Forecasting manufacturing industrial natural gas consumption of China using a novel time-delayed fractional grey model with multiple fractional order

Yu Hu[1,5] · Xin Ma[1,2] ⓘ · Wanpeng Li[3] · Wenqing Wu[1] · Daoxing Tu[4]

Abstract

Improving the proportion of natural gas consumption of the manufacturing industry would make significant contributions to the low-carbon and sustainable development of China, which is one of the largest manufacturers in the world. However, it is very difficult to catch the trend of natural gas consumption of the concerning manufacturing industry as not enough trustable data can be collected. To fill this gap, a novel time-delayed fractional grey model is developed to forecast the natural gas consumption concerning time-delayed effect. Theoretical analysis shows it has more general formulation, unbiasedness and higher flexibility than the existing similar model. Being optimized by the Particle Swarm Optimization algorithm, the proposed model presents higher accuracy in four validation cases. Finally, it is used to forecast the natural gas consumption of the manufacturing industry of China, and the results show that the proposed model significantly outperforms the other seven existing grey models.

Keywords Green manufacturing · Natural gas consumption · Low-carbon production · Time-delayed grey model · Fractional grey model · Particle swarm optimization

Mathematics Subject Classification 91B84

✉ Xin Ma
 cauchy7203@gmail.com

1 School of Science, Southwest University of Science and Technology, Mianyang, China

2 State Key Laboratory of Oil and Gas Reservoir Geology and Exploitation, Southwest Petroleum University, Chengdu, China

3 School of Computing, Mathematics and Digital Technology, Manchester Metropolitan University, Manchester, UK

4 School of Science, Southwest Petroleum University, Chengdu, China

5 College of Mathematics and Statistics, Chongqing University, Chongqing, China

1 Introduction

The manufacturing industry can directly reflect a country's productivity level, and energy is an important material basis for human survival and development. According to Ref National bureau of statistics (2002), in the past decade, the total energy consumption of the manufacturing industry has been on a steady trend, accounting for about 57% of the total energy consumption of the whole country; however, the gross domestic product (GDP) of the manufacturing industry only accounts for about 31% of the total, showing a sharp downward trend. It shows that the energy consumed by the manufacturing industry is not proportional to its contribution to the national GDP, and the consumption structure of the manufacturing energy needs to be further improved.

Natural gas is a kind of high-quality, efficient, and clean low-carbon energy. With the reform of natural gas prices and the vigorous promotion of natural gas development in the 13th five-year plan, the development of natural gas will usher in historic opportunities. According to Ref National bureau of statistics (2002), in the past decade, the consumption of natural gas in the manufacturing industry has been on an upward trend, accounting for about 40% of the national consumption of natural gas, indicating that the 18th national congress of the communist party of China proposed to vigorously promote the construction of ecological civilization and play a positive role in promoting the use of natural gas; however, the consumption of natural gas in manufacturing industry only accounts for about 0.34% of the total energy consumption in the manufacturing industry, and the natural gas consumption of manufacturing industry accounts for less than 0.4% of manufacturing GDP. It shows that the government's actions on the development of the natural gas manufacturing industry still needs to be accelerated. As a consequence, China is in the stage of reforming the energy consumption system, and the situation of the market and economics is changing fast. This brought more uncertainties to the energy consumption system of China. Meanwhile, under such circumstances, often the newest few data are available for accurate forecasting of energy consumption. Thus a tool which is efficient in dealing with uncertainties with small samples is needed.

Grey system theory proposed by Deng is such a tool which is available to deal with the problems described above (Ref Julong (1986)), in which the grey models play a key role. Unlike the white box models, such as the differential equations in Refs Wang et al. (2019), or the black box models, like the machine learning models in Refs Yang et al. (2019), Pei et al. (2019), Fan et al. (2019), the grey models essentially try to combine the merits of these models in order to take most advantages of the infomation. Moreover,it was proved to be very efficient in small sample modeling for time series forecasting in Ref Lifeng et al. (2013). Within such priority, the grey models have been applied in a wide variety of fields in recent years, such as dollar to euro price forecasting in Ref Kayacan et al. (2010), passenger demand growth forecasting in the air transportation industry in Ref Benítez et al. (2013), the actual cost and the cost at completion of a project forecasting in Ref San Cristóbal et al. (2015), the scrapped vehicles forecasting in Ref Ene and Öztürk (2017), short-term freeway traffic parameter prediction in Ref Bezuglov and Comert (2016) , the e-waste in Washington in Ref Duman et al. (2019), total natural gas consumption forecasting in Ref Zeng et al. (2020), pollutant forecasting in Xiong et al. (2020), traffic flow prediction Ref Xiao et al. (2020), etc. But it was also pointed out by Wu that the conventional grey models based on first-order accumulation is not flexible enough to deal with more complex sequences; thus the fractional order accumulation was introduced for grey models in Ref Lifeng et al. (2013). A series of theoretical analysis was also provided in the following research, such as the sensitivity of

initial condition in Ref Lifeng et al. (2015), ability of mining new information in Ref Lifeng and Bin (2017). Within such advantages the fractional grey models soon become popular and were applied in many new fields in recent years, such as transaction counts forecasting in Ref Gatabazi et al. (2019), and even the new coronavirus (COVID-19) cases forecasting in Ref Utkucan and Tezcan (2020).

On the other hand, the fractional order grey models are also suitable for energy forecasting with its high flexibility and effectiveness in small sample modeling. Wu et al. proposed the FGM(1,1) and made a more accurate prediction of the coal mine drainage volume in Ref Lifeng et al. (2014). Shaikh et al. constructed China's natural gas consumption forecasting model by utilizing two optimized nonlinear grey models: the Grey Verhulst Model and the Nonlinear Grey Bernoulli Model in Ref Shaikh et al. (2017). Wang et al. established a novel hybrid forecasting model based on an improved grey forecasting mode optimized by a multi-objective ant lion optimization algorithm and solved the problem of accuracy and stability of annual power consumption data in Ref Wang et al. (2018). Wu et al. used the GM(1,1) model with the fractional order accumulation (FGM(1,1)) to predict the future trend of air quality, and the results can be directly exploited in the decision-making processes for air quality management in Ref Lifeng et al. (2018). Moonchai et al. proposed a novel method based on the modification of the multivariate grey forecasting model and applied it to the consumption forecast of renewable energy in Ref Moonchai and Chutsagulprom (2020). Based on the new information priority principle and combined with grey buffer operator technology, Zeng B. realized the scientific forecast of shale gas production in my country in Ref Zeng et al. (2020). Utkucan S. built a fractional nonlinear grey Bernoulli model, briefly as FANGBM(1,1) to forecast of Turkey's total renewable and hydro energy in Ref Şahin (2020).

However, it should be noticed that the existing fractional order grey models only use a unified fractional order. As will be discussed in this work, such operation will limit the advantages of the fractional order accumulation, leading to less flexibility of the fractional grey models. Io present a more flexible modeling formulation, a time-delayed fractional discrete grey model with multiple fractional orders (TDF-DGM$_M$) is established in this work. The particle swarm optimization (PSO) is used to calculate the optimal orders r_1 and r_2 of TDF-DGM$_M$ model. Four cases were used to verify the validity and accuracy of the model. Finally, the TDF-DGM$_M$ model is used to predict the gas consumption of China's manufacturing industry.

The rest of this paper is organized as follows: Brief overview of background is shown in Sect. 2; a brief introduction of the fractional grey model (FGM) is presented in Sect. 3; the representation and modeling procedures of the TDF-DGM$_M$ are described in Sect. 4; relationship and difference between the TDF-DGM$_M$ and FTDGM model is analyzed in Sect. 5; the PSO for optimizing the proposed model is presented in Sect. 6; four case studies to verify the validity of the model are shown in Sect. 7; the case study of forecasting the natural gas consumption in China's manufacturing industry is shown in Sect. 8, and the conclusions are drawn in Sect. 9.

2 Brief overview of background

By consulting relevant data in Ref National bureau of statistics (2002), we have collected the development trend of the GDP of various industries, as shown in Fig. 1.

As can be seen from the Fig. 1, the manufacturing industry accounts for a large proportion of GDP in comparison with other industries, maintaining at about 31%, and its GDP has
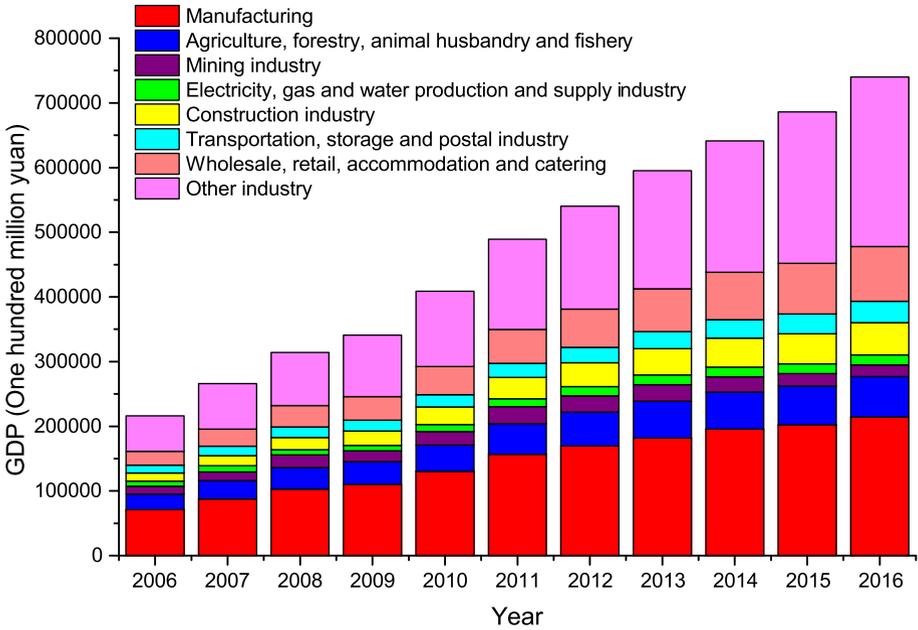
Fig. 1 The gross domestic product change trend of countrywide each industry
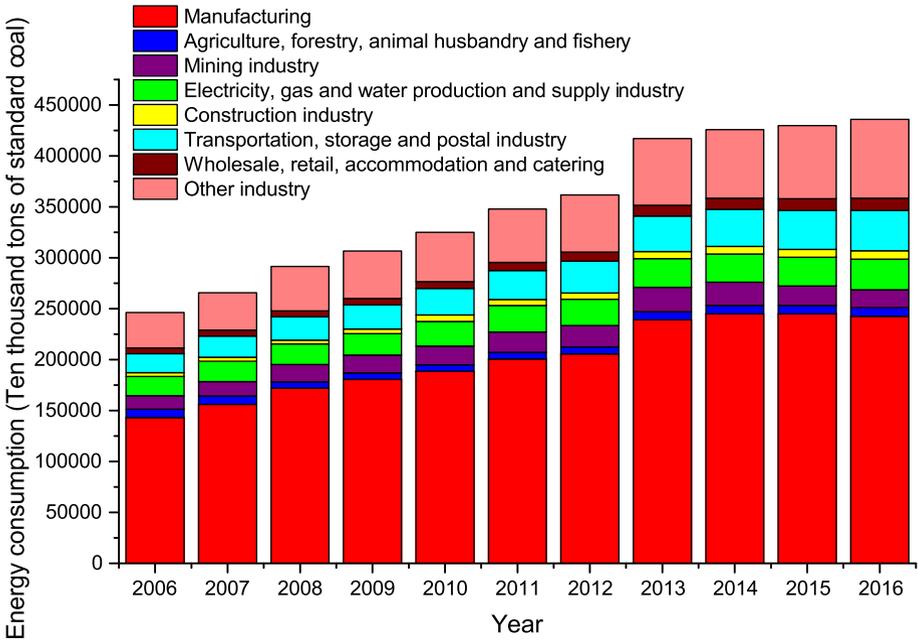
Fig. 2 Trends in energy consumption across the countrywide each industry
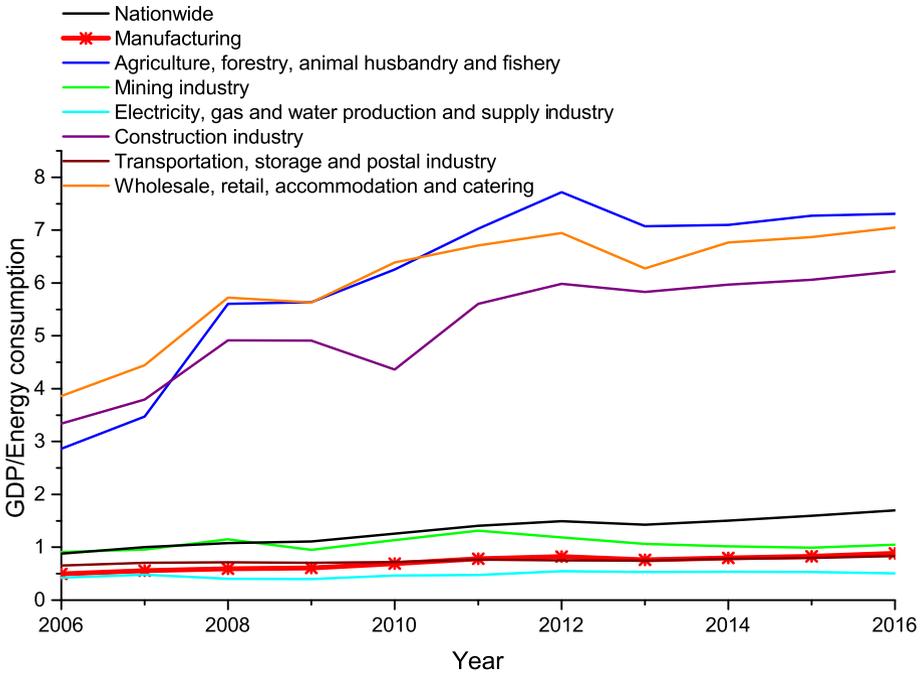
Fig. 3 Trends in the ratio of GDP to energy consumption across the country and in various industries

been on a steady rise, but the proportion is declining. Given the great contribution of the manufacturing industry to GDP, the development trend of national total energy consumption and energy consumption of individual industries is shown in Fig. 2. Compared with other industries, the manufacturing industry accounts for a very large proportion of the total energy consumption in China, which is maintained at about 57%, and its energy consumption shows a sharp rise. It shows that the energy consumed by the manufacturing industry is not proportional to its contribution to the national GDP, and the consumption structure of the manufacturing energy needs to be further improved.

The ratio between the GDP of various domestic industries and energy consumption and its development trend is shown in Fig. 3. It can be seen from Fig. 3 that both the national ratio and the manufacturing ratio are relatively low. It is precisely because the GDP of the manufacturing industry, mining industry, and other industries are not in direct proportion to the energy consumption so that the national GDP is not in direct proportion to the total energy consumption. For short, although manufacturing contributes a lot to GDP, its energy consumption is larger. Thus it is clear that its GDP contribution is relatively less inefficient.

Since both the gross domestic product and the energy consumption account for a very large proportion in the manufacturing industry in China, the changing trend of various energy consumption in the manufacturing industry, and the energy consumption structure in the manufacturing industry are further considered as shown in Fig. 4. Among them, other energy includes most of the polluting energy such as coal. As can be seen from the Fig. 4, in the energy consumption structure of the manufacturing industry, unclean energy sources such as coal in other energy, coke, crude oil account for a large proportion, while the consumption of clean energy sources, such as natural gas, accounts for a small proportion. Therefore, the energy consumption structure of the manufacturing industry is not in an optimal situation.
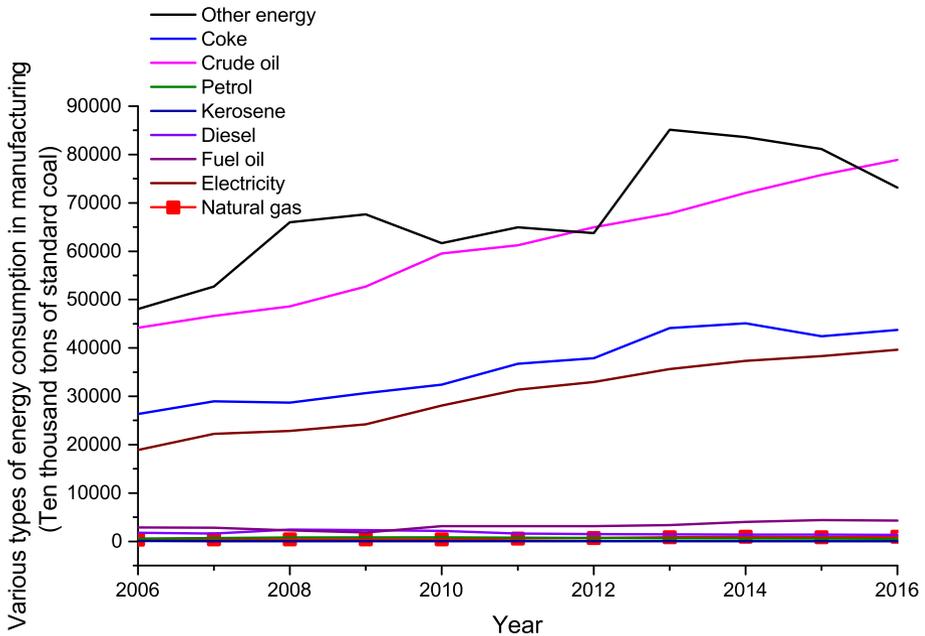
Fig. 4 Various energy consumption trends in the manufacturing industry

3 The grey model with fractional order accumulation

3.1 Definitions of the fractional order accumulation

Definition 1 (*See Ref* Lifeng et al. (2013)) Let $X^{(0)} = \left(x^{(0)}(1), x^{(0)}(2), ..., x^{(0)}(n)\right)$ be an original sequence. The corresponding r-order fractional order accumulation (FOA) $X^{(r)} = \left(x^{(r)}(1), x^{(r)}(2), ..., x^{(r)}(n)\right)$ is defined as

$$x^{(r)}(k) = \sum_{i=1}^{k} \binom{k-i+r-1}{k-i} x^{(0)}(i), k = 1, 2, ..., n, \tag{1}$$

where

$$\binom{k-i+r-1}{k-i} = \frac{(k-i+r-1)(k-i+r-2)...(r+1)r}{(k-i)!}$$

is the general Newton binomial coefficient, and r is the order of the FOA, which is often a non-negative real number. Particularly, $\binom{r-1}{0} = 1, \binom{k-1}{k} = 0, k = 1, 2, ..., n$.

Definition 2 (See Ref Lifeng et al. (2013)) Let $X^{(0)} = \left(x^{(0)}(1), x^{(0)}(2), ..., x^{(0)}(n)\right)$ be an original sequence, where $x^{(0)}(k)$ is the value at time. Then the corresponding r-order fractional order inverse accumulation (IFOA) $X^{(r)} = \left(x^{(r)}(1), x^{(r)}(2), ..., x^{(r)}(n)\right)$ is defined as

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

where

$$\binom{k-i-r-1}{k-i} = \frac{(k-i-r-1)(k-i-r-2)...(-r+1)(-r)}{(k-i)!}.$$

Particularly, $\binom{r-1}{0} = 1$, $\binom{k-1}{k} = 0$, $k = 1, 2, ..., n$.

3.2 The fractional order grey model

Let the r-order accumulation sequence of the non-negative sequence $X^{(0)} = \left(x^{(0)}(1), x^{(0)}(2), ..., x^{(0)}(n)\right)$ be $X^{(r)} = \left(x^{(r)}(1), x^{(r)}(2), ..., x^{(r)}(n)\right)$. In Ref Lifeng et al. (2014), the fractional order additive grey model is represented as the following differential equation:

$$\frac{dx^{(r)}(t)}{dt} + ax^{(r)}(t) = b, \tag{3}$$

which is often called the whitening equation of the FGM. The discrete form is often represented as the following difference equation:

$$x^{(r)}(k) - x^{(r)}(k-1) + az^{(r)}(k) = b, \tag{4}$$

where

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

is called the background value.

Once given the fractional order r, the linear parameters a, b of the FGM are often estimated by the least squares method as

$$[a, b]^T = (B^T B)^{-1} B^T Y, \tag{5}$$

where

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

Set $x^{(r)}(1) = x^{(0)}(1)$, the solution of the Eq. (4) be given by

$$\hat{x}^{(r)}(k) = [x^{(0)}(1) - \frac{b}{a}]e^{-ak} + \frac{b}{a}, k = 1, 2, ..., n. \tag{6}$$

The restored values $\hat{x}^{(0)}$ can be obtained by the r-order IFOA as

$$\hat{x}^{(0)}(k) = \sum_{i=1}^{k} \binom{k-i-r-1}{k-i} \hat{x}^{(r)}(i).$$

4 Time-delayed fractional discrete grey model with multiple fractional order

Let the r_1-order accumulation sequence of the non-negative sequence $X^{(0)} = \left(x^{(0)}(1), x^{(0)}(2), ..., x^{(0)}(n) \right)$ be $X^{(r_1)} = \left(x^{(r_1)}(1), x^{(r_1)}(2), ..., x^{(r_1)}(n) \right)$. Let the r_2-order cumulative sequence of the sequence $N^{(0)} = (1, 2, ..., n)$ be $N^{(r_2)} = \left(1^{(r_2)}, 2^{(r_2)}, ..., n^{(r_2)} \right)$.

Considering the fractional time-delayed effect, the Eq. (3) can be extended to

$$\frac{dx^{(r_1)}(t)}{dt} + ax^{(r_1)}(t) = bt^{(r_2)} + c. \tag{7}$$

The derivative in Eq. (7) can be approximated by

$$\left.\frac{dx^{(r_1)}(t)}{dt}\right|_{t=k} \approx \lim_{\Delta t \to 1} \left.\frac{\Delta x^{(r_1)}(t)}{\Delta t}\right|_{t=k} = \frac{x^{(r_1)}(k+1) - x^{(r_1)}(k)}{(k+1) - k} = x^{(r_1)}(k+1) - x^{(r_1)}(k). \tag{8}$$

Substituting Eq. (8) into Eq. (7), we have

$$x^{(r_1)}(k+1) - x^{(r_1)}(k) + ax^{(r_1)}(k) = bk^{(r_2)} + c,$$

that is

$$x^{(r_1)}(k+1) = (1+a)x^{(r_1)}(k) + bk^{(r_2)} + c.$$

Let $\beta_1 = 1 + a$, $\beta_2 = b$, $\beta_3 = c$; then we get the basic form of TDF-DGM$_M$ as

$$x^{(r_1)}(k+1) = \beta_1 x^{(r_1)}(k) + \beta_2 k^{(r_2)} + \beta_3, k = 2, 3, ..., n-1. \tag{9}$$

Once given the fractional order r_1 and r_2, the linear parameters $\beta_1, \beta_2, \beta_3$ of the TDF-DGM$_M$ can be estimated by the least squares method as

$$[\beta_1, \beta_2, \beta_3]^T = (B^T B)^{-1} B^T Y, \tag{10}$$

where

$$B = \begin{bmatrix} x^{(r_1)}(1) & 1^{(r_2)} & 1 \\ x^{(r_1)}(2) & 2^{(r_2)} & 1 \\ \vdots & \vdots & \vdots \\ x^{(r_1)}(n-1) & (n-1)^{(r_2)} & 1 \end{bmatrix}, Y = \begin{bmatrix} x^{(r_1)}(2) \\ x^{(r_1)}(3) \\ \vdots \\ x^{(r_1)}(n) \end{bmatrix}.$$

Set $\hat{x}^{(r_1)}(1) = x^{(0)}(1)$; by recursively solving the Eq. (9), the discrete response function of TDF-DGM$_M$ can be obtained as

$$\hat{x}^{(r_1)}(k+1) = \hat{\beta}_1^k x^{(0)}(1) + \hat{\beta}_2 \sum_{i=1}^{k} \hat{\beta}_1^{k-i} i^{(r_2)} + \frac{1 - \hat{\beta}_1^k}{1 - \hat{\beta}_1} \hat{\beta}_3, k = 2, 3..., n-1. \tag{11}$$

The restored values $\hat{x}^{(0)}(k)$ can be obtained using the r_1-order IFOA as

$$\hat{x}^{(0)}(k) = \sum_{i=1}^{k} \binom{k - i - r_1 - 1}{k - i} \hat{x}^{(r_1)}(i). \tag{12}$$

The detailed computational processes are summarized in Algorithm 1.

Algorithm 1 The algorithm of TDF-DGM$_M$ to calculate the prediction values.

Require: An original sequence, $X^{(0)}$; A sequence of natural Numbers, $K^{(0)} = (1, 2, ..., n)$;
Ensure: The best fractional order, $r_1 \& r_2$;
1: Calculate the accumulative sequence $X^{(r_1)}$ of the r_1-order of $X^{(0)}$ by the Eq. (1);
2: Calculate the accumulative sequence $K^{(r_2)}$ of the r_2-order of $K^{(0)}$ by the Eq. (1);
3: The least square method was used to estimate the parameters of $[\hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3]^T$ by the Eq. (10);
4: Predict the values of $\hat{X}^{(r_1)} = \left(\hat{x}^{(r_1)}(1), \hat{x}^{(r_1)}(2), ..., \hat{x}^{(r_1)}(n) \right)$ by the Eq. (11);
5: Calculate the restored values of $\hat{X}^{(r_1)}$ with r_1-order by the Eq. (12);
6: **return** the restored values $\hat{X}^{(0)} = \left(\hat{x}^{(0)}(1), \hat{x}^{(0)}(2), ..., \hat{x}^{(0)}(n) \right)$;

5 Relationship and difference between the TDF-DGM$_M$ and FTDGM model

As described above, the proposed TDF-DGM$_M$ is derived from a whitening equation of a grey system using the discrete modeling technique. To further analyze the properties of this model, another similar time-delayed model FTDGM in Ref Ma et al. (2019) is used for theoretical comparison, including the modeling mechanism, unbiasedness, and flexibility.

5.1 Difference in modeling mechanism

For convenience, the modeling details of the FTDGM in Ref Ma et al. (2019) and the proposed TDF-DGM$_M$ are summarized in Table 1.

First, it can be noticed that the TDF-DGM$_M$ is essentially a more general formulation of FTDGM as it can yield FTDGM when $r_1 = r_2$. And this generality will make it more flexible which will be discussed in the last subsection in this section.

Second, the basic form of the FTDGM is obtained by integrating and discretizing the two ends of its whitening equation. However, the basic form of the TDF-DGM$_M$ is obtained by discretizing the derivatives of its whitening equation. This will make the modeling procedures of the TDF-DGM$_M$ easier to implement. As shown in the last second row, the solution of the FTDGM is obtained by solving the whitening equation through the general solution formula of the ordinary differential equation, and its discrete-time response function is obtained from this solution by a numerical formula. However, the solution of the TDF-DGM$_M$ is obtained by recursing its basic form directly, making it more convenient for practical application.

5.2 Difference in unbiasedness

Actually, a general analysis of the unbiasedness of the fractional discrete multivariate grey model has been proved in Ref Ma et al. (2019). Similarly, the analysis method can also be used in this work as the proposed model TDF-DGM$_M$ also used a similar methodology, namely, the discrete modeling technique and fractional order accumulation.

According to Ref Ma et al. (2019), a grey model which is an unbiased model should satisfy the condition that its response function should satisfy its discrete formulation. For FTDGM, there should hold equality when substituting its discrete function to its basic form. However,

Table 1 Comparison of the modeling procedures of TDF-DGM$_M$ the FTDGM

Modeling method	TDF-DGM$_M$ (Discrete modeling)	FTDGM (Continuous modeling)
Whitening equation	$\dfrac{dx^{(r_1)}(t)}{dt} + ax^{(r_1)}(t) = bt^{(r_2)} + c$	$\dfrac{dx^{(r)}(t)}{dt} + ax^{(r)}(t) = bt^{(r)} + c$
Dealing with whitening equation	Differential discrete the derivative of whitening equation $\dfrac{dx^{(r_1)}(t)}{dt}$ to get $x^{(r_1)}(k+1) - x^{(r_1)}(k) + ax^{(r_1)}(k) = bk^{(r_2)} + c$ Be simplified to $x^{(r_1)}(k+1) = (1+a)x^{(r_1)}(k) + bk^{(r_2)} + c$	In interval[k, k+1], integrate both ends of whitening equation to get $\int_k^{k+1} dx^{(r)}(t) + a\int_k^{k+1} x^{(r)}(t)dt = \int_k^{k+1} bt^{(r)}dt + \int_k^{k+1} c\,dt$ Be simplified to $x^{(r)}(k+1) - x^{(r)}(k) + a\int_k^{k+1} x^{(r)}(t)dt = b\int_k^{k+1} t^{(r)}dt + c$
Basic form	Set $\beta_1 = 1+a, \beta_2 = b, \beta_3 = c$ to get Basic form of discrete modeling model $x^{(r_1)}(k+1) = \beta_1 x^{(r_1)}(k) + \beta_2 k^{(r_2)} + \beta_3, k = 1, 2, ..., n-1$	Use Two-point Trapezoidal formula to discretize and simplify to get Basic form of continuous modeling model $x^{(r)}(k+1) - x^{(r)}(k) + az^{(r)}(k) = bm^{(r)}(k) + c, k = 1, 2, ..., n-1$ where $z^{(r)}(k) = \frac{1}{2}[x^{(r)}(k) + x^{(r)}(k+1)], m^{(r)}(k) = \frac{1}{2}[k^{(r)} + (k+1)^{(r)}]$
Parameter estimation	$[\beta_1, \beta_2, \beta_3]^T = (B^T B)^{-1} B^T Y$ where $$B = \begin{bmatrix} x^{(r_1)}(1) & 1^{(r_2)} & 1 \\ x^{(r_1)}(2) & 2^{(r_2)} & 1 \\ \vdots & \vdots & \vdots \\ x^{(r_1)}(n-1) & (n-1)^{(r_2)} & 1 \end{bmatrix}, Y = \begin{bmatrix} x^{(r_1)}(2) \\ x^{(r_1)}(3) \\ \vdots \\ x^{(r_1)}(n) \end{bmatrix}$$	$[a, b, c]^T = (B^T B)^{-1} B^T Y$ where $$B = \begin{bmatrix} -z^{(r)}(2) & m^{(r)}(2) & 1 \\ -z^{(r)}(3) & m^{(r)}(3) & 1 \\ \vdots & \vdots & \vdots \\ -z^{(r)}(n) & m^{(r)}(n) & 1 \end{bmatrix}, Y = \begin{bmatrix} x^{(r)}(2) - x^{(r)}(1) \\ x^{(r)}(3) - x^{(r)}(2) \\ \vdots \\ x^{(r)}(n) - x^{(r)}(n-1) \end{bmatrix}$$
Solution	Recursing from the basic form of discrete modeling to get Discrete-time response function (Discrete Solution) $\hat{x}^{(r_1)}(k+1) = \hat{\beta}_1^k \hat{x}^{(0)}(1) + \hat{\beta}_2 \sum_{i=1}^k \hat{\beta}_1^{k-i} i^{(r_2)} + \frac{1-\hat{\beta}_1^k}{1-\hat{\beta}_1} \hat{\beta}_3$ where $\hat{x}^{(r_1)}(1) = x^{(0)}(1), k = 1, 2, ..., n-1$	Using general solution formula to solve whitening equation to get Continuous-time corresponding function (Continuous Solution) $x^{(r)}(t) = x^{(0)}(1)e^{a(1-t)} + \int_1^t f(s)e^{a(s-t)}ds$ where $f(s) = bs^{(r)} + c$ The discrete form of continuous solution is given by Gauss formula $\hat{x}^{(r)}(k) = x^{(0)}(1)e^{\hat{a}(1-k)} + \sum_{s=1}^{k-1} \frac{1}{2}[\hat{f}(s+1) + \hat{f}(s)]e^{\hat{a}(s-k+\frac{1}{2})}$ where $\hat{f}(s) = \hat{b}s^{(r)} + \hat{c}, k = 1, 2...n$
Restored value	$\hat{x}^{(0)}(k) = \sum_{i=1}^k \binom{k-i-r_1-1}{k-i} \hat{x}^{(r_1)}(i)$	$\hat{x}^{(0)}(k) = \sum_{i=1}^k \binom{k-i-r-1}{k-i} \hat{x}^{(r)}(i)$

the left-side of the FTDGM is actually

$$L_C(k) = x^{(r)}(k+1) - x^{(r)}(k) + az^{(r)}(k)$$
$$= x^{(r)}(k+1) - x^{(r)}(k) + \frac{a}{2}[x^{(r)}(k+1) + x^{(r)}(k)]$$
$$= \frac{a+2}{2}\left(x^{(0)}(1)e^{-ka} + \sum_{s=1}^{k}\frac{1}{2}[f(s+1) + f(s)]e^{a(s-k-\frac{1}{2})}\right)$$
$$+ \frac{a-2}{2}\left(x^{(0)}(1)e^{a(1-k)} + \sum_{s=1}^{k-1}\frac{1}{2}[f(s+1) + f(s)]e^{a(s-k+\frac{1}{2})}\right)$$
$$= \frac{a+2}{2}\left(x^{(0)}(1)e^{-ka} + \sum_{s=1}^{k-1}\frac{1}{2}[b(s+1)^{(r)} + bs^{(r)} + 2c]e^{a(s-k-\frac{1}{2})}\right)$$
$$+ \frac{1}{2}[b(k+1)^{(r)} + bk^{(r)} + 2c]e^{-\frac{1}{2}a}\Big)$$
$$+ \frac{a-2}{2}\left(x^{(0)}(1)e^{-ka}e^{a} + \sum_{s=1}^{k-1}\frac{1}{2}[b(s+1)^{(r)} + bs^{(r)} + 2c]e^{a(s-k+\frac{1}{2})}e^{a}\right)$$
$$= [\frac{a+2}{2} + \frac{a-2}{2}e^{a}]\left(x^{(0)}(1)e^{-ka} + \sum_{s=1}^{k-1}\frac{1}{2}[b(s+1)^{(r)} + bs^{(r)} + 2c]e^{a(s-k-\frac{1}{2})}\right)$$
$$+ \frac{a+2}{4}[b(s+1)^{(r)} + bs^{(r)} + 2c]e^{\frac{a}{2}}$$

And the right-side of the FTDGM basic form in Table 1 is

$$R_C(k) = bm^{(r)}(k) + c$$
$$= \frac{b}{2}[k^{(r)} + (k+1)^{(r)}] + c$$
$$= \frac{1}{2}[bk^{(r)} + b(k+1)^{(r)} + 2c]$$

Obviously, we have

$$L_C(k) \neq R_C(k).$$

And when $|a|$ is small, the discrete response function the FTDGM approximates its basic form. On the contrary, when $|a|$ is larger, the discrete response function is more different from its basic form, which leads to the larger error of the FTDGM. Thus it is obvious that the FTDGM is a biased model.

Similarly, we can also check the unbiasedness of the proposed TDF-DGM$_M$. Substituting discrete solution into the left-side of the TDF-DGM$_M$ basic form in Table 1, there is
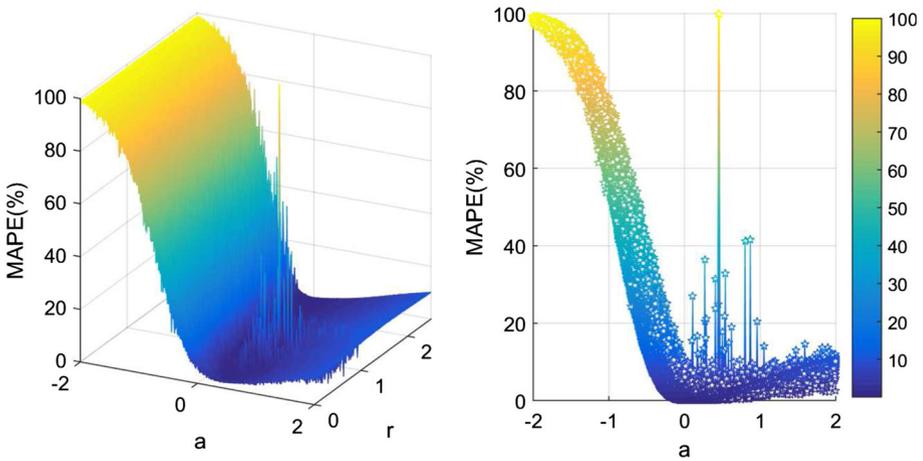
Fig. 5 Testing MAPEs of FTDGM with different values of r and a

$$L_D(k) = x^{(r_1)}(k+1) - \beta_1 x^{(r_1)}(k)$$

$$= \beta_1^k x^{(0)}(1) + \beta_2 \sum_{i=1}^{k} \beta_1^{k-i} i^{(r_2)} + \frac{1 - \beta_1^k}{1 - \beta_1} \beta_3$$

$$- \beta_1 [\beta_1^{k-1} x^{(0)}(1) + \beta_2 \sum_{i=1}^{k-1} \beta_1^{k-i-1} i^{(r_2)} + \frac{1 - \beta_1^{k-1}}{1 - \beta_1} \beta_3]$$

$$= \beta_2 \sum_{i=1}^{k} \beta_1^{k-i} i^{(r_2)} - \beta_2 \sum_{i=1}^{k-1} \beta_1^{k-i} i^{(r_2)} + \frac{1 - \beta_1^k}{1 - \beta_1} \beta_3 - \frac{\beta_1 - \beta_1^k}{1 - \beta_1} \beta_3$$

$$= \beta_2 k^{(r_2)} + \beta_3$$

$$= R_D(k)$$

In short we have:

$$L_D(k) = R_D(k).$$

The above discussions mean that the solution and the basic form of the TDF-DGM$_M$ are equivalent. Thus the TDF-DGM$_M$ is an unbiased model.

For a better explanation, several numerical tests are presented to show the unbiasedness of these two models. For FTDGM, the original series $X^{(0)}$ is generated using its response function in Table 1 as the ideal data. The parameters a is given in the interval[-2,2] by the step of 0.01, and r is given in the interval[0.01,2] by the step of 0.01, respectively. The other parameters b and c are randomly generated in the interval(0,5) by the uniform distribution, and the initial point $x^{(0)}(1)$ is randomly generated in the interval(0,1) by the uniform distribution, respectively. Ten points are generated for each series, in which the first six points are used for modeling, and the rest four points are used for testing. Then the FTDGM models are established based on these ideal data, and the mean absolute percentage error (MAPE) for testing is as shown in Fig. 5.

Figure 5 clearly illustrates the biasedness of the FTDGM. It also clear that when $|a|$ is small, errors of FTDGM are smaller with smaller $|a|$, and they are larger with larger $|a|$.
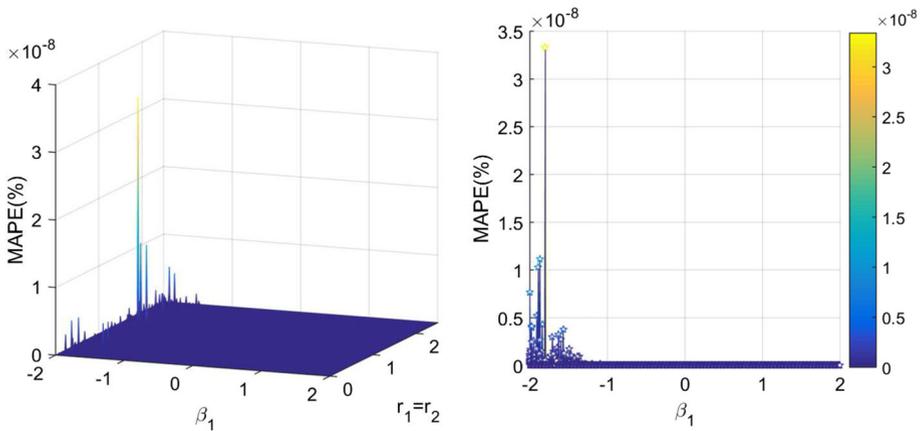
Fig. 6 Testing MAPEs by TDF-DGM$_M$ with different values of $r(= r_1 = r2)$ and β_1

Similar to the above experiment of FTDGM, the series $X^{(0)}$ is generated by the discrete solution of the TDF-DGM$_M$ as ideal data. To make the verification results comparable and more intuitive, the fractional order is set to be equivalent $r = r_1 = r_2$. Then the parameter β_1 is given in the interval[-2,2] by the step of 0.01, and r is given in the interval[0.01,2] by the step of 0.01, respectively. The other parameters β_2 and β_3 are randomly generated in the interval(0,5) by the uniform distribution, and the initial point $x^{(0)}(1)$ is randomly generated in the interval(0,1) by the uniform distribution, respectively. Data scale and divisions for modeling and testing are set to be the same as the above experiment. Then the TDF-DGM$_M$ models ($r = r_1 = r_2$) are established for these ideal data, and the MAPEs for testing are as shown in Fig. 6. It can be clearly seen that all the MAPEs of TDF-DGM$_M$ are smaller than 10^{-8}, which are only truncated errors caused by computer precision. And the parameters do not affect the accuracy of the TDF-DGM$_M$.

5.3 Difference in flexibility

As mentioned above, the multiple fractional order will make the TDF-DGM$_M$ more flexible. This subsection mainly discusses the flexibility of the multivariate fractional order of FTDGM to illustrate the flexibility of the multivariate fractional order of TDF-DGM$_M$. Recalling the analysis in Ref Ma et al. (2019) that the fractional time-delayed term is actually a function more than integer order polynomials, $i. e.$

$$k^{(r)} = \begin{cases} k & r = 0 \\ \frac{1}{2}k(k+1) & r = 1 \\ \frac{1}{6}k(k+1)(k+2) & r = 2 \end{cases} \tag{13}$$

It can be noticed that if the r is completely free, then the form of $k^{(r)}$ can be richer, and this will make the model more flexible. However, the FTDGM uses a unified fractional order for the time-delayed term $k^{(r)}$ and $x^{(r)}$, this makes the variation of the time-delayed term not be a free polynomial, and further limits the flexibility of the FTDGM.

For more intuitive analysis, a simple example is illustrated to show such flexibility with multiple fractional orders of these two models. Figure 7 plots the cases of these two models with unified fractional order and multiple fractional order. Out of interest, we also tried
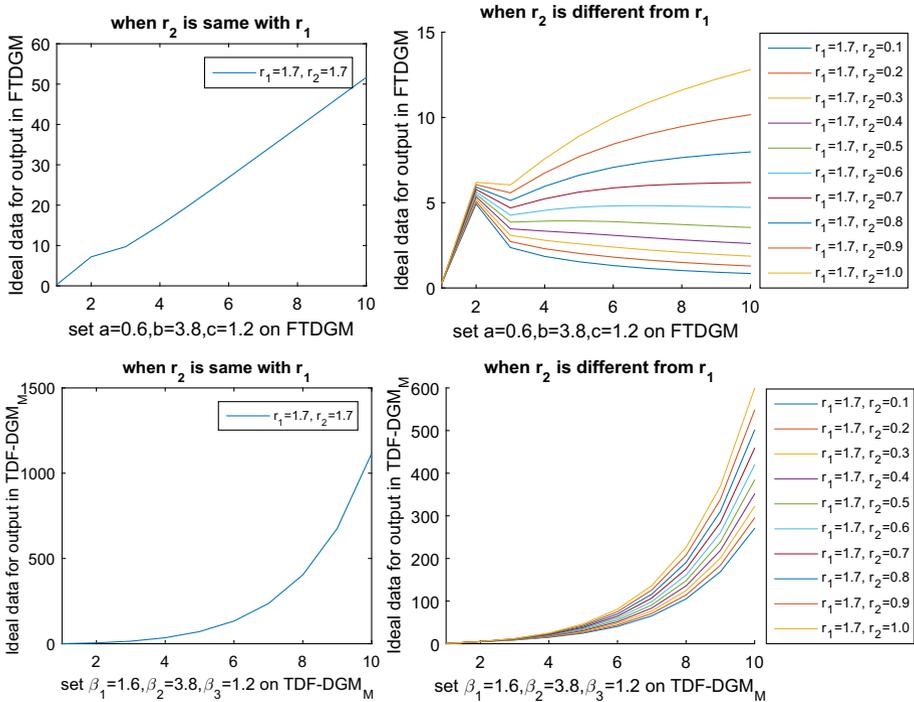
Fig. 7 Output series by FTDGM and TDF-DGM$_M$ when $r_2 = r_1$ and $r_2 \neq r_1$

to make FTDGM with multiple orders in this case. It is clear to see that if the r_2 of the time-delayed term changes the produced curves of both FTDGM and TDF-DGM$_M$ have more shapes, and this property will provide more possibilities for the models to better fit the sample data. Further, it can be easily deduced that this property can make the models more flexible and make them capable to deal with more complex time series.

6 Optimization of the fractional order r_1 and r_2 based on particle swarm optimization

6.1 Formulating the nonlinear optimization problem for r_1 and r_2

The main idea of finding the optimal value of r_1 and r_2 is to minimize the errors of the TDF-DGM$_M$ with independent fractional orders. Generally, we use the MAPE as the main criteria and then the optimization problem for finding the optimal r_1 and r_2 can be formulated as

$$\min J(r_1, r_2) = \frac{1}{V} \sum_{k=1}^{V} \left| \frac{\hat{x}^{(0)}(k) - x^{(0)}(k)}{x^{(0)}(k)} \right| \times 100\%$$

$$s.t. \begin{cases} [\beta_1, \beta_2, \beta_3]^T = (B^T B)^{-1} B^T Y \\ B = \begin{bmatrix} x^{(r_1)}(1) & x^{(r_1)}(2) & \cdots & x^{(r_1)}(n-1) \\ 1^{(r_2)} & 2^{(r_2)} & \cdots & (n-1)^{(r_2)} \\ 1 & 1 & \cdots & 1 \end{bmatrix}^T \\ Y = [x^{(r_1)}(2), x^{(r_1)}(3), \dots, x^{(r_1)}(n)]^T \\ x^{(r_1)}(k+1) = \beta_1 x^{(r_1)}(k) + \beta_2 k^{(r_2)} + \beta_3, k = 1, 2, \dots, n-1 \\ \hat{x}^{(r_1)}(k+1) = \hat{\beta}_1^{k} x^{(0)}(1) + \hat{\beta}_2 \sum_{i=0}^{k-1} \hat{\beta}_1^{i} (k-i)^{(r_2)} \\ + \frac{1-\hat{\beta}_1^{k}}{1-\hat{\beta}_1} \hat{\beta}_3, k = 2, 3, \dots, n-1 \\ \hat{x}^{(0)}(k) = \sum_{i=1}^{k} C_{k-i-r_1-1}^{k-i} \hat{x}^{(r_1)}(i), k = 1, 2, \dots, n \end{cases} \quad (14)$$

where V represents the number of data points used for estimating parameters β_1, β_2, and β_3, that is used for modeling.

It can be seen that the objective function is a nonlinear function of r_1 and r_2, and there exist several nonlinear constraints; thus this optimization problem is essentially nonlinear programming. The explicit expression of the objective function and the constraints are very complex, and thus it cannot be solved analytically. It should be noticed that such formulation is often used for the optimization of the existing nonlinear grey models in Ref Pei et al. (2018), Zheng-Xin (2014).

6.2 Solving the nonlinear programming using particle swarm optimization

Particle swarm optimization is a population-based stochastic optimization technique developed by Eberhart and Kennedy in 1995, inspired by the social behavior of bird flocking or fish schooling. In the past several years, the PSO has been successfully applied in many research and application areas. It is demonstrated that the PSO gets better results in a faster and cheaper way compared with other methods. Another reason for choosing PSO is that there are few parameters to adjust. One version with slight variations works well in a wide variety of applications.

In each iteration, the particle updates its speed and position through the individual extremum and the group extremum. The change of them is defined as

$$B = \begin{cases} V_{id}^{k+1} + c_1 r_1 (P_{id}^k - X_{id}^k) + c_2 r_2 (P_{gd}^k - X_{id}^k) \\ X_{id}^{k+1} = X_{id}^k + V_{id}^{k+1} \end{cases} \quad (15)$$

where ω is the weight of inertia, $d = 1, 2, \dots, D$, $i = 1, 2, \dots, n$, k is the current iteration number, V_{id} is the speed of the particle, c_1 and c_2 are non-negative constant which was called the acceleration factor, and set to $c_1 = c_2 = 2$ normally, the random number of r_1 and r_2 distribution in the interval[-2,2]. To prevent the blind searching of particles, it is generally recommended to limit their position and speed to a certain range $[-X_{max}, X_{max}]$, $[-V_{max}, -V_{max}]$.

The overall calculation steps of the TDF-DGM$_M$ model based on the PSO can be briefly summarized in Fig. 8.
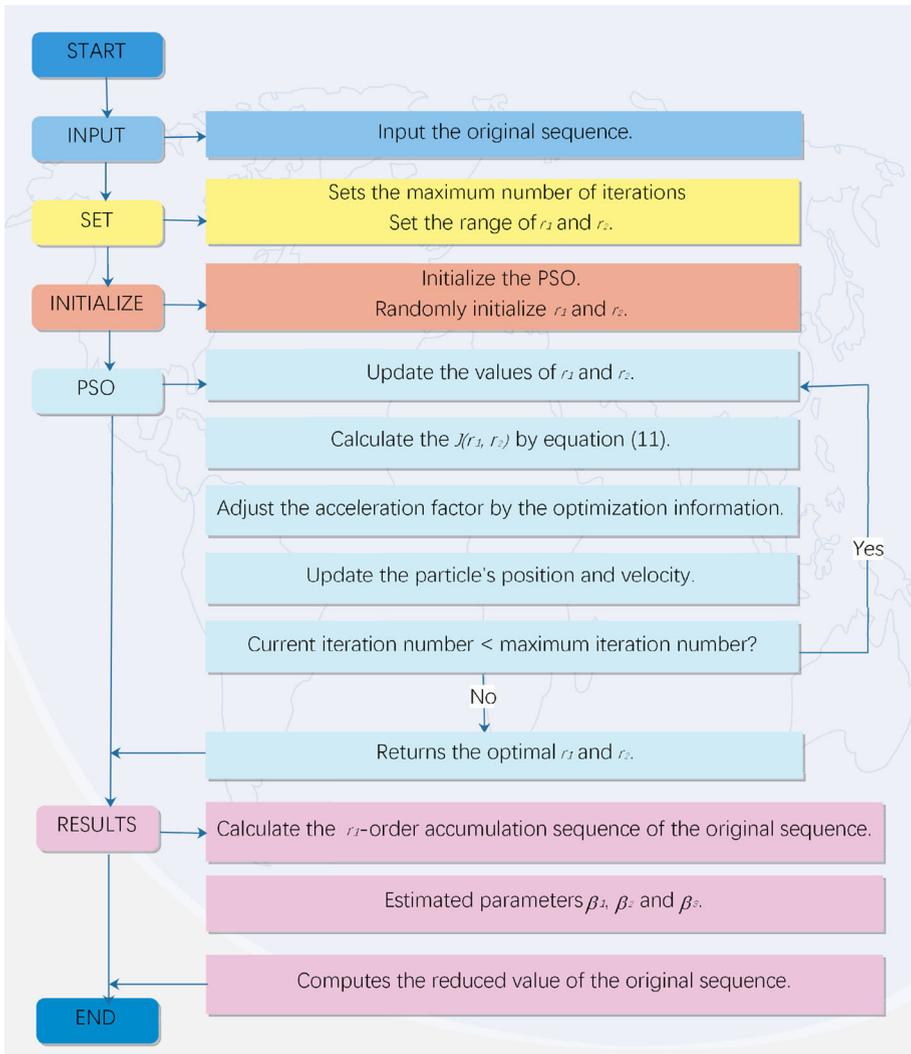
Fig. 8 The flowchart of TDF-DGM$_M$ model based on the PSO

7 Validation

In this paper, the parameter optimization problem has changed from the one-dimensional nonlinear programming problem of FGM to the multi-dimensional nonlinear programming problem of TDF-DGM$_M$, which makes parameter optimization more difficult. So we used four cases to verify the validity and accuracy of the TDF-DGM$_M$. For all cases, the PSO will be compared with Grey Wolf Optimizer (GWO) and Genetic Algorithm (GA), and the TDF-DGM$_M$ will be compared to other grey models, including the FTDGM, the time-delayed fractional discrete grey model with unique fractional order (TDFDGM$_U$), the fractional non-homogeneous discrete grey model (FNDGM), the fractional discrete grey model (FDGM), the GM(1,1), the nonhomogeneous discrete grey model (NDGM), and the discrete grey model

Table 2 Raw data and relevant information of the four validation cases

NO.	Raw data	Source	Modeling points	Predicted points
Case 1	247.839, 273.021, 289.014, 285.208, 288.818, 297.078, 293.662, 290.404, 279.143	In Ref Lifeng (2015)	8	1
Case 2	4998.00, 5309.01, 6029.88, 6510.94, 7182.10, 7942.88, 8696.55, 9997.47, 11242.85, 12264.55, 13471.50, 15160.90, 16674.30	In Ref Xin (2016)	6	7
Case 3	809.340, 867.550, 965.830, 1099, 1291.402, 1501.924, 1734.832, 2013.674, 2095.019, 2343.846, 2691.520, 2984.904	In Ref Han (2014)	6	6
Case 4	15.93, 18.38, 21.35, 23.60, 25.87, 28.00, 29.14, 30.66, 32.50, 34.80	In Ref Shiquan et al. (2014)	8	2
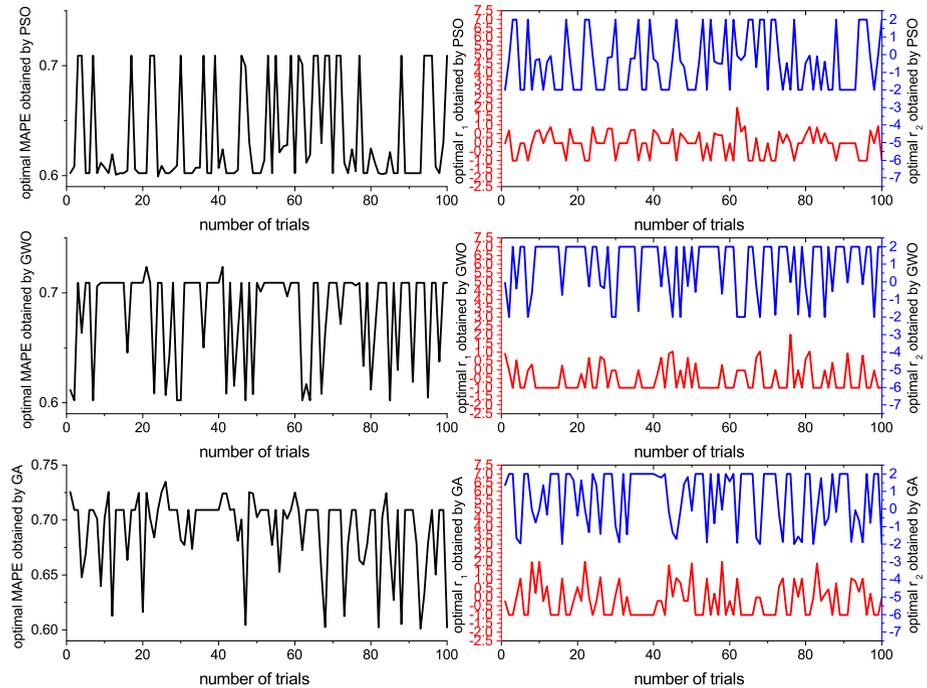
Fig. 9 The optimal parameters and MAPE of TDF-DGM$_M$ in each trial of Case 1

(DGM). Its data sources are shown in Table 2. The population size of PSO is set at 30, the maximum number of iterations is taken as the stop condition and set as 100 times, and each experiment is repeated 100 times. The GWO and GA parameter settings are the same as PSO. All the calculations have been done in Matlab 2015a.

To compare the performance of the PSO, GWO, and GA, the minimum MAPE and the corresponding r_1 and r_2 of the TDF-DGM$_M$ among the 100 trials obtained by the three

Table 3 For the four validation cases, the minimum MAPE and the corresponding r_1 and r_2, and the meantime in the 100 experiments of TDF-DGM$_M$ with PSO, GWO, and GA

NO.	PSO				GWO				GA			
	MAPE(%)	r_1	r_2	Meantime	MAPE(%)	r_1	r_2	Meantime	MAPE(%)	r_1	r_2	Meantime
Case 1	**0.5993**	0.8859	-0.0747	**7.9375**	0.6023	-0.0273	-2	8.0967	0.6010	0.8968	-0.0654	8.9481
Case 2	**0.0000**	-0.1511	1.6030	**12.6497**	0.0002	-0.1510	1.6020	12.6696	0.0009	-0.1508	1.5979	12.9006
Case 3	**0.0002**	-0.0075	1.2171	**10.9856**	0.1541	-0.0084	1.1493	11.0254	0.1660	-1.0112	-1.9419	11.2473
Case 4	**0.2952825**	-1.1987	0.2851	**9.0092**	0.2952916	-1.1998	0.3074	9.0365	0.2952826	-1.1986	0.2842	9.2792
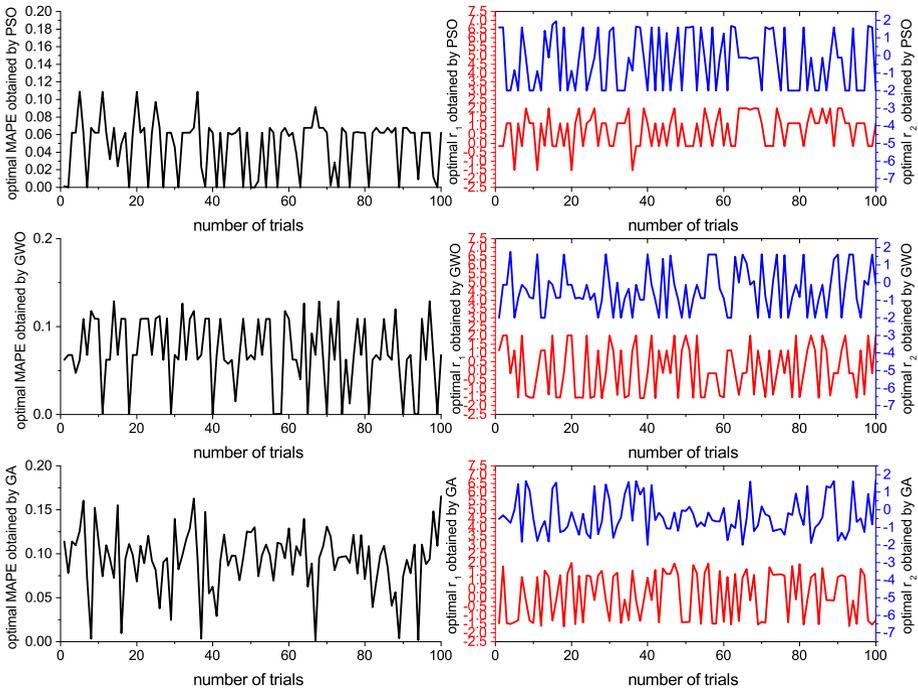
Fig. 10 The optimal parameters and MAPE of TDF-DGM$_M$ in each trial of Case 2
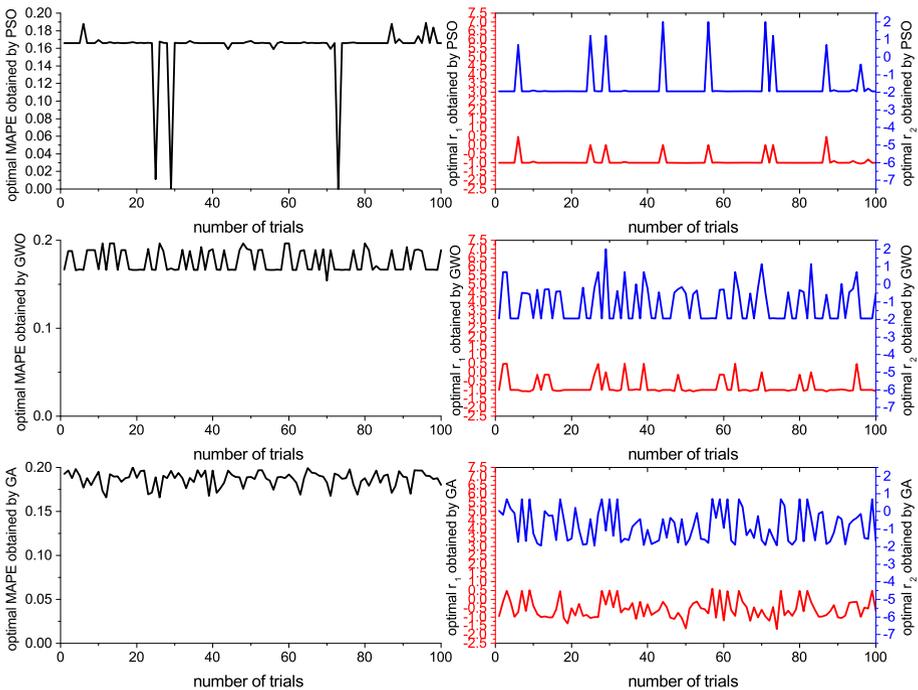
Fig. 11 The optimal parameters and MAPE of TDF-DGM$_M$ in each trial of Case 3
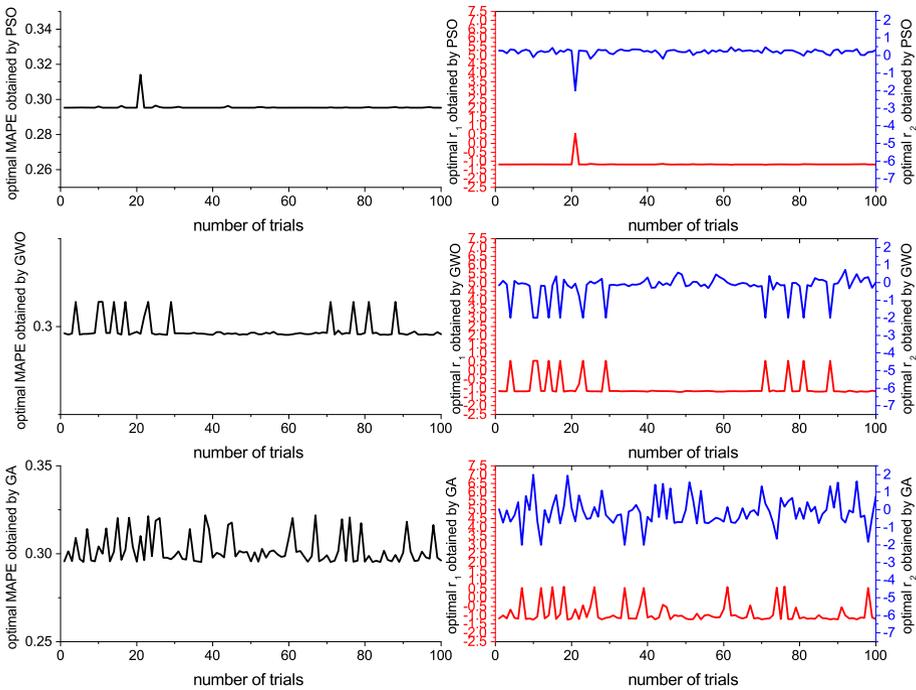
Fig. 12 The optimal parameters and MAPE of TDF-DGM$_M$ in each trial of Case 4

algorithms are presented in Table 3. Meanwhile, the average time of 100 experiments is also given in the table. Besides, the optimal parameters and MAPE in each trial are shown in Figs. 9, 10, 11, and 12.

From Table 3, it can be seen that the PSO can achieve the best optimization effect and smaller objective function value in the same 100 trials, which shows that its convergence is better than the GWO and GA. In the four cases, the average test time of PSO is shorter, which shows that its convergence speed is faster than the GWO and GA. It can be seen from Figs. 10, 11, and 12 that the PSO has a bit slightly higher stability than GWO and GA in Case 2, Case 3, and Case 4. It is interesting to see in Fig. 9 that fluctuations of MAPE by PSO, GWO, and GA are similar, but that of PSO is more stable with smaller MAPE values.

Also, a set of optimization results in 100 trials is shown in Fig. 13. According to the experimental results from Fig. 13, the PSO has the fastest convergence speed and best convergence comparing with GWO and GA. This shows that the PSO needs fewer iterations to converge to the optimal value.

Above all, the PSO is finally selected to optimize the r_1 and r_2 of the above four cases. To make the grey models comparative, the PSO is used to optimize the parameters of the above eight grey models. The initial population size is set as 30, the stop criteria is set as 10^{-6}, and the maximum number of iteration is set as 500 times. Among them, the processes of r_1 and r_2 for optimizing TDF-DGM$_M$ parameters by PSO are shown in Fig. 14. Then we calculated the MAPEs of fitting and prediction of the eight models, and the MAPEs of fitting and prediction of each case were obtained, as shown in Table 4.

According to Table 4, except for the fitting error of case4, the MAPEs of fitting and prediction of the discrete grey model with the independent fractional time-delayed term is
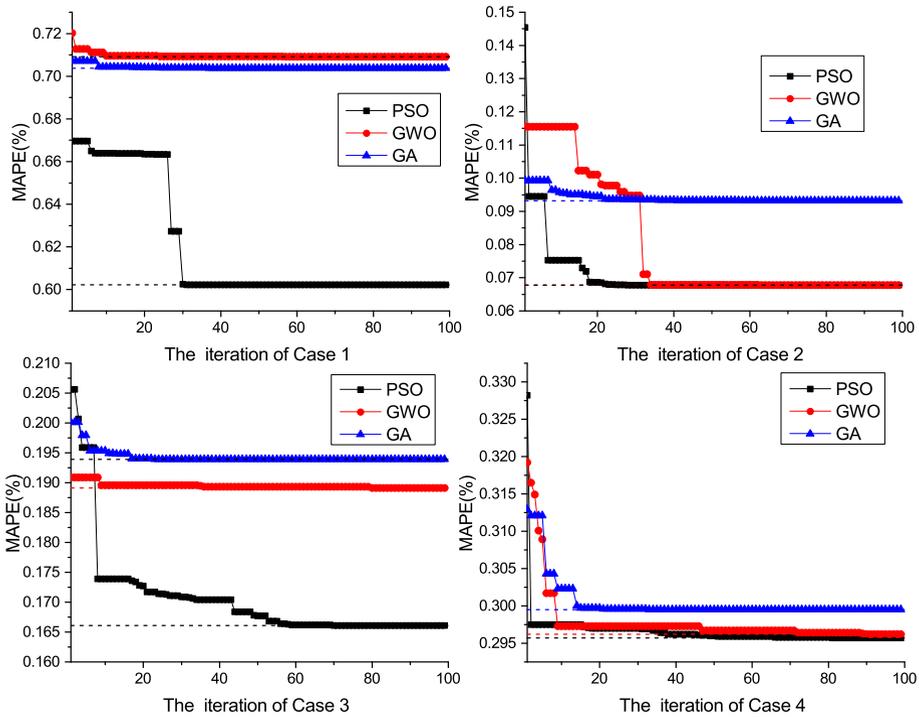
Fig. 13 The change of MAPE of TDF-DGM$_M$ in the process of PSO, GWO, and GA

lower than that of the other models in the above four cases. So the discrete grey model with the independent fractional time-delayed term is more appropriate for the four cases.

8 Application in forecasting natural gas consumption of the manufacturing industry of China

The raw data are collected from the statistics of energy consumption by industry in China's statistical yearbook in the range of 2007–2017 (http://www.stats.gov.cn/tjsj/ndsj/).

Table 5 indicates that the consumption of natural gas in the national manufacturing industry is increasing year by year. The data from 2006 to 2010 will be used to build the models, and the data from 2011 to 2015 will be used to test their out-of-sample performance. The minimum MAPE and the corresponding r_1 and r_2 of the TDF-DGM$_M$ among the 100 trials obtained by the three algorithms are presented in Table 6. Also, the average time of 100 experiments is given in the table.

It can be seen in Table 6 that the PSO has higher accuracy than GWO and GA, and the PSO has faster convergence speed. The optimized parameters of TDF-DGM$_M$, FTDGM, TDFDGM$_U$, FNDGM, FDGM by PSO are r_1=2, r_2=-0.8679, r=-0.1156, r=-0.3203, r=-2, r=0.8565, respectively. The fitting and prediction results of the eight models are shown in Table 7. And the absolute value of fitting error and prediction error of the eight models are shown in Fig. 15.
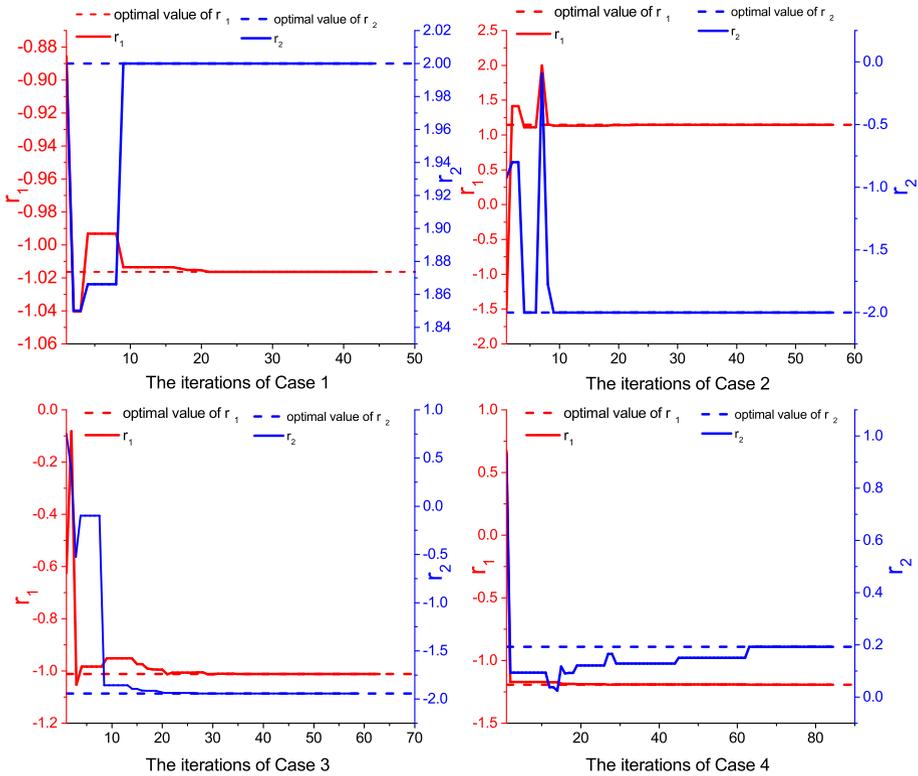
Fig. 14 The change of r_1 and r_2 of TDF-DGM$_M$ in the process of PSO

Table 7 shows that the MAPE of TDF-DGM$_M$ is smaller than that of the others. Figure 15 shows that the errors of the TDF-DGM$_M$ are better than that of the others. So the TDF-DGM$_M$ is more appropriate for forecasting the data of national manufacturing gas consumption. Thus, the fitting and prediction results of the eight models are plotted in Fig. 15.

Figure 16 further illustrates the details of the fitting effect and prediction performance of these models. It is clear to see that the predicted values of TDF-DGM$_M$ are much closer to the raw data, while most other models failed to catch the overall trend of the testing values. It is also very interesting to see that all these models perform quite well in fitting, especially the fitting errors of TDFDGM$_U$ are smaller than $1e - 2\%$. Thus it is obvious that these models have over-fitted the sample data. On the contrast, this further presents the higher generality of the TDF-DGM$_M$.

9 Conclusions

In this paper, a novel time-delayed fractional grey model with multiple fractional order, abbreviated as TDF-DGM$_M$, was proposed and the PSO algorithm was employed to select its optimal values of two independent fractional orders. Results of the numerical validation with four real-world data sets were used to show the effectiveness of PSO and the priority of TDF-DGM$_M$ over the seven existing grey models.

Table 4 The MAPEs obtained by fitting and predicting four validation cases with eight models, respectively

NO.	TDF-DGM$_M$		FTDGM		TDFDGM$_U$		FNDGM		FDGM		GM(1,1)		NDGM		DGM	
	Fit	Pre	Fit	Pre	Fit	Pre	Fit	Pre	Fit	Pre	Fit	Pre	Fit	Pre	Fit	Pre
Case 1	**0.71**	**1.03**	0.81	2.73	0.78	2.25	0.73	6.22	0.81	3.39	1.17	7.00	0.84	4.79	1.17	6.98
Case 2	**0.06**	**2.68**	1.58	18.41	0.49	14.31	0.13	14.24	0.60	6.76	0.59	4.34	0.64	8.96	0.61	4.25
Case 3	**0.17**	**3.74**	0.33	16.66	0.19	8.41	0.19	11.72	0.24	18.75	1.07	8.71	0.43	32.15	1.07	8.88
Case 4	**0.2953**	4.38	0.52	5.52	0.30	**3.52**	0.2957	4.29	0.33	3.59	2.49	5.73	0.35	3.97	2.50	5.76

Table 5 The raw data of national manufacturing gas consumption (Billion cubic meters)

Year	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
National manufacturing gas consumption	263.12	309.08	337.92	321.14	357.7	483.07	572.72	715.74	781.92	718.63

Table 6 For the natural gas in the national manufacturing industry of China, the minimum MAPE and the corresponding r_1 and r_2, and the meantime in the 100 experiments of TDF-DGM$_M$ with PSO, GWO, and GA

NO.	PSO MAPE(%)	r_1	r_2	Meantime	GWO MAPE(%)	r_1	r_2	Meantime	GA MAPE(%)	r_1	r_2	Meantime
National manufacturing gas consumption	**1.97 × 10^{-12}**	1.5935	-2	**7.7435**	6.36×10^{-7}	2	-0.8679	7.7543	1.58×10^{-7}	1.6592	-1.7678	8.0356

Table 7 The fitting and prediction results of forecasting the natural gas consumption of the manufacturing industry of China by the eight models

Raw data		TDF-DGM$_M$		FTDGM		TDFDGM$_U$		FNDGM		FDGM		GM(1,1)		NDGM		DGM	
Year	Data	Data	E(%)	Data	E(%)	Data	E(%)	Data	E(%)	Data	E(%)	Data	E(%)	Data	E(%)	Data	E(%)
2006	263.12	263.12	**0.00**	263.12	**0.00**	263.12	**0.00**	263.12	**0.00**	263.12	**0.00**	263.12	**0.00**	263.12	**0.00**	263.12	**0.00**
2007	309.08	309.08	**0.00**	298.55	3.41	309.08	**0.00**	304.81	1.38	309.08	**−0.00**	312.22	−1.02	304.46	1.49	312.50	−1.11
2008	337.92	337.92	**−0.00**	322.15	4.67	337.92	**−0.00**	326.49	3.38	328.32	2.84	324.68	3.92	345.88	−2.35	324.82	3.88
2009	321.14	321.14	**0.00**	341.13	−6.22	321.14	**0.00**	324.56	−1.06	340.42	−6.00	337.63	−5.14	327.58	−2.01	337.62	−5.13
2010	357.70	357.70	**−0.00**	357.70	**−0.00**	357.70	**−0.00**	356.68	0.29	348.80	2.49	351.10	1.84	335.66	6.16	350.93	1.89
MAPE			**0.00**		2.86		**0.00**		1.22		2.27		2.38		2.40		2.40
2011	483.07	418.51	13.37	372.79	22.83	267.05	44.72	462.22	**4.32**	354.84	26.54	365.11	24.42	332.09	31.25	364.76	24.49
2012	572.72	499.67	**12.75**	386.88	32.45	412.69	27.94	686.01	−19.78	359.28	37.27	379.68	33.71	333.67	41.74	379.14	33.80
2013	715.74	602.41	**15.83**	400.25	44.08	81.08	88.67	1071.25	−49.67	362.55	49.35	394.83	44.84	332.97	53.48	394.08	44.94
2014	781.92	730.03	**6.64**	413.07	47.17	683.29	12.61	1661.63	−112.51	364.91	53.33	410.58	47.49	333.28	57.38	409.61	47.61
2015	718.63	887.28	**−23.47**	425.45	40.80	−564.50	178.55	2500.68	−247.98	366.57	48.99	426.96	40.59	333.15	53.64	425.76	40.75
MAPE			**14.41**		37.46		70.50		86.85		43.10		38.21		47.50		38.32
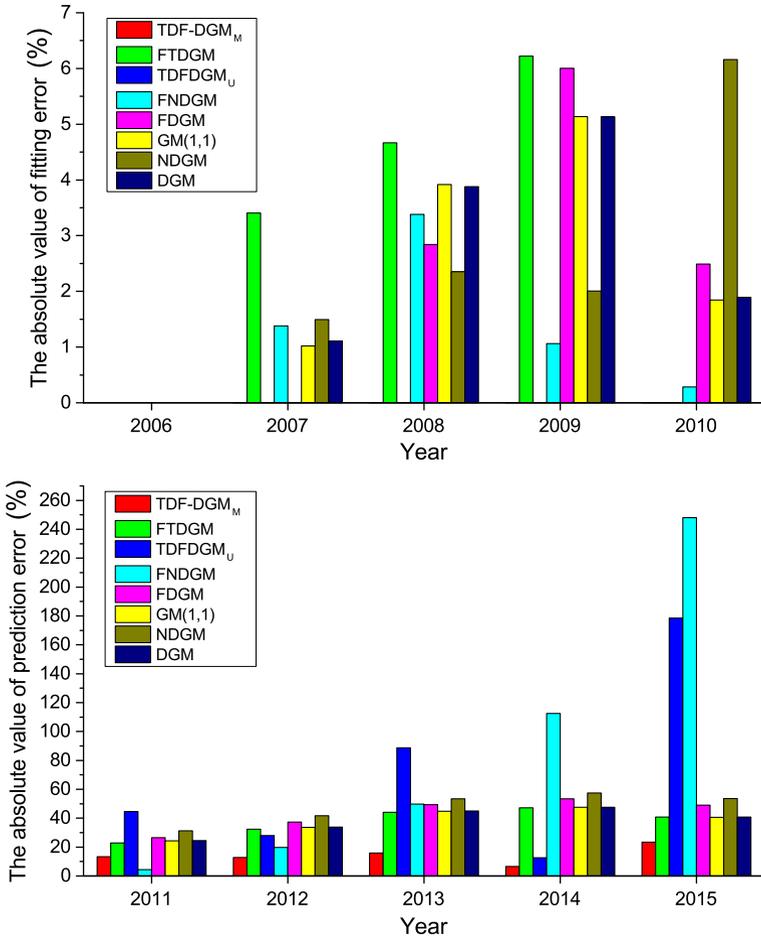
Fig. 15 The absolute value of fitting error and prediction error of the eight models

Real-world application of forecasting the natural gas consumption of the manufacturing industry of China was executed with real-world data. The results showed that the proposed TDF-DGM$_M$ model was significantly better than the other seven existing models. And it is also very interesting to see that the TDF-DGM$_M$ was also more effective than its special form TDFDGM$_U$ with unified fractional order. Further the results obtained in this paper illustrated that the TDF-DGM$_M$ was eligible to forecast the natural gas consumption of the manufacturing industry of China.

What's more, the methodology used to build the TDF-DGM$_M$ model can also be regarded as a new way of the fractional grey modeling technique, which can be expected to build more fractional grey models with higher accuracy in the future.
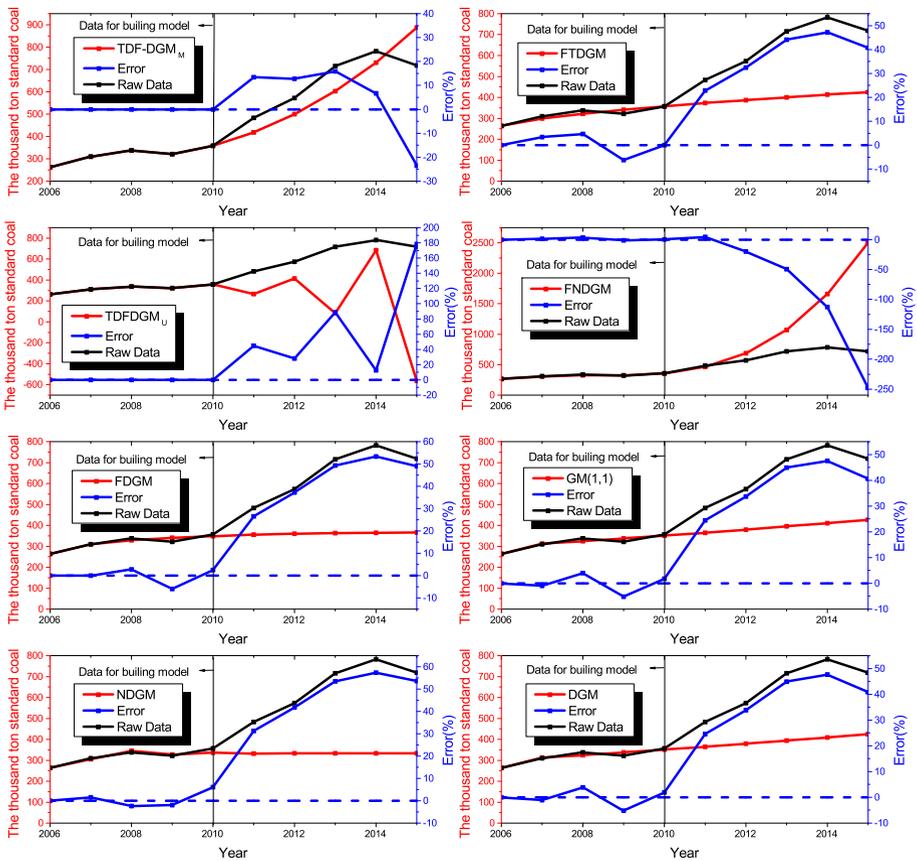
Fig. 16 Plots of fitting and forecasting values of the natural gas consumption of the manufacturing industry of China by the eight models

References

Benítez Rafael Bernardo Carmona, Paredes Rafael Bernardo Carmona, Lodewijks Gabriel, Nabais Joao Lemos (2013) Damp trend grey model forecasting method for airline industry. Expert Syst Appl 40(12):4915–4921

Bezuglov Anton, Comert Gurcan (2016) Short-term freeway traffic parameter prediction: application of grey system theory models. Expert Syst Appl 62:284–292

Duman Gazi Murat, Kongar Elif, Gupta Surendra M (2019) Estimation of electronic waste using optimized multivariate grey models. Waste Manag 95:241–249

Ene Seval, Öztürk Nursel (2017) Grey modelling based forecasting system for return flow of end-of-life vehicles. Technol Forecasting Soc Change 115:155–166

Fan Junliang, Lifeng Wu, Zhang Fucang, Cai Huanjie, Ma Xin, Bai Hua (2019) Evaluation and development of empirical models for estimating daily and monthly mean daily diffuse horizontal solar radiation for different climatic regions of china. Renew Sustain Energy Rev 105:168–186

Gatabazi P, Mba JC, Pindza E (2019) Modeling cryptocurrencies transaction counts using variable-order fractional grey lotka-volterra dynamical system. Chaos Solitons Fractals 127:283–290

Han Xie (2014) Electric power consumption forecasting based on gray neural network model—taking hebei province as all example. J Baoding Univ 27(2):55–60

Julong Deng (1986) Main methods of intrinsic gray system. Syst Eng Theory Practice 01:60–65

Kayacan Erdal, Ulutas Baris, Kaynak Okyay (2010) Grey system theory-based models in time series prediction. Expert Syst Appl 37(2):1784–1789

Lifeng, Wu (2015) *Fractional order grey forecasting models and their application*. PhD thesis, Nanjing university of aeronautics and astronautics

Lifeng Wu, Bin Fu (2017) Gm(11) model with fractional order opposite direction accumulated generation and its properties. Stat Decision-Making 18:33–36

Lifeng Wu, Liu Sifeng, Yao Ligen, Yan Shuli, Liu Dinglin (2013) Grey system model with the fractional order accumulation. Commun Nonlinear Sci Numer Simul 18(7):1775–1785

Lifeng Wu, Liu Sifeng, Chen Ding, Yao Ligen, Cui Wei (2014) Using gray model with fractional order accumulation to predict gas emission. Natu Hazards 71(3):2231–2236

Lifeng Wu, Liu Sifeng, Fang Zhigeng, Haiyan Xu (2015) Properties of the gm (1, 1) with fractional order accumulation. Appl Math Comput 252:287–293

Lifeng Wu, Li Nu, Yang Yingjie (2018) Prediction of air quality indicators for the beijing-tianjin-hebei region. J Clean Prod 196:682–687

Ma Xin, Mei Xie, Wenqing Wu, Xinxing Wu, Zeng Bo (2019) A novel fractional time delayed grey model with grey wolf optimizer and its applications in forecasting the natural gas and coal consumption in chongqing china. Energy 178:487–507

Ma Xin, Xie Mei, Wenqing Wu, Zeng Bo, Wang Yong, Xinxing Wu (2019) The novel fractional discrete multivariate grey system model and its applications. Appl Math Model 70:402–424

Moonchai Sompop, Chutsagulprom Nawinda (2020) Short-term forecasting of renewable energy consumption: augmentation of a modified grey model with a kalman filter. Appl Soft Comput 87:105994

National bureau of statistics. *Statistical yearbook of China*. Beijing: China statistical publishing house, 2002-2007

Pei Lingling, Li Qin, Wang Zhengxin (2018) The nls-based nonlinear grey bernoulli model with an application to employee demand prediction of high-tech enterprises in china. Theory and Application, Grey Systems

Pei Du, Wang Jianzhou, Yang Wendong, Niu Tong (2019) A novel hybrid model for short-term wind power forecasting. Appl Soft Comput 80:93–106

Şahin, Utkucan (2020) Projections of turkey's electricity generation and installed capacity from total renewable and hydro energy using fractional nonlinear grey bernoulli model and its reduced forms. *Sustainable Production and Consumption*

San Cristóbal José Ramón, Correa Francisco, González María Antonia, Diaz Ruiz de Navamuel Emma, Madariaga Ernesto, Ortega Andrés, López Sergio, Trueba Manuel (2015) A residual grey prediction model for predicting s-curves in projects. Proc Comput Sci 64:586–593

Shaikh Faheemullah, Ji Qiang, Shaikh Pervez Hameed, Mirjat Nayyar Hussain, Uqaili Muhammad Aslam (2017) Forecasting china s natural gas demand based on optimised nonlinear grey models. Energy 140:941–951

Shiquan Jiang, Sifeng Liu, Xingcai Zhou (2014) Optimization of background value in gm(1,1) based on compound trapezoid formula. Control Decision 29(12):2221–2225

Utkucan, Şahin, Tezcan, Şahin (2020) Forecasting the cumulative number of confirmed cases of covid-19 in italy, uk and usa using fractional nonlinear grey bernoulli model. *Chaos, Solitons & Fractals*, page 109948

Wang Jianzhou, Pei Du, Haiyan Lu, Yang Wendong, Niu Tong (2018) An improved grey model optimized by multi-objective ant lion optimization algorithm for annual electricity consumption forecasting. Appl Soft Comput 72:321–337

Wang Yong, Tian Donghong, Li Guofeng, Zhang Chen, Chen Tao (2019) Dynamic analysis of a fractured vertical well in a triple media carbonate reservoir. Chem Technol Fuels Oils 55(1):56–65

Xiao Xinping, Duan Huiming, Wen Jianghui (2020) A novel car-following inertia grey model and its application in forecasting short-term traffic flow. Appl Math Model

Xin, Ma (2016) *Study on dynamical prediction methods based on gray system and kernel method*. PhD thesis, Chengdu: Southwest Petroleum University

Xiong Ping-ping, Huang Shen, Peng Mao, Xiang-hua Wu (2020) Examination and prediction of fog and haze pollution using a multi-variable grey model based on interval number sequences. Appl Math Model 77:1531–1544

Yang Wendong, Wang Jianzhou, Haiyan Lu, Niu Tong, Pei Du (2019) Hybrid wind energy forecasting and analysis system based on divide and conquer scheme: a case study in china. J Clean Prod 222:942–959

Zeng Bo, Ma Xin, Shi Juanjuan (2020) Modeling method of the grey gm (1, 1) model with interval grey action quantity and its application. Complexity 2020:

Zeng Bo, Zhou Meng, Liu Xianzhou, Zhang Zhiwei (2020) Application of a new grey prediction model and grey average weakening buffer operator to forecast chinas shale gas output. Energy Rep 6:1608–1618

Zheng-Xin Wang (2014) Nonlinear grey prediction model with convolution integral ngmc (1, n) and its application to the forecasting of china's industrial so2 emissions. Journal of Applied Mathematics 2014: