k)}{x^{(0)}(k)}\right|\times 100\%,$$
$$s.t.\begin{cases}\hat{x}^{(0)}(k)=\dfrac{1}{\omega_k}\hat{y}^{(0)}(k),k=1,2,\cdots,n,\\ \hat{y}^{(1)}(1)=y^{(0)}(1),\quad \omega_k>0,k=1,2,\cdots,n,\\ \hat{y}^{(0)}(k+1)=\hat{y}^{(1)}(k+1)-\hat{y}^{(1)}(k)=(\alpha_1-1)\left(y^{(0)}(1)-\dfrac{\alpha_2}{1-\alpha_1}\right)\alpha_1^k,\\ k=1,2,\cdots n-1,\\ \hat{y}^{(1)}(k+1)=\alpha_1^k\left(y^{(0)}(1)-\dfrac{\alpha_2}{1-\alpha_1}\right)+\dfrac{\alpha_2}{1-\alpha_1},k=1,2,\cdots n-1,\\ \hat{\alpha}=\left(B_2^TB_2\right)^{-1}B_2^T\Theta_2.\end{cases} \tag{19}$$

Considering the complexity of equation (19), solving the optimal $\omega=(\omega_1,\omega_2,\cdots,\omega_n),\omega_k>0,k=1,2,\cdots,n$ are very difficult. Based on this, this paper uses the particle swarm optimization algorithm to find the optimal $\omega=(\omega_1,\omega_2,\cdots,\omega_n),\omega_k>0,k=1,2,\cdots,n$ value.

The Particle Swarm Optimization (PSO) algorithm was proposed by Kennedy and Eberhart [41]. The algorithm is based on the simulation of the social activities of the flocks, and proposes a global random search algorithm based on swarm intelligence by simulating the behavior of the flocks interacting with each other.The particle swarm algorithm first randomly initializes the particle swarm in the solution space and initializes the velocity and position.The dimension of the solution space is determined by the number of variables to be optimized.Each position of the particle in the search space is a solution to the problem to be optimized, and each particle is given a velocity, which determines the flight distance and direction of the particle, so that the particle can fly within to the solution space and land on the optimal solution. Each particle in the swarm is evaluated by the objective function to determine the fitness value to determine the pros and cons of the current position, while the particles endowed with memory function record the current optimal position searched.Through iterative optimization, each particle in the group keeps track of two extremes case.Where, the individual best is recorded in $pbest$, the global best is recorded in $gbest$, and the position and flight speed of the particle in the solution space are updated according to the two records.The particle swarm then follows the current optimal particle and continues searching in the solution space.The steps of the algorithm are specifically shown below.

Step1: Initialize the population particle number $M=100$, particle dimension $n=7$, maximum iteration number k_{max}, learning factor $l_1=1.5, l_2=1.5$, inertia weight maximum value $w_{max}=0.8$ and minimum value $w_{min}=0.4$.

Step2: Initialize the population particle maximum position $\omega_{max}=\left(\omega_{1,max}, \omega_{2,max}, \cdots, \omega_{n,max}\right)$, minimum position $\omega_{min}=\left(\omega_{1,min}, \omega_{2,min}, \cdots, \omega_{n,min}\right)$, maximum speed $v_{max}=\left(v_{1,max}, v_{2,max}, \cdots, v_{n,max}\right)$, minimum speed $v_{min}=\left(v_{1,min}, v_{2,min}, \cdots, v_{n,min}\right)$, particle individual optimal position $pbest_i^1$ and optimal value p_i^1, and particle group global optimal position $gbest^1$ and optimal value g^1;

Step3: calculating the fitness value $MAPE\left(\omega_{i,1}^k, \omega_{i,2}^k, \cdots, \omega_{i,n}^k\right)$ of each particle in the particle group;

Step 4: Compare each particle fitness value $MAPE\left(\omega_{i,1}^k, \omega_{i,2}^k, \cdots, \omega_{i,n}^k\right)$ with the individual extreme value p_i^k and the particle group global optimal value g^k, respectively. If $MAPE\left(\omega_{i,1}^k, \omega_{i,2}^k, \cdots, \omega_{i,n}^k\right) < p_i^k$, replace p_i^k with $MAPE\left(\omega_{i,1}^k, \omega_{i,2}^k, \cdots, \omega_{i,n}^k\right)$ and replace the particle's individual optimal position $pbest_i^k$. If $MAPE\left(\omega_{i,1}^k, \omega_{i,2}^k, \cdots, \omega_{i,n}^k\right) < g^k$, replace g^k with $MAPE\left(\omega_{i,1}^k, \omega_{i,2}^k, \cdots, \omega_{i,n}^k\right)$ and replace the global optimal position $gbest^k$ of the particle group;

Step 5: Calculate the dynamic inertia weight w according to the following formula;

$$w = w_{max} - k\left(w_{max} - w_{min}\right) / k_{max}.$$

Step6: Update the velocity value $v_{i,j}^k$ and the position $\omega_{i,j}^k$ according to the following iteration formula and perform boundary condition processing, where $i = 1, 2, \cdots, M$, $j = 1, 2, \cdots, n$;

$$\begin{aligned} v_{i,j}^{k+1} &= wv_{i,j}^k + l_1 \times rand(0,1) \times \left(pbest_{i,j}^k - \omega_{i,j}^k\right) + \\ &\quad l_2 \times rand(0,1) \times \left(gbest_j^k - \omega_{i,j}^k\right), \\ \omega_{i,j}^{k+1} &= \omega_{i,j}^k + v_{i,j}^{k+1}. \end{aligned} \tag{20}$$

Step7: Judge whether the algorithm termination condition is satisfied: if yes, end the algorithm and output the optimization result: otherwise return to Step3.

Compared with the classical DGM (1,1) model, the AVWA-DGM (1,1) model proposed in this paper, namely the adaptive variable weight accumulation DGM (1,1) model, is more widely applicable.After combining the PSO algorithm, the classical DGM (1,1) model is optimized with a fixed weight for the first-order accumulation process, and the adaptive change of the weighting coefficients is realized.The accumulation of the raw data sequence with adaptive weights is more likely to exploit the underlying internal information of the raw data sequence than the fixed weight accumulation of the raw data sequence. Moreover, after the raw data sequence is accumulated by using the adaptive weights method, the *1th-order* accumulation generation sequence can be made to conform to the characteristic requirements of the data of the DGM (1, 1) model.

4 Application of AVWA-DGM(1,1) model

This part will show the accuracy of the adaptive weighted optimized AVWA-DGM(1,1) model under actual data.The modeling results were compared with the classical DGM (1, 1) model.Where, the weighting coefficient $\omega = \left(\omega_1, \omega_2, \cdots, \omega_n\right), \omega_k > 0, k = 1, 2, \cdots, n$ of the AVWA-DGM (1,1) model is determined by the PSO. The article uses the actual enrollment of Chinese higher education from the China Statistical Yearbook [42] 2005-2016 as an example to illustrate the superiority of the AVWA-DGM (1,1) model. This paper divides the data into two parts, namely, the modeling data from 2005 to 2011 and the test data of the model from 2012 to 2016. The raw data is shown in Table 1.

Table 1 Actual enrollment of Chinese higher education in 2005-2016

Year	postgraduate	master degree	undergraduate and junior college	undergraduate
2005	36.4831	31.0037	504.5	236.3647
2006	39.7925	34.197	546.1	253.0854
2007	41.8612	36.059	565.9	282.0971
2008	44.6422	38.6658	607.7	297.0601
2009	51.0953	44.9042	639.5	326.1081
2010	53.8177	47.4415	661.8	351.2563
2011	56.0168	49.4609	681.5	356.6411
2012	58.9673	52.1303	688.8	374.0574
2013	61.1381	54.0919	699.8	381.4331
2014	62.1323	54.8689	721.4	383.4152
2015	64.5055	57.0639	737.8	389.4184
2016	66.7064	58.9812	748.6	405.4007

4.1 Number of postgraduates enrolled in China

This section combines the particle swarm optimization algorithm and the actual data provided by the China Statistical Yearbook to study the number of postgraduates enrollment scale in China by establishing the DGM (1,1) model and the AVWA-DGM (1,1) model.The final calculation results and weighting coefficients (both reserved for four decimal places) are given in Table 2, Table 3 and Fig.1, Fig.2.It can be seen from Table 2 that when the AVWA-DGM (1,1) model is accumulated, the weights of the raw data are not all constant 1, but the corresponding optimal weight coefficients are given according to the characteristics of the raw data sequence itself. As can be seen from Table 3 and Fig.1, both grey models reflect the changing trend of the number of postgraduates enrolled in China. As can be seen from Table 3, the simulated MAPE of the DGM (1,1) model, the MAPE of the test data and the overall MAPE were 1.6791%, 13.7769% and 6.7199%, respectively, while the AVWA-DGM (1,1) was 3.53×10^{-13}%, 0.3485% and 0.1452%, respectively.These results indicate that the AVWA-DGM (1,1) model is more accurate than the DGM (1,1) model in predicting the trend of postgraduates enrollment in China.

Table 2 Weighting coefficients of the two models

Model	Weight coefficient
DGM(1,1)	$(1,1,1,1,1,1,1)$
AVWA-DGM(1,1)	$(1.0000,1.2376,1.2123,1.1714,1.0547,1.0318,1.0215)$

Table 3 Calculation results and errors of the two models

Year	Raw data	DGM(1,1)	APE(%)	AVWA-DGM(1,1)	APE(%)
2005	36.4831	36.4831	0.0000	36.4831	0.0000
2006	39.7925	39.5405	0.6333	39.7925	0.0000
2007	41.8612	42.5517	1.6494	41.8612	0.0000
2008	44.6422	45.7921	2.5758	44.6422	0.0000
2009	51.0953	49.2793	3.5541	51.0953	0.0000
2010	53.8177	53.0321	1.4597	53.8177	0.0000
2011	56.0168	57.0707	1.8814	56.0168	0.0000
2012	58.9673	61.4168	4.1540	58.9673	0.0000
2013	61.1381	66.0939	8.1059	60.7646	0.6108
2014	62.1323	71.1271	14.4769	62.6168	0.7797
2015	64.5055	76.5437	18.6623	64.5253	0.0308
2016	66.7064	82.3728	23.4856	66.4921	0.3213
simulation MAPE		1.6791		3.53×10^{-13}	
forecast MAPE		13.7769		0.3485	
overall MAPE		6.7199		0.1452	

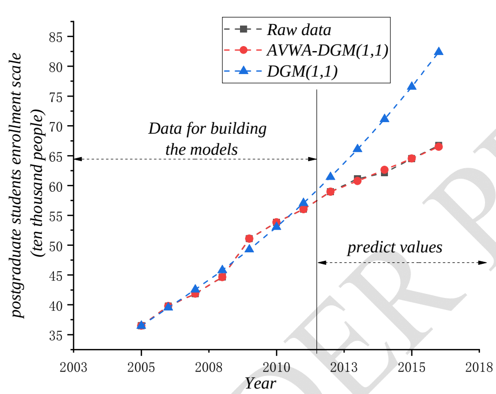

Fig.1 Comparison of simulation and prediction of two models in postgraduate's enrollment

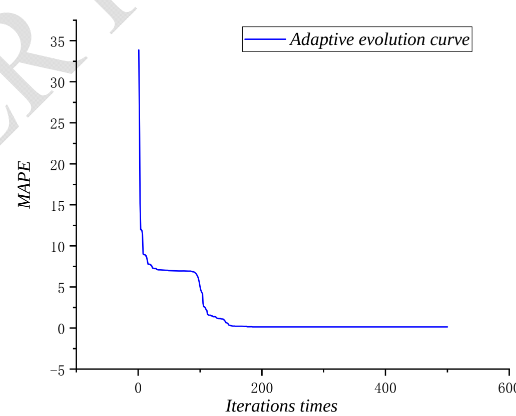

Fig.2 PSO algorithm fitness evolution curve

4.2 China's master degree student's enrollment

Similar to the previous section, the AVWA-DGM (1,1) model and AVWA-DGM (1,1) model were established, and the parameters of AVWA-DGM (1,1) model were solved by particle swarm optimization.The resulting final calculation results and weighting coefficients (both reserved for four decimal places) are given in Table 4, Table 5 and Fig.3, Fig.4.Table 4 also shows that the weight coefficients of the AVWA-DGM (1,1) model are not all constant 1.It can be seen from Fig.3 that compared with the DGM (1,1) model, the AVWA-DGM (1,1) model can more accurately predict the changing trend of the number of master degree students in China. As can be seen from table 5 and table 1, the simulated MAPE of DGM

(1,1), the MAPE of test data and the overall MAPE were 1.9163%, 16.0442% and 7.8029%, respectively, while the AVWA-DGM (1,1) are 4.31×10^{-11} %, 0.3764% and 0.1568%, respectively.

Table 4 Weighting coefficients of the two models

Model	Weight coefficient
DGM(1,1)	$(1,1,1,1,1,1,1)$
AVWA-DGM(1,1)	$(1.0000,1.2682,1.2402,1.1926,1.0589,1.0335,1.0221)$

Table 5 Calculation results and errors of the two models

Year	Raw data	DGM(1,1)	APE(%)	AVW-DGM(1,1)	APE(%)
2005	31.0037	31.0037	0.0000	31.0037	0.0000
2006	34.1970	33.9552	0.7070	34.1970	0.0000
2007	36.0590	36.7624	1.9507	36.0590	0.0000
2008	38.6658	39.8017	2.9377	38.6658	0.0000
2009	44.9042	43.0922	4.0353	44.9042	0.0000
2010	47.4415	46.6548	1.6583	47.4415	0.0000
2011	49.4609	50.5119	2.1248	49.4609	0.0000
2012	52.1303	54.6878	4.9061	52.1303	0.0000
2013	54.0919	59.2091	9.4601	53.7540	0.6247
2014	54.8689	64.1041	16.8313	55.4282	1.0193
2015	57.0639	69.4037	21.6246	57.1546	0.1589
2016	58.9812	75.1416	27.3992	58.9347	0.0788
simulation MAPE		1.9163		4.31×10^{-11}	
forecast MAPE		16.0442		0.3764	
overall MAPE		7.8029		0.1568	

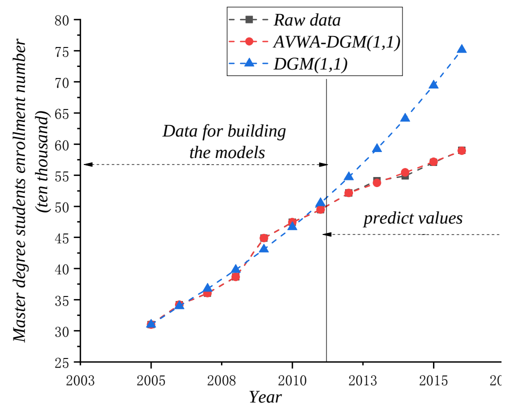

Fig.3 Comparison of simulation and prediction of two models in master degree student's enrollment

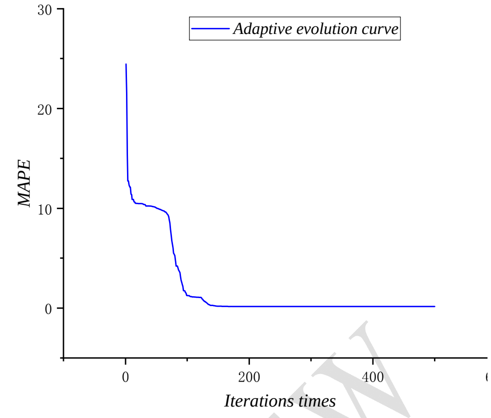

Fig.4 PSO algorithm fitness evolution curve

4.3 Enrollment scale of Chinese undergraduates and junior college students

Similarly, the DGM(1,1) model and AVWA-DGM (1,1) model were used to model and predict the enrollment scale of undergraduate and junior college students in China, and the parameters of AVWA-DGM (1,1) model were optimized and solved by particle swarm optimization algorithm.The final calculations for both models (all retaining four decimal places) are given in Table 6, Table 7, and Fig.5, Fig.6.Table 6 shows that the weight coefficients of the AVWA-DGM (1,1) model varies with the raw data and is not a fixed constant. As can be seen from Fig.5, when modeling and forecasting the enrollment scale of undergraduate and junior college students in China, the simulation and prediction accuracy of AVWA-DGM (1,1) model is higher than the DGM(1,1) model. In Table 5, the simulated MAPE of DGM (1,1), the MAPE of the test data, and the overall MAPE are 0.8899%, 10.1810%, and 4.7612%, respectively, while AVWA-DGM(1,1) are $5.4059\times10^{-13}\%$, 0.3184% and 0.1327%, respectively.

Table 6 Weighting coefficients of the two models

Model	Weight coefficients
DGM(1,1)	$(1,1,1,1,1,1,1)$
AVWA-DGM (1,1)	$(1.0000,1.1006,1.0865,1.0350,1.0062,0.9946,0.9880)$

Table 7 Calculation results and errors of the two models

Year	Raw data	DGM(1,1)	APE(%)	AVWA-DGM(1,1)	APE(%)
2005	504.5	504.5	0.0000	504.5	0.0000
2006	546.1	548.6084	0.4593	546.1	0.0000
2007	565.9	574.3444	1.4922	565.9	0.0000
2008	607.7	601.2876	1.0552	607.7	0.0000

2009	639.5	629.4948	1.5645	639.5	0.0000
2010	661.8	659.0253	0.4193	661.8	0.0000
2011	681.5	689.9411	1.2386	681.5	0.0000
2012	688.8	722.3071	4.8646	688.8	0.0000
2013	699.8	756.1915	8.0582	704.6215	0.6890
2014	721.4	791.6655	9.7402	720.8065	0.0823
2015	737.8	828.8036	12.3345	737.3632	0.0592
2016	748.6	867.6839	15.9075	754.3002	0.7614
simulation MAPE		0.8899		5.4059×10^{-13}	
forecast MAPE		10.1810		0.3184	
overall MAPE		4.7612		0.1327	

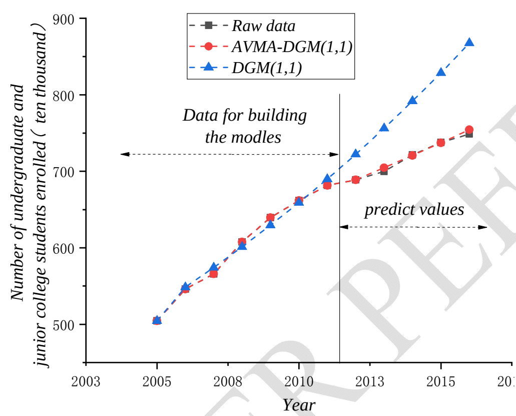

Fig.5 Comparison of simulation and prediction of two models in the number of enrollment scale undergraduate and junior college students

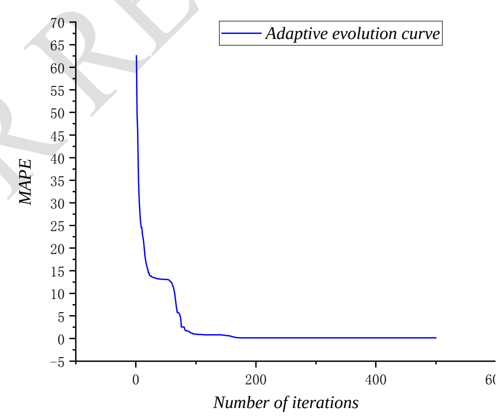

Fig.6 PSO algorithm fitness evolution curve

4.4 Number of students enrolled in undergraduates in China

In this section, we use grey theory to study the number of undergraduate enrollments scale in China. The DGM (1,1) model was established and compared with the AVWA-DGM (1,1) model. Similarly, the particle swarm algorithm is used to optimize the parameters of the AVWA-DGM (1,1) model. The final calculations for both models (all retaining four decimal places) are given in Table 8, Table 9, and Fig.7, Fig.8. The weight coefficients of the *1th-order* accumulation generation sequence of the DGM (1, 1) model and the AVWA-DGM (1, 1) model are compared in Table 8.The results show that the weight coefficients of the AVWA-DGM (1,1) model are also a sequence that varies with the raw data sequence and is not a fixed constant. In Table 9, the simulation MAPE of DGM(1,1), the MAPE of the test data, and the overall MAPE are 1.6922%, 16.7266%, and 7.9565%, respectively, while AVWA-DGM(1,1) are 0.0028%, 0.7061% and 0.2958%, respectively. It can be seen from Fig. 7 that the simulation accuracy and prediction accuracy of the AVWA-DGM (1, 1) model is higher than the DGM (1, 1) model.

Table 8 Weighting coefficients of the two models

Model	Weight coefficients
DGM(1,1)	$(1,1,1,1,1,1,1)$
AVWA-DGM(1,1)	$(1.0001,1.3147,1.2029,1.1648,1.0819,1.0242,1.0285)$

Table 9 Calculation results and errors of the two models

Year	Raw data	DGM(1,1)	APE(%)	AVWA-DGM(1,1)	APE(%)
2005	236.3647	236.3647	0.0000	236.3647	0.0000
2006	253.0854	260.4513	2.9104	253.0854	0.0000
2007	282.0971	278.8862	1.1382	282.0834	0.0049
2008	297.0601	298.6259	0.5271	297.0289	0.0105
2009	326.1081	319.7628	1.9458	326.1081	0.0000
2010	351.2563	342.3959	2.5225	351.2563	0.0000
2011	356.6411	366.6308	2.8011	356.6553	0.0040
2012	374.0574	392.5812	4.9521	374.0574	0.0000
2013	381.4331	420.3683	10.2076	381.4232	0.0026
2014	383.4152	450.1223	17.3981	388.9340	1.4394
2015	389.4184	481.9822	23.7698	396.5927	1.8423
2016	405.4007	516.0972	27.3055	404.4023	0.2463
simulation MAPE		1.6922		0.0028	
forecast MAPE		16.7266		0.7061	
overall MAPE		7.9565		0.2958	

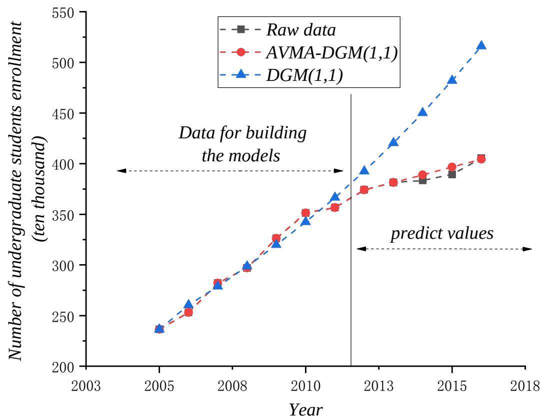

Fig.7 Comparison of simulation and prediction of two models in undergraduate enrollment students

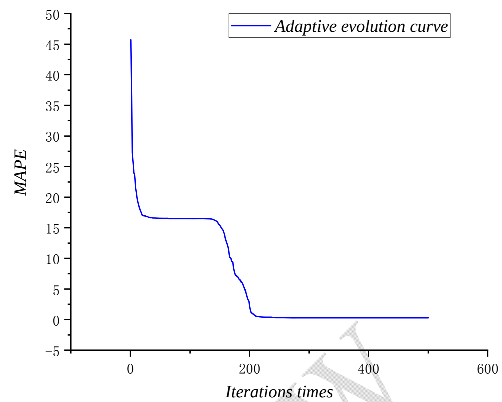

Fig.8 PSO algorithm fitness evolution curve

5 Conclusion

In this paper, the *1th-order* accumulation sequence of the classical DGM (1,1) model is changed by weighting, and the weighting coefficients are optimized by particle swarm optimization algorithm to obtain the optimal weight coefficients, and the AVWA-DGM (1,1) model is proposed. The results show that when performing the *1th-order* accumulation, the original data is given an appropriate weight, and then the *1th-order* accumulation is performed, which can change the variation characteristics of the original data sequence. Through the sequence of the optimal weight coefficient, the data variation characteristics of the sequence are more in line with the data requirements of the DGM (1,1) model, so that the solution of the parameters becomes more accurate, thus making the DGM(1,1) model and the simulation and prediction accuracy are effectively improved. According to the results of the four numerical examples provided in this paper, the AWVA-DGM(1,1) model can effectively improve the prediction accuracy of the DGM(1,1) model, and has certain theoretical significance and application value.

7 References

[1] Aksenova SS , Zhang D , Lu ML . Enrollment Prediction through Data Mining IEEE International Conference on Information Reuse & Integration. IEEE, 2006.
[2] Chen SL. Modeling and prediction of postgraduates enrollment scale. Computer Simulation, 2012, 29(02): 396-399.
[3] Zhang M, Jiao HJ. Application of least squares support vector machine for forecast of regular higher learning institution enrollment. Journal of Zhengzhou Institute of Aeronautical Industry Management, 2008(01): 142-144.
[4] Xu J, Yang Y, Zhang R. Graduate enrollment prediction by an error back propagation algorithm based on the multi-experiential particle swarm optimization 2015 11th International Conference on Natural Computation (ICNC). IEEE, 2015: 1159-1164.
[5] Li HX, Li CW. Application of Artificial Neural Network in Prediction the number of

Postgraduates. Mathematics in Practice and Theory,2009,39(12):27-33.

[6] Sah M, Konstantin Y. Forecasting enrollment model based on first-order fuzzy time series. In World Academy of Science, Engineering and Technology, 2005, 1: 375-378.

[7] Lee MH, Efendi R, Ismail Z. Modified weighted for enrollment forecasting based on fuzzy time series. Matematika, 2009, 25: 67-78.

[8] Sun MJ, Chen BF, Wen CH, Ren JZ. Modeling and Forecasting of Graduate Enrollment Scale Based on ARIMA Model. Statistics & Decision,2010(12):60-62.

[9] Gao YF. Research and application of grey system theory in students' enrollment management in general institutions of higher learing. Liaoning Technical University, 2005.

[10] Yang F, Wu YJ. forecasting of the student intake of a certain college based on GM(1,1) model. Journal of Changchun University,2008(04):29-30.

[11] Wang D. The application of grey theory in enrollment system. Sci-Tech Information Development & Economy,2008(13):183-184.

[12] Yang YC, Wang XY, Zhang XB. Equal dimension and new information model in graduate recruitment with correction factors. Journal of University of Science and Technology Liaoning,2010,33(03):285-288.

[13] Deng JL, Control problems of grey systems, Systems & Control Letters 1 (5)(1982) 288–294.

[14] Zeng B, Tan YT, Xu H, Quan J, Wang LY, Zhou XY. Forecasting the Electricity Consumption of Commercial Sector in Hong Kong Using a Novel Grey Dynamic Prediction Model. Journal of Grey System, 2018, 30(1).

[15] Ding S, Hipel KW, Dang YG. Forecasting China's electricity consumption using a new grey prediction model. Energy, 2018, 149: 314-328.

[16] Ou SL. Forecasting agricultural output with an improved grey forecasting model based on the genetic algorithm. Computers and electronics in agriculture, 2012, 85: 33-39.

[17] Javed S A, Liu SF. Predicting the research output/growth of selected countries: application of Even GM (1, 1) and NDGM models. Scientometrics, 2018, 115(1): 395-413.

[18] Dai HW, Chen XX, Sun XT. The gray prediction model of the atmospheric pollutants in Shenzhen based on GM(1,1) model group. Mathematics in Practice and Theory,2014,44(01):131-136.

[19] Zhang F, Sun T, Fan LH. Application of GM(1,1) model in the prediction of social medical insurance pooling fund. China Health Resources,2009,12(06):269-270+283.

[20] Wu LF. Fractional order grey forecasting models and their application. Nanjing University of Aeronautics and Astronautics, 2015.

[21] Cui J, Liu SF, Zeng B, Xie NM. A novel grey forecasting model and its optimization. Applied Mathematical Modelling, 2013, 37(6): 4399-4406.

[22] Luo D, Wei BL. Grey forecasting model with polynomial term and its optimization. optimization, 2017, 29(3): 58-69.

[23] Wei BL, Xie NM, Hu A. Optimal solution for novel grey polynomial prediction mode. Applied Mathematical Modelling, 2018, 62: 717-727.

[24] Chen PY, Yu HM. Foundation settlement prediction based on a novel NGM model. Mathematical Problems in Engineering, 2014, 2014.

[25] W. Qian, Y. Dang, S. Liu, Grey GM(1,1,t^{α}) model with time power and its application,

Systems Engineering–Theory & Practice 32 (10) (2012) 2247–2252.

[26] Tien TL. The indirect measurement of tensile strength of material by the grey prediction model GMC (1, n). Measurement Science and Technology, 2005, 16(6): 1322.

[27] Tien TL. A research on the grey prediction model GM (1, n). Applied mathematics and computation, 2012, 218(9): 4903-4916.

[28] Zeng B, Liu SF. A self-adaptive intelligence gray prediction model with the optimal fractional order accumulating operator and its application. Mathematical Methods in the Applied Sciences, 2017, 40(18): 7843-7857.

[29] Zeng B, Li C. Improved multi-variable grey forecasting model with a dynamic background-value coefficient and its application. Computers & Industrial Engineering, 2018, 118: 278-290.

[30] Wang ZX. A predictive analysis of clean energy consumption, economic growth and environmental regulation in China using an optimized grey dynamic model. Computational Economics, 2015, 46(3): 437-453.

[31] Ma X, Liu ZB. The GMC (1, n) model with optimized parameters and its application. Journal of grey system, 2017, 29(4): 121-137.

[32] Ma X. A brief introduction to the Grey Machine Learning. arXiv preprint arXiv:1805.01745, 2018.

[33] Xie NM, Liu SF. Discrete GM(1,1)and mechanism of grey forecasting model. Systems Engineering —Theory & Practice,2005(01):93-99.

[34] Yin CH, Gu PL. Energy forecast based on grey series spanning buffering operator. Journal of Systems Engineering,2003(02):189-192.

[35] Han R. Study on the weakening buffer operator based on the strictly monotone function. Journal of Communication University of China(Natural Science Edition), 2018, 25(04): 27-31.

[36] Cui J, Dang YG, Liu SF. Analysis on modeling accuracy of GM(1,1) based on new weakening operators. Systems Engineering —Theory & Practice,2009,29(07):132-138.

[37] Dai WZ, Xiong W, Yang AP. Gray modeling based on $\cot(x^{\alpha})$ transformation and background value optimization. Journal of Zhejiang University(Engineering Science),2010,44(07):1368-1372.

[38] Guan YQ, Liu SF. Grey GM(1,1) An approach to grey modeling based on $\cot(x^{\alpha})$ transformation. System Engineering,2008(09):89-93.

[39] Wei Y, Zhang Y. A Criterion for Comparing the Function Transformations to Raise Smooth Degree in Grey Modeling Data. Journal of Grey System, 2007, 19(1).

[40] Zeng B, Liu SF. DGM(1,1) direct modeling approach of DGM(1,1) with approximate non-homogeneous exponential sequence. Systems Engineering —Theory & Practice,2011,31(02):297-301.

[41] Kennedy J, Eberhart RC. Particle swarm optimization (PSO)Proc. IEEE International Conference on Neural Networks, Perth, Australia. 1995: 1942-1948.

[42] National Bureau of Statistics of China. China Statistical Yearbook (2018). Beijing: China Statistics Press, 2018.