An optimized grey model for predicting non-renewable energy consumption in China

Jianlong Guo [a,b], Lifeng Wu [c], Yali Mu [a,b,*]

[a] College of Economics and Management, Nanjing Forestry University, Nanjing, 210037, China
[b] Research Center for Economics and Trade in Forest Products of the State Forestry Administration (SINO-RCETFOR), Nanjing, 210037, China
[c] School of Management Engineering and Business, Hebei University of Engineering, Handan, 056038, China

A R T I C L E   I N F O

A B S T R A C T

The large amount of the non-renewable energy consumption in China brings certain challenges to the realization of carbon neutrality. This paper proposes a new grey model to predict the consumption of non-renewable energy in China. Based on the traditional grey model, the proposed model introduces two parameters to adjust the weight of information. Simultaneously, the intelligent optimization algorithm determines the optimal parameters. Three cases verify the feasibility of the model. The forecast results show that the amount of oil and natural gas consumption will continue to grow at a faster rate. By 2026, the amount of oil consumption will exceed 37 EJ (EJ) and natural gas consumption will exceed 22 EJ. Compared to 2021, oil consumption is up nearly 24%, and natural gas consumption is up more than 60%. While the consumption of coal will maintain a small up rate and gradually be leveled off.

1. Introduction

Natural resources play an important role in social development. Non-renewable resources are closely related to people's daily life [1]. The non-renewable resources that humans are most exposed to in their daily lives are coal, oil and natural gas. The massive consumption of these resources has caused very serious problems to the environment on which people live [2]. At the same time, if the consumption of these resources is not controlled, they will be depleted. It impedes sustainable development. In recent years, the focus on carbon neutrality is high, and China has an important role to play in achieving carbon neutrality [3,4]. The realization of carbon neutrality can help China adjust its energy structure and protect the environment. The massive consumption is bound to prevent China from becoming carbon neutral [5]. It is a general trend to vigorously develop clean energy for replacing non-renewable energy. Therefore, this thesis mainly takes China as the research area and non-renewable energy consumption as the research object. The future development trend is forecasted and analyzed. And aim to provide reference for relevant departments to issue policies.

The grey system was proposed in the 1980s [6], the theory has been developed by many scholars. The theory is based on differential equations, with little data information to model and predict. Compared with other forecasting methods, such as neural network algorithm [7–9], long and short memory network method [10,11], linear regression method [12–14] and other models which require a lot of statistical data. Grey model can make accurate predictions with less data. With this characteristic, it has been widely used in air quality [15,16], water resources [17,18], transportation [19,20], the energy consumption [21,22], etc.

However, as a subject which has only been developed for more than 30 years, it still needs further optimization and adjustment. By

---

studying literatures, the improvement methods are mainly divided into the following parts. The first aspect is optimization of background values. The traditional grey model is $W^{(1)}(k) = 0.5 * [m^{(1)}(k+1) + m^{(1)}(k)]$. A new optimization method, $W^{(1)}(k) = \frac{m^{(1)}(k) - m^{(1)}(k-1)}{\ln m^{(1)}(k) - \ln m^{(1)}(k-1)}$ is proposed, which can better reduce the error [23]. Some scholars suggest that the traditional method will cause a large error to the result, so a regulating factor is introduced in the calculation process. $M^{(1)}(k) = \alpha y^{(1)}(k) + (1 - \alpha)$ is proposed by Wang [24]. The second aspect is the improvement of generating operator. It can skillfully deal with the unsmoothness in the sequence. So, the data series can be further optimized and the prediction accuracy can be improved. In contrast, the sequence development law after accumulation is more clear. However, does first-order accumulation make the sequence fit best? Based on this problem, Wu proposes a fractional order generating operator, which gives different orders according to different wave conditions of the sequence [25]. Some scholars believe that different information should be given different weights [26]. There are also other operators, such as the weakening buffer operator [27], damping accumulation generating operator [28], etc. Another aspect is the optimization of error correction. When the result of model prediction is not good, it is usually necessary to correct the error. Usually, the traditional model predicts the error again after predicting the result. Jia believes that this is inappropriate. Therefore, he uses Markov method instead of grey model to correct the error [29]. Some scholars also use the triangular residual correction method to reduce the uncertainty of the sequence [30].

Reasonable accumulation method makes the original sequence smoother, so that makes the prediction better. The accumulation method combines fractional order and weighted accumulation of new information. The fractional order accumulation method gives the same order to each information in the original sequence for the accumulation, and also makes the error disturbance order smaller. The weighted accumulation method of new information takes into account the timeliness of the information. The newer the data, the greater the weight, which also confirms the basic principle. The combination of the two methods can adjust the weight of information from many angles, thus increasing the accuracy. Simultaneously, the stability and the disturbance of the error boundary are optimized.

The overall structure is as follows. In Section 2, the compound cumulant grey model (CGM(1,1)) and the properties of cumulant generator are given. In Section 3, the properties are analyzed, mainly the perturbation. Then, verify whether the first value in the original sequence of the model is valuable. In Section 4, the accuracy is verified by citing several cases and comparing with other models. Section 5 is the application part. The last Section summarizes this article.

2. The compound accumulation model

The CGM(1,1) model is analyzed in detail here. The first part mainly elaborates the modeling process. The second part introduces the selection of parameters and the methods used in this model. The third part describes the properties of generating operators in this model.

2.1. The modeling process

equations (1)–(10) show the modeling process. $X^{(0)} = \{x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(n)\}$ is the non-negative sequence, two non-negative parameters are $r, \lambda \subset [0, 1]$, the sequence after fractional accumulation is

$$X^{(r)} = \{x^{(r)}(1), x^{(r)}(2), \cdots, x^{(r)}(n)\}$$

among them,

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

$$C_{k-i+r-1}^{k-i} = \frac{(k-i+r-1)(k-i+r-2)\cdots(r+2)(r+1)r}{(k-i)!}, k = 1, 2, \cdots, n, \tag{2}$$

$$C_{r-1}^{0} = 1, C_{k}^{k+1} = 0, x^{(r)}(1) = x^{(0)}(1).$$

The sequence after the weights are assigned is

$$X^{(r\lambda)} = \{x^{(r\lambda)}(1), x^{(r\lambda)}(2), \cdots, x^{(r\lambda)}(n)\},$$

among them,

$$x^{(r\lambda)}(k) = \sum_{i=1}^{k} \lambda^{k-i} x^{(r)}(i), k = 1, 2, \cdots, n. \tag{3}$$

So, the conformable fractional weighted accumulation sequence is

$$X^{(r\lambda)} = \{x^{(r\lambda)}(1), x^{(r\lambda)}(2), \cdots, x^{(r\lambda)}(n)\}.$$

Next,

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

(4)

is the whitenization equation. Then,

$$\widehat{x}^{(r\lambda)}(k) = \left(x^{(0)}(1) - \frac{b}{a}\right)e^{-a(k-1)} + \frac{b}{a}, k = 1, 2, \cdots, n.$$

(5)

Estimating the values of a and b by least square method,

$$\begin{bmatrix} a \\ b \end{bmatrix} = (B^T B)^{-1} B^T Y,$$

(6)

among them,

$$B = \begin{bmatrix} -\frac{(x^{(r\lambda)}(1) + x^{(r\lambda)}(2))}{2} & 1 \\ -\frac{0.5(x^{(r\lambda)}(2) + x^{(r\lambda)}(3))}{2} & 1 \\ \vdots & \vdots \\ -\frac{0.5(x^{(r\lambda)}(n-1) + x^{(r\lambda)}(n))}{2} & 1 \end{bmatrix}, Y = \begin{bmatrix} x^{(r\lambda)}(2) - x^{(r\lambda)}(1) \\ x^{(r\lambda)}(3) - x^{(r\lambda)}(2) \\ \vdots \\ x^{(r\lambda)}(n) - x^{(r\lambda)}(n-1) \end{bmatrix}.$$

(7)

The decreasing sequence is

$$\widehat{X}^{(r)} = \{\widehat{x}^{(r)}(1), \widehat{x}^{(r)}(2), \cdots, \widehat{x}^{(r)}(n),$$

among them,

$$\widehat{x}^{(r)}(k) = \widehat{x}^{(r\lambda)}(k) - \lambda\widehat{x}^{(r\lambda)}(k-1),$$

(8)

the predicted values $\widehat{X}^{(0)}$ is

$$\widehat{X}^{(0)} = \{\widehat{x}^{(0)}(1), \widehat{x}^{(0)}(2), \cdots, \widehat{x}^{(0)}(n)\},$$

among them,

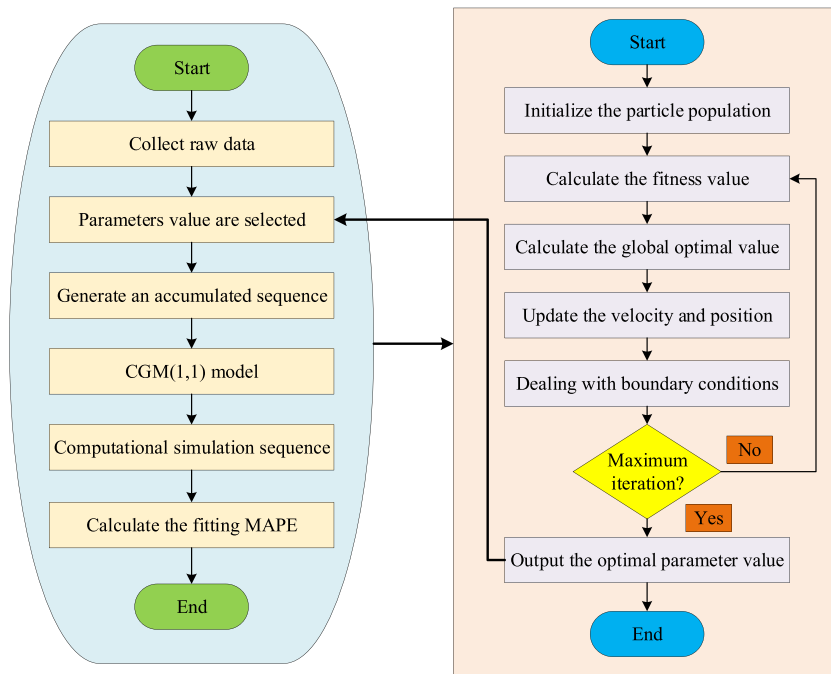$$\widehat{X}^{(0)}(k) = \widehat{x}^{(r)(1-r)}(k) - \widehat{x}^{(r)(1-r)}(k-1).$$

(9)

Fig. 1. Flowchart of searching parameters using particle swarm optimization.

Usually, the accuracy of a model is judged by the indicators. There are many evaluation indicators [28], such as MAPE, RMSPE, MSE, etc. In this paper, the MAPE value will be used as an indicator. The expression for MAPE is shown below:

$$\text{MAPE} = \frac{1}{n} \sum_{k=1}^{n} \frac{|\widehat{x}^{(0)}(k) - x^{(0)}(k)|}{x^{(0)}(k)} \times 100\%. \tag{10}$$

2.2. Parameter selection method

The model uses particle swarm optimization algorithm (PSO). The algorithm is proposed after receiving the foraging behavior of birds [31]. A flock of birds forages in an area of space where the food location is unknown. But all birds know how far away the food is. The most effective way to find food is to look around for the birds closest to the food.

Suppose there are m particles at any position in a search space. So, every particle has a position in space. The best position for each particle to pass through itself in each iteration of the generation is P_{best}. Now, the optimal position that exists in the entire particle swarm is g_{best}. In the iteration process, the change formula of particle velocity is shown in Eq. (11).

$$V_{i+1} = \omega v_i^d = c_1 r_1 \left(p_i^d - x_i^d \right) + c_2 r_2 \left(p_g^d - x_i^d \right). \tag{11}$$

where, ω stands for inertia coefficient, c_1 and c_2 stand for the acceleration coefficients, r_1 and r_2 are randomly generated numbers. The change formula of the position is shown in Eq. (12).

$$x_{i+1} = x_i + v_i. \tag{12}$$

If the particle is outside the boundary during the change, it is set equal to the boundary value. The termination condition of the search is generally set to meet the target requirement or to meet the set number of iterations. The flow chart of the CGM(1,1) combined with this algorithm is shown in Fig. 1.

2.3. Properties of generating operator

Proposition 1. The grey cumulative generating operator of the CGM(1,1) model satisfies the fixed point theorem.

Proof The expression is shown as Eq. (13).

$$x^{(r\lambda)}(k) = \sum_{i=1}^{k} \lambda^{(k-i)} \frac{(r+k-i)!}{(k-i+1)!r!} x^{(0)}(i). \tag{13}$$

When $k = 1$,.

$$x^{(r\lambda)}(1) = \sum_{i=1}^{k} \lambda^{(k-i)} \frac{(r+k-i)!}{(k-i+1)!r!} x^{(0)}(i) = \sum_{i=1}^{1} \lambda^{(1-1)} \frac{(r+1-1)!}{(1-1+1)!r!} x^{(0)}(1) = x^{(0)}(1).$$

So, $x^{(0)}(1)$ is the fixed point of $x^{(r\lambda)}(k)$.

Proposition 2. When $r \in (0,1), \lambda \in (0,1),$ the grey cumulative generator of CGM(1,1) model is new information first.

Proof From $x^{(r\lambda)}(k)$, the coefficient of $x^{(0)}(i)$ is shown in Eq. (14).

$$a_{ki} = \lambda^{(k-i)} \frac{(r+k-i)!}{(k-i+1)!r!}, \tag{14}$$

the coefficient of $x^{(0)}(i-1)$ is shown as Eq. (15) and Eq. (16).

$$a_{k(i-1)} = \lambda^{(k-i+1)} \frac{(r+k-i+1)!}{(k-i+2)!r!}, \tag{15}$$

$$\frac{a_{ki}}{a_{k(i-1)}} = \frac{\lambda^{(k-i)} \frac{(r+k-i)!}{(k-i+1)!r!}}{\lambda^{(k-i+1)} \frac{(r+k-i+1)!}{(k-i+2)!r!}} = \frac{1}{\lambda} \frac{(r+k-i)!}{(k-i+1)!r!} \frac{(k-i+2)!r!}{(r+k-i+1)!} = \frac{1}{\lambda} \frac{(r+k-i+2-r)}{(r+k-i+1)} = \frac{1}{\lambda} \left(1 + \frac{1-r}{r+k-i+1} \right). \tag{16}$$

When $r \in (0,1), \lambda \in (0,1), i \le k, \frac{1}{\lambda} > 1, 1 + \frac{1-r}{r+k-i+1} > 1,$.

so,

$$\frac{a_{ki}}{a_{k(i-1)}} = \frac{1}{\lambda} \left(1 + \frac{1-r}{r+k-i+1} \right) > 1.$$

This shows that in the expression of $x^{(r\lambda)}(k)$, the coefficient of $x^{(0)}(i)$ is larger than that of $x^{(0)}(i-1)$, and has greater weight. Therefore, it is proved.

Proposition 3. When $r > 1, \lambda > 1,$ the old information in the grey accumulation operator of the CGM(1,1) model has greater weight.

Proof Same as the formula in Proposition 2.

$$\frac{a_{ki}}{a_{k(i-1)}} = \frac{\lambda^{(k-i)}\frac{(r+k-i)!}{(k-i+1)!r!}}{\lambda^{(k-i+1)}\frac{(r+k-i+1)!}{(k-i+2)!r!}} = \frac{1}{\lambda}\frac{(r+k-i)!}{(k-i+1)!r!}\frac{(k-i+2)!r!}{(r+k-i+1)!} = \frac{1}{\lambda}\frac{(r+k-i+2-r)}{(r+k-i+1)}$$

$$= \frac{1}{\lambda}\left(1 + \frac{1-r}{r+k-i+1}\right).$$

When $r > 1, \lambda > 1, i \le k,$.

$$\frac{1}{\lambda} < 1, 1 + \frac{1-r}{r+k-i+1} < 1.$$

So,

$$\frac{a_{ki}}{a_{k(i-1)}} = \frac{1}{\lambda}\left(1 + \frac{1-r}{r+k-i+1}\right) < 1.$$

This shows that in the expansion expression of $x^{(r\lambda)}(k)$, the coefficient of $x^{(0)}(i)$ is larger than that of $x^{(0)}(i-1)$, and has greater weight, so it indicates that the old information has priority.

Proposition 4. When $r = 1, \lambda = 1$, it has equal information of grey accumulation generating operator.

Proof Same as the formula in Proposition 2.

$$\frac{a_{ki}}{a_{k(i-1)}} = \frac{\lambda^{(k-i)}\frac{(r+k-i)!}{(k-i+1)!r!}}{\lambda^{(k-i+1)}\frac{(r+k-i+1)!}{(k-i+2)!r!}} = \frac{1}{\lambda}\frac{(r+k-i)!}{(k-i+1)!r!}\frac{(k-i+2)!r!}{(r+k-i+1)!} = \frac{1}{\lambda}\frac{(r+k-i+2-r)}{(r+k-i+1)}$$

$$= \frac{1}{\lambda}\left(1 + \frac{1-r}{r+k-i+1}\right).$$

When $r = 1, \lambda = 1,$.

$$\frac{a_{ki}}{a_{k(i-1)}} = \frac{1}{\lambda}\left(1 + \frac{1-r}{r+k-i+1}\right) = 1,$$

$$a_{ki} = a_{k(i-1)}.$$

Therefore, it is proved.

When $r = 1$ and $\lambda = 0$, the CGM(1,1) model is transformed into the GM(1,1) model. when $r = 1, \lambda \in (0,1)$, it is transformed into the WDGM(1,1) model. When $\lambda = 0, r \in (0,1)$, it is transformed into the FGM(1,1) model.

Two cases are introduced to analyze the four models. Table 1 studies the energy consumption in Henan Province. Table 2 takes the carbon emission as the research object. Data comes from literature [21]. The GM(1,1), discrete grey model (DGM(1,1)), WDGM(1,1) and FGM(1,1) models are used in both cases for comparison. The purpose is to analyze whether the CGM(1,1) model has improved accuracy.

According to the statistical data in Table 1, the CGM(1,1) model's MAPE is 0.21%. Compared with 0.85% for the WDGM(1,1) and 0.5% for the FGM(1,1), the CGM(1,1) model works best. From Table 2, the CGM(1,1) model's MAPE is 1.6%, which is also the smallest in comparison. Although in some years, the data error is greater than that of other models, but this is acceptable. When looking for parameters, the objective is to minimize the overall error value. Therefore, there may be large errors in individual years.

To sum up, compared with the two models before composite, the accuracy of the composite model is indeed improved.

3. Properties of the CGM(1,1) model

The properties need to be proved. The first part mainly proves the perturbation bound of the model, as shown in equations (17)–(22). The second part mainly proves whether the first value in the original sequence of the model has any effect.

Table 1
The fitting results on the energy consumption in Henan Province

Year	Actual value	GM(1,1)	DGM(1,1)	WDGM(1,1)	*FGM(1,1)*	*CGM(1,1)*
2003	10595	10595	10595	10595	10595	10595
2004	13074	13307	13319	12830	13038	13072
2005	14625	14593	14606	14791	14736	14625
2006	16234	16003	16017	16414	16234	16284
2007	17838	17549	17565	17758	17657	17733
2008	18976	19244	19262	18872	19058	19043
MAPE(%)		1.01	1.27	0.85	0.50	**0.21**

Table 2
The carbon emission fitting error of the five models in the Beijing-Tianjin-Hebei region

Year	Actual value	GM(1,1)	DGM(1,1)	WDGM(1,1)	FGM(1,1)	CGM(1,1)
2000	35726.78	35726.78	35726.78	35726.78	35726.78	35726.78
2001	36471.51	37949.98	37981.67	32408.96	34660.35	36446.97
2002	38402.42	40762.73	40795.89	38435.42	39073.59	37765.94
2003	41530.01	43783.94	43818.64	43945.52	43433.44	42728.67
2004	46828.24	47029.08	47065.35	48983.48	47640.24	47309.86
2005	53241.51	50514.74	50552.62	53589.78	51749.30	51670.96
2006	55868.82	54258.74	54298.28	57801.38	55813.53	55904.16
2007	61360.13	58280.24	58321.47	61652.12	59873.66	60067.93
2008	62511.46	62599.81	62642.76	65172.92	63960.81	64202.33
2009	66640.81	67239.52	67284.23	68392.04	68099.53	68336.39
2010	72472.94	72223.12	72269.60	71335.33	72309.79	72492.25
2011	79335.39	77576.09	77624.36	74026.43	76608.49	76687.40
2012	79890.07	83325.80	83375.89	76486.94	81010.33	80936.23
MAPE(%)		3.08	3.07	3.51	2.33	**1.60**

3.1. Disturbance analysis

Lemma 1. [32,33] Let A $\in$ C$^{m \times n}$, b $\in$ C^m, there is a matrix A whose generalized inverse is matrix A†, if the columns of A are linearly independent, then

$$\|Ax - b\|_2 = \min, \tag{17}$$

has only one solution x $= A^\dagger b$.

Lemma 2. [32,33] Let A $\in$ C$^{m \times n}$, b $\in$ C^m, there is a matrix A whose generalized inverse is matrix A†,

$$B = A + E, c = b + k \in C^m.$$

Let the solutions of the linear least squares problem

$$\|Bx - c\|_2 = min, \tag{18}$$

and

$$\|Ax - b\|_2 = min,$$

be $x + h$ and x.
If $rank(A) = rank(B) = n$, and $\|A^\dagger\|_2 \|E\|_2 < 1$, then

$$\|h\| \leq \frac{\kappa_\dagger}{\gamma_\dagger} \left(\frac{\|E\|_2}{\|A\|_2} \|x\| + \frac{\|k\|}{\|A\|} + \frac{\kappa_\dagger}{\gamma_\dagger} \frac{\|E\|_2}{\|A\|_2} \frac{\|r_x\|}{\|A\|} \right), \tag{19}$$

where

$$\kappa_\dagger = \|A^\dagger\|_2 \|A\|, \gamma_\dagger = 1 - \|A^\dagger\|_2 \|E\|_2, r_x = b - Ax. \tag{20}$$

From these two lemmas, E is the disturbance matrix, and k is the disturbance vector. The perturbation of the solution is h. The perturbation bound of least squares can be obtained by perturbation analysis of matrices.

Theorem 1. *The CGM(1,1) model is* $x^{(r\lambda)}(k + 1) = \beta_1 x^{(r\lambda)}(k) + \beta_2$. *the least square estimation of the parameter satisfies*

$$\begin{bmatrix} \beta_2 \\ \beta_1 \end{bmatrix} = \left(B^T B \right)^{-1} B^T Y, \tag{21}$$

where

$$B = \begin{bmatrix} 1 & x^{(r\lambda)}(1) \\ 1 & x^{(r\lambda)}(1) \\ \vdots & \vdots \\ 1 & x^{(r\lambda)}(n-2) \\ 1 & x^{(r\lambda)}(n-1) \end{bmatrix}, Y = \begin{bmatrix} x^{(r\lambda)}(2) \\ x^{(r\lambda)}(3) \\ \vdots \\ x^{(r\lambda)}(n-1) \\ x^{(r\lambda)}(n) \end{bmatrix}. \tag{22}$$

Theorem 2. *According to the least square method*

$$min\|Bx - Y\|_2,$$

the solutions of the CGM(1,1) model $x^{(r\lambda)}(k+1) = \beta_1 x^{(r\lambda)}(k) + \beta_2$ is x. If there is a disturbance,

$$\widehat{x}^{(0)}(1) = x^{(0)}(1) + \varepsilon,$$

$$\widehat{B} = B + \Delta B = \begin{bmatrix} 1 & x^{(r\lambda)}(1) \\ 1 & x^{(r\lambda)}(2) \\ \vdots & \vdots \\ 1 & x^{(r\lambda)}(n-1) \end{bmatrix} + \begin{bmatrix} 0 & \varepsilon \\ 0 & r\lambda\varepsilon \\ \vdots & \vdots \\ 0 & \lambda^{n-2} C_{n-3+r}^{n-2} \varepsilon \end{bmatrix},$$

$$\widehat{Y} = Y + \Delta Y = \begin{bmatrix} x^{(r\lambda)}(2) \\ x^{(r\lambda)}(3) \\ \vdots \\ x^{(r\lambda)}(n) \end{bmatrix} + \begin{bmatrix} \lambda r\varepsilon \\ \lambda^2 C_{1+r}^2 \varepsilon \\ \vdots \\ \lambda^{n-1} C_{n-2+r}^{n-1} \varepsilon \end{bmatrix}.$$

The solution of the least squares problem $\min\|\widehat{B}x - \widehat{Y}\|_2$ is $\widehat{x}$, the perturbation of the solution is Δx. Let $k(B) = \text{rank}(\widehat{B}) = 2$, $\|B^\dagger\|_2 \|\Delta B\|_2 < 1$. Then

$$\|\Delta x\| \leq |\varepsilon| \frac{\kappa_\dagger}{\gamma_\dagger} \left(\frac{\left(\sqrt{\sum_{k=1}^{n-1} \lambda^{2(k-1)} \left(C_{k+r-2}^{k-1} \right)^2} \right)^2}{\|B\|} \|x\| + \frac{\sqrt{\sum_{k=2}^{n} \lambda^{2(k-1)} \left(C_{k+r-2}^{k-1} \right)^2}}{\|B\|} + \frac{\kappa_\dagger}{\gamma_\dagger} \frac{\left(\sqrt{\sum_{k=1}^{n-1} \lambda^{2(k-1)} \left(C_{k+r-2}^{k-1} \right)^2} \right)^2}{\|B\|} \frac{\|r_x\|}{\|B\|} \right).$$

Proof When $\widehat{x}^{(0)}(1) = x^{(0)}(1) + \varepsilon$,

$$\|\Delta Y\|_2 = \sqrt{(\lambda r)^2 + \left(\lambda^2 C_{1+r}^2 \right)^2 + \cdots + \left(\lambda^{(n-1)} C_{n-2+r}^{n-1} \right)^2 |\varepsilon|^2} = |\varepsilon| \sqrt{\sum_{k=2}^{n} \lambda^{(k-1)} \left(C_{k+r-2}^{k-1} \right)^2}$$

$$\Delta B^T \Delta B = \begin{bmatrix} 0 & 0 \\ 0 & \left[1 + \lambda^2 r^2 + \cdots + \lambda^{2(n-2)} \left(C_{n-3+r}^{n-2} \right)^2 \right] \varepsilon^2 \end{bmatrix}.$$

the maximum eigenvalue of $\Delta B^T \Delta B$ is $[1 + \lambda^2 r^2 + \cdots + \lambda^{2(n-2)} (C_{n-3+r}^{n-2})^2]\varepsilon^2$, thus,

$$\|\Delta B\|_2 = \sqrt{1 + \lambda^2 r^2 + \cdots + \lambda^{2(n-2)} \left(C_{n-3+r}^{n-2} \right)^2} |\varepsilon| = \sqrt{\sum_{k=1}^{n-1} \lambda^{2(k-1)} \left(C_{k+r-2}^{k-1} \right)^2} |\varepsilon|,$$

according to Lemma 2,

$$\|\Delta x\| \leq \frac{\kappa_\dagger}{\gamma_\dagger} \left(\frac{\|\Delta B\|_2}{\|B\|} \|x\| + \frac{\|\Delta Y\|}{\|B\|} + \frac{\kappa_\dagger}{\gamma_\dagger} \frac{\|\Delta B\|_2}{\|B\|} \frac{\|r_x\|}{\|B\|} \right) =$$

If $\widehat{x}^{(0)}(1) = x^{(0)}(1) + \varepsilon$, the perturbation bound of the CGM(1,1) model is

$$L\left[x^{(0)}(1) \right] = |\varepsilon| \frac{\kappa_\dagger}{\gamma_\dagger} \left(\frac{\sqrt{\sum_{k=1}^{n-1} \lambda^{2(k-1)} \left(C_{k+r-2}^{k-1} \right)^2}}{\|B\|} \|x\| + \frac{\sqrt{\sum_{k=2}^{n} \lambda^{2(k-1)} \left(C_{k+r-2}^{k-1} \right)^2}}{\|B\|} + \frac{\kappa_\dagger}{\gamma_\dagger} \frac{\sqrt{\sum_{k=1}^{n-1} \lambda^{2(k-1)} \left(C_{k+r-2}^{k-1} \right)^2}}{\|B\|} \frac{\|r_x\|}{\|B\|} \right).$$

$$|\varepsilon| \frac{\kappa_\dagger}{\gamma_\dagger} \left(\frac{\sqrt{\sum_{k=1}^{n-1} \lambda^{2(k-1)} \left(C_{k+r-2}^{k-1} \right)^2}}{\|B\|} \|x\| + \frac{\sqrt{\sum_{k=2}^{n} \left(\lambda^{k-1} C_{k+r-2}^{k-1} \right)^2}}{\|B\|} + \frac{\kappa_\dagger}{\gamma_\dagger} \frac{\sqrt{\sum_{k=1}^{n-1} \left(\lambda^{k-1} C_{k+r-2}^{k-1} \right)^2}}{\|B\|} \frac{\|r_x\|}{\|B\|} \right).$$

Theorem 3. Compared with the Theorem 2, under the premise of unchanged conditions. If $\widehat{x}^{(0)}(2) = x^{(0)}(2) + \varepsilon$, the perturbation bound of the CGM(1,1) model is

$$L[x^{(0)}(2)] = |\varepsilon| \frac{\kappa_\dagger}{\gamma_\dagger} \left(\frac{\sqrt{\sum_{k=1}^{n-2} \lambda^{2(k-1)} \left(C_{k+r-2}^{k-1} \right)^2}}{\|B\|} \|x\| + \frac{\sqrt{\sum_{k=1}^{n-1} \lambda^{2(k-1)} \left(C_{k+r-2}^{k-1} \right)^2}}{\|B\|} + \frac{\kappa_\dagger}{\gamma_\dagger} \frac{\sqrt{\sum_{k=1}^{n-2} \lambda^{2(k-1)} \left(C_{k+r-2}^{k-1} \right)^2}}{\|B\|} \frac{\|r_x\|}{\|B\|} \right).$$

Similarly, if $\widehat{x}^{(0)}(s) = x^{(0)}(s) + \varepsilon, s = 3, 4, \cdots, n-1,$ the perturbation bound of the CGM(1,1) model is

$$L[x^{(0)}(r)] =$$

$$|\varepsilon| \frac{\kappa_\dagger}{\gamma_\dagger} \left(\frac{\sqrt{\sum_{k=1}^{n-s} \lambda^{2(k-1)} \left(C_{k+r-2}^{k-1} \right)^2}}{\|B\|} \|x\| + \frac{\sqrt{\sum_{k=1}^{n-s+1} \lambda^{2(k-1)} \left(C_{k+r-2}^{k-1} \right)^2}}{\|B\|} + \frac{\kappa_\dagger}{\gamma_\dagger} \frac{\sqrt{\sum_{k=1}^{n-r} \lambda^{2(k-1)} \left(C_{k+r-2}^{k-1} \right)^2}}{\|B\|} \frac{\|r_x\|}{\|B\|} \right).$$

$s = 3, 4, \cdots, n-1.$

If $\widehat{x}^{(0)}(n) = x^{(0)}(n) + \varepsilon$, then the perturbation boundary is

$$L[x^{(0)}(n)] = \frac{\kappa_\dagger}{\gamma_\dagger} \frac{|\varepsilon|}{\|B\|}.$$

Proof When $\widehat{x}^{(0)}(2) = x^{(0)}(2) + \varepsilon,$

$$\Delta B = \begin{bmatrix} 0 & 0 \\ 0 & \varepsilon \\ \vdots & \vdots \\ 0 & \lambda^{n-3} C_{n-4+r}^{n-3} \varepsilon \end{bmatrix}, \Delta Y = \begin{bmatrix} \varepsilon \\ \lambda r \varepsilon \\ \vdots \\ \lambda^{n-2} C_{n-1+r}^{n-2} \varepsilon \end{bmatrix}.$$

the perturbation boundary is

$$L[x^{(0)}(2)] =$$

$$|\varepsilon| \frac{\kappa_\dagger}{\gamma_\dagger} \left(\frac{\sqrt{\sum_{k=1}^{n-2} \lambda^{2(k-1)} \left(C_{k+r-2}^{k-1} \right)^2}}{\|B\|} \|x\| + \frac{\sqrt{\sum_{k=1}^{n-1} \lambda^{2(k-1)} \left(C_{k+r-2}^{k-1} \right)^2}}{\|B\|} + \frac{\kappa_\dagger}{\gamma_\dagger} \frac{\sqrt{\sum_{k=1}^{n-2} \lambda^{2(k-1)} \left(C_{k+r-2}^{k-1} \right)^2}}{\|B\|} \frac{\|r_x\|}{\|B\|} \right).$$

When $\widehat{x}^{(0)}(s) = x^{(0)}(s) + \varepsilon, s = 3, 4, \cdots, n-1,$ then

$$L[x^{(0)}(r)] =$$

$$|\varepsilon| \frac{\kappa_\dagger}{\gamma_\dagger} \left(\frac{\sqrt{\sum_{k=1}^{n-s} \lambda^{2(k-1)} \left(C_{k+r-2}^{k-1} \right)^2}}{\|B\|} \|x\| + \frac{\sqrt{\sum_{k=1}^{n-s+1} \lambda^{2(k-1)} \left(C_{k+r-2}^{k-1} \right)^2}}{\|B\|} + \frac{\kappa_\dagger}{\gamma_\dagger} \frac{\sqrt{\sum_{k=1}^{n-r} \lambda^{2(k-1)} \left(C_{k+r-2}^{k-1} \right)^2}}{\|B\|} \frac{\|r_x\|}{\|B\|} \right)$$

$r = 3, 4, \cdots, n-1$

If $\widehat{x}^{(0)}(n) = x^{(0)}(n) + \varepsilon$, the perturbation boundary is

$$L[x^{(0)}(n)] = \frac{\kappa_\dagger}{\gamma_\dagger} \frac{|\varepsilon|}{\|B\|}.$$

Thus, the result is proved.

When $\widehat{x}^{(0)}(1) = x^{(0)}(1) + \varepsilon$, the perturbation boundary is

$$L\left[x^{(0)}(1)\right]=|\varepsilon|\frac{\kappa_{\dagger}}{\gamma_{\dagger}}\left(\frac{\sqrt{\sum_{k=1}^{n-s}\lambda^{2(k-1)}\left(C_{k+r-2}^{k-1}\right)^{2}}}{\|\mathbf{B}\|}\|\mathbf{x}\|+\frac{\sqrt{\sum_{k=1}^{n-s+1}\lambda^{2(k-1)}\left(C_{k+r-2}^{k-1}\right)^{2}}}{\|\mathbf{B}\|}+\frac{\kappa_{\dagger}}{\gamma_{\dagger}}\frac{\sqrt{\sum_{k=1}^{n-r}\lambda^{2(k-1)}\left(C_{k+r-2}^{k-1}\right)^{2}}}{\|\mathbf{B}\|}\frac{\|\mathbf{r}_{\mathbf{x}}\|}{\|\mathbf{B}\|}\right).$$

Compared with the FGM(1,1) model

$$L\left[x^{(0)}(1)\right]=|\varepsilon|\frac{\kappa_{\dagger}}{\gamma_{\dagger}}\left(\frac{\sqrt{\sum_{k=1}^{n-1}\left(C_{k+r-2}^{k-1}\right)^{2}}}{\|B\|}\|x\|+\frac{\sqrt{\sum_{k=2}^{n}\left(C_{k+r-2}^{k-1}\right)^{2}}}{\|B\|}+\frac{\kappa_{\dagger}}{\gamma_{\dagger}}\frac{\sqrt{\sum_{k=1}^{n-1}\left(C_{k+r-2}^{k-1}\right)^{2}}}{\|B\|}\frac{\|r_{x}\|}{\|B\|}\right)$$

and the WDGM(1,1) *model*

$$L\left[x^{(0)}(1)\right]=|\varepsilon|\frac{\kappa_{\dagger}}{\gamma_{\dagger}}\left(\frac{\sqrt{\sum_{i=1}^{n-2}\lambda^{2i}}}{\|B\|}\|x\|+\frac{\sqrt{\sum_{i=1}^{n-1}\lambda^{2i}}}{\|B\|}+\frac{\kappa_{\dagger}}{\gamma_{\dagger}}\frac{\sqrt{\sum_{i=0}^{n-2}\lambda^{2i}}}{\|B\|}\frac{\|r_{x}\|}{\|B\|}\right),$$

it decreased significantly.

Because the perturbation bound is smaller than the WDGM(1,1) and FGM(1,1). Therefore, it is proved.

3.2. Analysis of the validity of initial value

An important principle is to mine the information contained in data. But in traditional model, the first data information is not mined because changing this value has no effect on the fitting sequence. Therefore, the first numerical information in the original sequence of traditional model is not mined, which has been proved in Li's paper on air quality [34].

Next, verify that the first value of the original sequence of the CGM(1,1). In other words, we observe whether the changes of the initial values affect the fitting results. In order to get a more convenient and intuitive conclusion, a case is introduced for analysis. The grain output of Hebei Province during 2013–2018 was studied, the data is from the National Bureau of Statistics (http://www.ststs. gov.cn/). The operation results of the GM(1,1) and CGM(1,1) models are shown in Table 3.

The analysis shows the data growth rate fitted by CGM(1,1) is variable, but the growth rate of GM(1,1) is fixed. So, the CGM(1,1) model can more truly reflect the development trend when predicting cases.

In order to more directly reflect whether the first value in the original sequence of the CGM(1,1) model has any effect, the first value of the original sequence 3584.87 is replaced by 3560. Among them, the replacement data can be arbitrarily selected. The purpose is to verify whether the fitting sequence will change after the initial value is changed. The operation results are shown in Table 4.

In order to prevent the occurrence of coincidence, the first value 3584.87 is replaced by 3600. Of course, this data is also arbitrary and the operation results are shown in Table 5.

According to Tables 4 and tbl5, after the first value changes, the results of the GM(1,1) model has not changed. Therefore, the first value is invalid. However, in contrast, after the first value changes, the fitting results have changed. This fully demonstrates the importance of the first number in the numerical sequence to the CGM(1,1) model. that is to say, the first value in the sequence is utilized. Thus, the first value in the CGM(1,1) model sequence is fully mined for information.

4. Empirical validation

This chapter introduces three practical cases. At the same time, other models are cited to further verify the accuracy.

Case 1. The research object is natural gas, the research area is China, and the time node is from 2002 to 2010 [35].

In this case, natural gas (Units: Billion Cubic Meters (BCM)) is used as research target. Use data from 2002 to 2010 to predict results for the next four years. The MAPE value is used to compare the models. The simulation results are shown in Table 6. Table 7 collates the

Table 3
The operation results of the models.

Time	value	growth rate	GM(1,1)	growth rate	CGM(1,1)	growth rate
2013	3584.87		3584.87		3584.87	
2014	3568.98	−0.45%	3600.51	0.44%	3558.88	−0.73%
2015	3602.19	0.92%	3648.03	1.31%	3601.76	1.19%
2016	3782.99	4.78%	3696.18	1.31%	3782.34	4.77%
2017	3829.25	1.21%	3744.96	1.31%	3804.15	0.57%
2018	3700.86	−3.47%	3794.39	1.31%	3723.03	−2.18%

Table 4
The operation results of the models.

Time	value	GM(1,1)	CGM(1,1)	New sequence	GM(1,1)	CGM(1,1)
2013	3584.87	3584.87	3584.87	3560.00	3560.00	3560.00
2014	3568.98	3600.51	3558.88	3568.98	3600.51	3561.01
2015	3602.19	3648.03	3601.76	3602.19	3648.03	3598.01
2016	3782.99	3696.18	3782.34	3782.99	3696.18	3781.02
2017	3829.25	3744.96	3804.15	3829.25	3744.96	3805.46
2018	3700.86	3794.39	3723.03	3700.86	3794.39	3726.90

Table 5
The operation results of the models.

Time	value	GM(1,1)	CGM(1,1)	New sequence	GM(1,1)	CGM(1,1)
2013	3584.87	3584.87	3584.87	3600.00	3600.00	3600.00
2014	3568.98	3600.51	3558.88	3568.98	3600.51	3551.87
2015	3602.19	3648.03	3601.76	3602.19	3648.03	3603.09
2016	3782.99	3696.18	3782.34	3782.99	3696.18	3794.44
2017	3829.25	3744.96	3804.15	3829.25	3744.96	3805.64
2018	3700.86	3794.39	3723.03	3700.86	3794.39	3700.87

annual error percentages.

Table 6 shows that the CGM(1,1)'s MAPE(2002–2010) is only 1.72%, which is significantly better than other models. The CGM (1,1)'s MAPE(2011–2014) is 2.84%, while the MAPE of the GM(1,1) and SIGM(1,1) models is more than 5%. From Table 7 and Fig. 2, in contrast, CGM(1,1) has smallest annual relative error percentage in the simulation results. Obviously, CGM(1,1) has the best accuracy in this case.

Case 2. The research object is energy consumption, the research area is Jiangsu Province, and the time node is from 2001 to 2008 [36].

The energy consumption is chosen as our research object (Units: 10^4 tons of standard coal). First, use data from 2001 to 2008 to predict results for the next four years. Then, other models are introduced. The data simulated by the models are organized into Table 8.

From Tables 8 and 9, in contrast, the MAPE value of the CGM(1,1) model is the smallest. As can be seen from Fig. 3, the CGM(1,1) model has the smallest relative error except for individual years. This indicates that the simulation results are very effective. Therefore, the CGM(1,1) model has the best accuracy in this case.

Case 3. The case for wind turbine capacity in Europe [37].

In this case, wind turbine capacity is selected for the study (Units: megawatts). Use data from 2007 to 2015 to predict results. Meanwhile, other models are introduced. The data simulated by the models are organized into Table 10.

From Table 10, the MAPE of the CGM(1,1) model are 0.41% and 0.14%. Among the other four models, the DGM(1,1) model's fitting MAPE is 1.54%. The NGBM(1,1) model's prediction MAPE is 3.07%. From Table 11 and Fig. 4, the CGM(1,1) model's simulation results are relatively stable, and the largest percentage error occurred in 2009, with an error of 1.12%. The simulation results are very ideal.

Based on the above information, The accuracy of CGM(1,1) has improved. Therefore, CGM(1,1) has high fitting and prediction ability. It can be applied to real case studies.

Table 6
Consumption of natural gas (Units: BCM).

Year	Actual	GM(1,1)	SIGM(1,1)	FGM(1,1)	FHGM(1,1)	CGM(1,1)
2002	29.2	29.2	29.2	29.2	29.2	29.2
2003	33.9	34.7	32.98	33.4	32.9	33.9
2004	39.7	40.8	40.1	39.7	40.3	39.7
2005	46.8	47.9	48.1	47.3	48.2	47.8
2006	56.1	56.2	57.1	56.1	56.9	56.7
2007	69.5	66.1	67.2	66.3	66.9	66.9
2008	80.7	77.6	78.7	78.0	78.2	78.3
2009	87.5	91.1	91.5	91.4	91.1	91.2
2010	107.5	107.0	106.0	106.8	106.0	106.1
MAPE(%)		2.64	2.50	2.00	2.36	**1.72**
2011	131.3	125.7	122.2	124.4	123.0	123.0
2012	147.1	147.6	140.6	144.6	142.6	142.3
2013	165.0	173.4	161.2	167.7	165.0	164.3
2014	187.0	203.6	184.4	194.2	190.9	189.5
MAPE(%)		6.19	5.02	3.11	2.87	**2.84**

Table 7
Percentage of simulated prediction errors for the five models.

| Year	Simulation error(%) of GM (1, 1)	Simulation error (%)of SIGM(1,1)	Simulation error (%)of FGM(1,1)	Simulation error(%) of FHGM(1,1)	Simulation error(%) of CGM(1,1)
2002	0.00	0.00	0.00	0.00	0.00
2003	2.41	2.70	1.47	2.95	0.00
2004	2.71	0.99	0.00	1.51	0.00
2005	2.33	2.77	1.07	2.99	2.14
2006	0.26	1.79	0.00	1.43	1.07
2007	4.95	3.25	4.60	3.74	3.74
2008	3.86	2.53	3.35	3.10	2.97
2009	4.13	4.58	4.46	4.11	4.23
2010	0.45	1.43	0.65	1.40	1.30
2011	4.28	6.90	5.26	6.32	6.32
2012	0.35	4.44	1.70	3.06	3.26
2013	5.07	2.31	1.64	0.00	0.42
2014	8.88	1.39	3.85	2.09	1.34
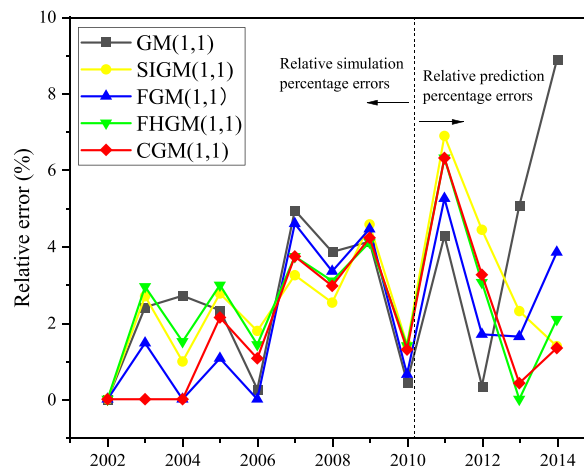
Fig. 2. Relative errors of five models.

Table 8
Prediction results of five models (Units: 10^4 tons of standard coal).

Year	Actual	GM (1, 1)	FGM(1,1)	NIPGM(1,1) 1)1)	WFGM(1,1)	CGM(1,1)
2001	8881	8881	8881	8881	8881	8881
2002	9609	10483	9998	9609	9621	9596
2003	11061	11988	11990	11826	11850	11062
2004	13652	13710	14048	14031	14087	13970
2005	17167	15678	16129	16224	16276	16561
2006	18742	17930	18239	18404	18411	18862
2007	20948	20505	20391	20572	20493	20898
2008	22232	23449	22599	22728	22522	22690
MAPE(%)		4.81	3.55	2.63	2.61	**1.12**
2009	23709	26816	24875	24872	24499	24258
2010	25774	30667	27228	27004	26426	25621
2011	27589	35071	29669	29124	28304	26795
2012	28850	40108	32207	31232	30134	27796
MAPE(%)		24.56	7.43	5.87	3.23	**2.36**

5. Empirical results and analysis

Non-renewable energy refers to the energy that cannot be renewed after people exploit and use it at the present stage, mainly including oil, coal, natural gas, nuclear energy, etc. Society cannot develop without these sources of energy, especially in some developing countries, which consume large amounts of these non-renewable. As the existing amount of non-renewable energy is limited, if people continue to exploit it without control, these resources will soon be exhausted. Although now people have begun to focus on replace the non-renewable energy, consumption is still high.

Table 9
Percentage of simulation errors for the five models.

| Year	Simulation error(%) of GM(1,1)	Simulation error(%) of FGM(1,1)	Simulation error(%) of NIPGM(1,1)	Simulation error(%) of WFGM(1,1)	Simulation error(%) of CGM(1,1)
2001	0.00	0.00	0.00	0.00	0.00
2002	9.09	4.05	0.00	1.26	0.14
2003	8.38	8.40	6.92	7.14	0.00
2004	0.42	2.90	2.78	3.18	2.33
2005	8.67	6.05	5.50	5.19	3.53
2006	4.33	2.69	1.80	1.76	0.64
2007	2.12	2.66	1.80	2.17	0.24
2008	5.47	1.65	2.23	1.31	2.06
2009	13.12	4.92	4.90	3.33	2.32
2010	18.99	5.64	4.77	2.53	0.59
2011	27.12	7.54	5.56	2.59	2.88
2012	39.02	11.64	8.26	3.23	3.65
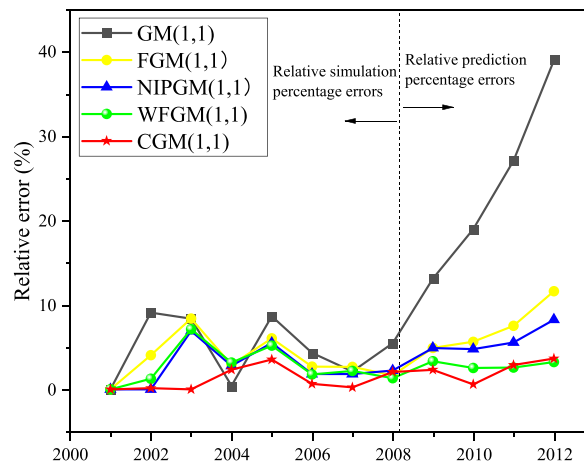
Fig. 3. The relative error percentage of the energy consumption in Jiangsu province by five models.

Table 10
Five models simulate wind turbine capacity (Units: megawatts).

Year	Actual	GM(1, 1)	DGM(1,1)	NGBM(1,1)	NGM(1,1,k,c)	CGM(1,1)
2007	56748.89	56748.89	56748.89	56748.89	56748.89	56748.89
2008	64943.48	68733.07	68819.11	68053.13	62092.57	65284.87
2009	77019.99	76808.49	76910.91	76293.36	72082.88	76154.00
2010	86721.97	85832.7	85954.16	85366.11	82532.46	86751.70
2011	96603.13	95917.15	96060.72	95432.01	93462.41	97610.32
2012	109884.87	107186.43	107355.62	106629.26	104894.83	108972.89
2013	120994.68	119779.73	119978.58	119099.88	116852.81	121003.59
2014	133915.44	133852.61	134085.77	132997.36	129360.5	133833.00
2015	147637.65	149578.91	149851.68	148490.51	142443.19	147576.92
MAPE(%)		1.58	1.54	1.79	4.22	**0.41**
2016	161939.87	167152.88	167471.36	165766.36	156127.30	162345.59
2017	178314.15	186791.62	187162.78	185032.76	170440.49	178248.83
MAPE(%)		3.99	4.19	3.07	4.00	**0.14**

According to the Statistical Review of World Energy (https://www.bp.com)(BP), China is one of the world's top coal consumers. Oil consumption is also considerable, the second largest in the world. Natural gas consumption ranks third in the world. The consumption of non-renewable energy is huge in China. Therefore, this section mainly forecasts China's non-renewable energy consumption. Table 12 collects consumption data for recent years. The data comes from the BP.

Next, take oil consumption as an example to carry out detailed calculation. The other two sources of energy only show results. China's oil consumption from 2010 to 2021 is sorted into the original sequence,

$$X^{(0)} = \{18.99, 19.41, 20.36, 21.27, 22.11, 23.80, 24.56, 25.86, 27.12, 28.49, 28.74, 30.60\}.$$

Table 11
Percentage error of five models.

| Year	Simulation error(%) of GM(1,1)	Simulation error(%) of DGM(1,1)	Simulation error(%) of NGBM(1,1,k)	Simulation error(%) of NGM(1,1,k,c)	Simulation error(%) of CGM(1,1)
2007	0.00	0.00	0.00	0.00	0.00
2008	5.84	5.97	4.79	4.39	0.53
2009	0.27	0.14	0.94	6.41	1.12
2010	1.03	0.89	1.56	4.83	0.03
2011	0.71	0.56	1.21	3.25	1.04
2012	2.46	2.30	2.96	4.54	0.83
2013	1.00	0.84	1.57	3.42	0.01
2014	0.05	0.13	0.69	3.40	0.06
2015	1.31	1.50	0.58	3.52	0.04
2016	3.22	3.42	2.36	3.59	0.25
2017	4.75	4.96	3.77	4.42	0.04
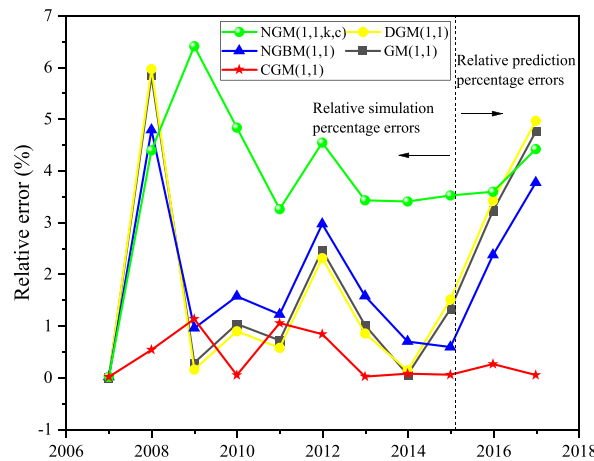
Fig. 4. Percentage error of five models.

Table 12
Energy consumption (Units: Exajoules).

Year	Coal	oil	Natural gas	Year	Coal	oil	Natural gas
2010	73.22	18.99	3.92	2016	80.19	24.56	7.54
2011	79.71	19.41	4.87	2017	80.56	25.86	8.69
2012	80.71	20.36	5.43	2018	81.05	27.12	10.22
2013	82.43	21.27	6.19	2019	81.70	28.49	11.10
2014	82.48	22.11	6.78	2020	82.38	28.74	12.12
2015	80.92	23.80	7.01	2021	86.17	30.60	13.63

The particle swarm optimization algorithm as shown in **Section 2.2** is used to obtain the best parameter values. The optimization process is shown in **Fig. 1**, which is realized by MATLAB. Where $r = 0.9685$, $\lambda = 0.1037$.

From **Eq. (1)** and **Eq. (2)**, the sequence after r-order accumulation is

$$X^{(r)} = \{18.99, 37.98, 57.65, 78.01, 99.03, 121.52, 144.62, 168.84, 194.15, 220.64, 247.29, 275.59\}.$$

From **Eq. (3)**, the weighted accumulation sequence of $X^{(r)}$ is

$$X^{(r\lambda)} = \{18.99, 39.95, 61.8, 84.42, 107.78, 132.7, 158.38, 185.26, 213.36, 242.77, 272.47, 303.84\}.$$

The coefficient a and b are solved by **Eq. (6),**

$$\begin{bmatrix} a \\ b \end{bmatrix} = (B^T B)^{-1} B^T Y = \begin{bmatrix} -0.0405 \\ 19.7846 \end{bmatrix},$$

where

$$B = \begin{bmatrix} -29.4700 & 1 \\ -50.8706 & 1 \\ & 1 \\ -73.1047 \vdots -288.1560 & \vdots \\ & 1 \end{bmatrix}, Y = \begin{bmatrix} 20.9601 \\ 21.8410 \\ 22.6261 \\ \vdots \\ 31.3759 \end{bmatrix}.$$

From Eq. (5),

$$\widehat{x}^{(r\lambda)}(k) = \left(x^{(0)}(1) - \frac{b}{a}\right)e^{-a(k-1)} + \frac{b}{a} = (18.99 + 488.5086)e^{0.0405(k-1)} - 488.5086$$

and

$$\widehat{X}^{(r\lambda)} = \{18.99, 39.97, 61.81, 84.55, 108.24, 132.9, 158.58, 185.32, 213.17, 242.17, 272.37, 303.81,$$

$$336.56, 370.65, 406.16, 443.13, 481.63\}$$

From Eq. (8), the weighted reduction sequence

$$\widehat{X}^{(r)} = \{18.99, 38.0, 57.66, 78.14, 99.47, 121.67, 144.8, 168.88, 193.95, 220.06, 247.25, 275.57,$$

$$305.05, 335.75, 367.72, 401.01, 435.68\}$$

From Eq. (9),

$$\widehat{X}^{(0)} = \{18.99, 19.43, 20.36, 21.39, 22.43, 23.49, 24.59, 25.71, 26.87, 28.08, 29.32, 30.62, 31.96,$$

$$33.36, 34.82, 36.33, 37.90\}$$

From Eq. (10), MAPE $= \frac{1}{n}\sum_{k=1}^{n}\frac{\left|\widehat{x}^{(0)}(k)-x^{(0)}(k)\right|}{x^{(0)}(k)} \times 100\% = 0.71\%$.

The results are shown in Table 13. Table 14 shows the corresponding accuracy levels of each index.

The formula of grey absolute correlation degree is shown in equations (23)–(26).

$$\varepsilon_{0i} = \frac{1 + |s_0| + |s_i|}{1 + |s_0| + |s_i| + |s_i - s_0|}, \tag{23}$$

$$|s_0| = \left|\sum_{k=2}^{n-1} x_0^0(k) + \frac{1}{2}x_0^0(n)\right|, \tag{24}$$

$$|s_i| = \left|\sum_{k=2}^{n-1} x_i^0(k) + \frac{1}{2}x_i^0(n)\right|, \tag{25}$$

Table 13
The fitting values of the CGM(1,1) mode (Units: Exajoules).

Year	Oil		Natural gas		Coal Natural gas	
	Actual	fitting value	Actual	fitting values	Actual	fitting values
2010	18.99	18.99	3.92	3.92	73.22	73.22
2011	19.41	19.43	4.87	4.87	79.71	80.05
2012	20.36	20.36	5.43	5.30	80.71	80.42
2013	21.27	21.39	6.19	5.89	82.43	80.76
2014	22.11	22.43	6.78	6.53	82.48	81.09
2015	23.80	23.50	7.01	7.24	80.92	81.40
2016	24.56	24.59	7.54	8.03	80.19	81.70
2017	25.86	25.71	8.69	8.91	80.56	82.00
2018	27.12	26.87	10.22	9.89	81.05	82.29
2019	28.49	28.08	11.10	10.97	81.70	82.58
2020	28.74	29.32	12.12	12.16	82.38	82.86
2021	30.60	30.62	13.63	13.49	86.17	83.15
MAPE(%)	0.71		2.44		1.29	
a	−0.0405		−0.1036		−0.0031	
b	19.7846		4.5905		79.6255	
r	0.9685		0.9993		0.9958	
λ	0.1037		0.1023		0.0021	

Table 14
Precision inspection reference table.

Accuracy Grade(Level)	Critical value NGM(1,1,k) CGM(1,1)			
	Relative Error(%)	Correlation degree	Mean square deviation ratio(%)	Small error probability
1	1	0.90	35	0.95
2	5	0.80	50	0.80
3	10	0.70	65	0.70
4	20	0.60	80	0.60

Next, the CGM(1,1) model are used to verify whether the model meets the requirements. In general, MAPE, correlation degree, mean square error ratio, small error probability are used as indicators.

$$|s_i - s_0| = \left| \sum_{k=2}^{n-1} \left(x_i^0(k) - x_0^0(k) \right) + \frac{1}{2}(x_i^0(n) - x_0^0(n)) \right|. \tag{26}$$

By calculating the absolute correlation degree, it is found that the fitting absolute correlation degree of oil, coal and natural gas is 0.998, 0.999, 0.997, respectively.

First, Eq. (27) is used to find the mean and variance.

$$\bar{x} = \frac{1}{n} \sum_{k=1}^{n} x^{(0)}(k), s_1^2 = \frac{1}{n} \sum_{k=1}^{n} \left(x^{(0)}(k) - \bar{x} \right)^2. \tag{27}$$

Next, the mean and variance of the residual sequence are calculated by Eq. (28).

$$\bar{\varepsilon} = \frac{1}{n} \sum_{k=1}^{n} \varepsilon(k), s_2^2 = \frac{1}{n} \sum_{k=1}^{n} \left(\varepsilon(k) - \bar{\varepsilon} \right)^2. \tag{28}$$

Then the mean square error ratio is shown in Eq. (29).

$$C = \frac{s_2}{s_1}. \tag{29}$$

Through calculation, the variance ratio of oil is close to 0, that of natural gas is 0.0478 and that of coal is 0.3177.
The calculation formula of small error probability is

$$p = P(|\varepsilon(k) - \bar{\varepsilon}| < 0.6745 S_1). \tag{30}$$

From Eq. (30), the small error probability of oil and natural gas is 1, and that of coal is 0.92. Table 15 summarizes the results of these indicators.

From Tables 14 and 15, the fitting MAPE value of oil meets the first-level accuracy standard, while that of natural gas and coal meets the second-level accuracy standard. The absolute correlation degree of the three energy sources all meet the first-level accuracy standard. The mean square error ratio of oil and nature gas meets the first-level accuracy standard, while the probability of small error of coal meets the second-level accuracy standard. To sum up, this model has good performance, and the results are summarized in Table 16. The development trends are summarized in Fig. 5.

Fig. 5 shows that China's oil consumption will continue to grow. By 2026, the consumption of oil will exceed 37 EJ, that's an increase of nearly 24% compared to 2021, which is inseparable from the rapid development of China. But oil reserves are limited, if uncontrolled exploitation, oil resources will soon face depletion. As the current alternative energy sources have not been popularized, the consumption of oil as a rigid demand of human beings will continue to increase. At present, the focus of people's task is to vigorously develop and promote new energy to gradually replace oil. The consumption of natural gas is expected to exceed 22 EJ by 2026, that's an increase of more than 60% compared to 2019. Coal consumption in China, will maintain a small upward trend and gradually level off.

Oil is mainly used as fuel and power energy. With the continuous development of social, oil consumption is also increasing year by year. China should speed up the adjustment of industrial, develop clean energy instead of oil consumption. At the same time, improve the petroleum technology. China is a big coal resource country, and coal consumption is huge in the development of the whole society. In order to improve people's ecological environment, China's coal provinces should develop diversified economy on the basis of their own development. According to regional characteristics, the construction of diversified pillar industries, steadily promote the

Table 15
The fitting index of three kinds of energy.

energy	MAPE	Absolute Correlation	Mean square deviation ratio	Small error probability
Oil	0.71%	0.998	0	1
Nature gas	2.44%	0.997	0.0478	1
Coal	1.29%	0.999	0.3177	0.917

Table 16
Prediction value of the CGM(1,1) model (Units: Exajoules).

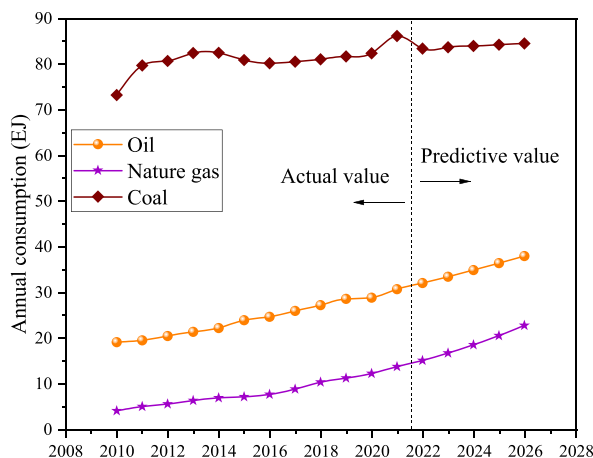| Year	Oil	Nature gas	Coal
2022	31.96	14.97	83.43
2023	33.36	16.60	83.71
2024	34.82	18.41	83.99
2025	36.33	20.43	84.27
2026	37.90	22.66	84.55

Fig. 5. Three energy consumption trends.

transformation of industrial structure. The government should improve talent management and promote the efficient and clean utilization of coal. As a clean energy source, natural gas consumption in China is increasing. To some extent, it reflects the work of the government in recent years. In the future, the government should continue to strengthen the implementation of relevant laws.

6. Conclusion and future study

This thesis studies a composite summation model, which is then used to study the consumption of non-renewable energy. Two adjustable parameters are introduced in GM(1,1) model. The optimal parameters are determined by the PSO algorithm. Improve the performance by giving greater weight to new information. Then, it studies the consumption of non-renewable energy in China. The results show that oil and natural gas will maintain a larger growth trend in terms of consumption in the future. By 2026, oil consumption will exceed 37 EJ. Natural gas consumption will exceed 22 EJ. However, coal will maintain a small growth trend in terms of consumption and gradually tend to level off.

Future work can be considered from the following three aspects. Firstly, this model is combined with other optimization algorithms. Secondly, the accuracy of the multi-variable model is further improved by considering other influencing factors comprehensively. Thirdly, the good accuracy of the model can be studied in more fields.

Author contribution statement

Jianlong Guo: Conceived and designed the experiments; Performed the experiments; Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper.

Lifeng Wu: Conceived and designed the experiments; Contributed reagents, materials, analysis tools or data.

Yali Mu: Analyzed and interpreted the data; Wrote the paper.

Data availability statement

Data will be made available on request.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This study was supported by the National Natural Science Foundation of China (Program No. 72073064), the "333 distinguished Talents Project" Foundation of Jiangsu Province in China (Grant No. BRA2018070).

References

[1] P. Zhao, Z. Lu, J. Fang, et al., Determinants of renewable and non-renewable energy demand in China, Struct. Change Econ. Dynam. 54 (2020) 202–209.
[2] R.L. Ibrahim, O.O. Julius, I.C. Nwokolo, et al., The role of technology in the non-renewable energy consumption-quality of life nexus: insights from sub-Saharan African countries, Econ. Change Restruct. 55 (2022) 257–284, https://doi.org/10.1007/s10644-020-09312-6.
[3] T. Peng, Z. Ning, H. Yang, Embodied CO_2 in China's trade of harvested wood products based on an MRIO model, Ecol. Indicat. 137 (2022), 108742, https://doi.org/10.1016/j. ecolind.2022. 108742.
[4] X. Li, W. Wang, H. Zhang, et al., Dynamic baselines depending on REDD+ payments: a comparative analysis based on a system dynamics approach, Ecol. Indicat. 140 (2022), 108983, https://doi.org/10.1016/j.ecolind.2022.108983.
[5] T.S. Adebayo, R.L. Ibrahim, E.B. Agyekum, et al., The environmental aspects of renewable energy consumption and structural change in Sweden: a new perspective from wavelet-based granger causality approach, Heliyon 8 (9) (2022), e10697, https://doi.org/10.1016/j.heliyon. 2022.e10697.
[6] S. Liu, Y. Lin, S. Liu, et al., Introduction to grey systems theory, Grey Syst. Theor. Appl. 68 (2011) 1–18, https://doi.org/10.1007/978-3-642-16158-2_1.
[7] S. Perla, R. Bisoi, P. Dash, A hybrid neural network and optimization algorithm for forecasting and trend detection of Forex market indices, Decision Analytics Journal 6 (2023), 100193, https://doi.org/10.1016/j.dajour.2023.100193.
[8] A. Kurani, P. Doshi, A. Vakharia, et al., A comprehensive comparative study of artificial neural network (ANN) and support vector machines (SVM) on stock forecasting, Annals of Data Science 10 (1) (2023) 183–208, https://doi.org/10.1007/s40745-021-00344-x.
[9] A. Izadi, M. Shahafve, P. Ahmadi, Neural network genetic algorithm optimization of a transient hybrid renewable energy system with solar/wind and hydrogen storage system for zero energy buildings at various climate conditions, Energy Convers. Manag. 260 (2022), 115593.
[10] W. Li, N. Sengupta, P. Dechent, et al., Online capacity estimation of lithium-ion batteries with deep long short-term memory networks, J. Power Sources 482 (2021), 228863.
[11] X. Yuan, L. Li, Y. Wang, Nonlinear dynamic soft sensor modeling with supervised long short-term memory network, IEEE Trans. Ind. Inf. 16 (5) (2019) 3168–3176. https://doi: 10.1109/TII.2019.2902129.
[12] G. Ciulla, A. D'Amico, Building energy performance forecasting: a multiple linear regression approach, Appl. Energy 253 (2019), 113500, https://doi.org/10.1016/j.apenergy.2019.113500.
[13] K. Joki, A.M. Bagirov, N. Karmitsa, et al., Clusterwise support vector linear regression, Eur. J. Oper. Res. 287 (1) (2020) 19–35, https://doi.org/10.1016/j.ejor.2020.04.032.
[14] M. Flores-Sosa, E. Avilés-Ochoa, J.M. Merigó, et al., The OWA operator in multiple linear regression, Appl. Soft Comput. 124 (2022), 108985, https://doi.org/10.1016/j.asoc.2022. 108985.
[15] W. Zhou, X. Wu, S. Ding, et al., Predictive analysis of the air quality indicators in the Yangtze River Delta in China: an application of a novel seasonal grey model, Sci. Total Environ. 748 (2020), 141428, https://doi.org/10.1016/j.scitotenv.2020.141428.
[16] L. Tu, Y. Chen, An unequal adjacent grey forecasting air pollution urban model, Appl. Math. Model. 99 (2021) 260–275, https://doi.org/10.1016/j. apm.2021.06.025.
[17] X. Long, S. Wu, J. Wang, et al., Urban water environment carrying capacity based on VPOSR-coefficient of variation-grey correlation model: a case of Beijing, China, Ecol. Indicat. 138 (2022), 108863, https://doi.org/10.1016/j.ecolind.2022.108863.
[18] Y. Huang, Y. Cai, Y. Xie, et al., An optimization model for water resources allocation in Dongjiang River Basin of Guangdong-Hong Kong-Macao Greater Bay Area under multiple complexities, Sci. Total Environ. 820 (2022), https://doi.org/10.1016/j.scitotenv.2022. 153198.
[19] L. Ye, N. Xie, A. Hu, A novel time-delay multivariate grey model for impact analysis of CO_2 emissions from China's transportation sectors, Appl. Math. Model. 91 (2021) 493–507, https://doi.org/10.1016/j.apm.2020.09.045.
[20] R. Rajamoorthy, G. Arunachalam, P. Kasinathan, et al., A novel intelligent transport system charging scheduling for electric vehicles using Grey Wolf Optimizer and Sail Fish Optimization algorithms, Energy Sources, Part A Recovery, Util. Environ. Eff. 44 (2) (2022) 3555–3575.
[21] S. Ding, Z. Tao, H. Zhang, et al., Forecasting nuclear energy consumption in China and America: an optimized structure-adaptative grey model, Energy 239 (2022) 121928.
[22] J. Guo, W. Liu, L. Tu, et al., Forecasting carbon dioxide emissions in BRICS countries by exponential cumulative grey model, Energy Rep. 7 (2021) 7238–7250.
[23] J. Ye, Y. Dang, B. Li, Grey-Markov prediction model based on background value optimization and central-point triangular whitenization weight function, Commun. Nonlinear Sci. Numer. Simulat. 54 (2018) 320–330, https://doi.org/10.1016/j.cnsns. 2017.06.004.
[24] Z. Wang, Y. Dang, S. Liu, et al., The optimization of background value in GM(1,1) model, J. Grey Syst. 10 (2) (2007) 69–74.
[25] L. Wu, S. Liu, Z. Fang, et al., Properties of the GM(1,1) with fractional order accumulation, Appl. Math. Comput. 252 (2015) 287–293, https://doi.org/10.1016/j.amc. 2014.12.014.
[26] L. Wu, H. Zhao, Discrete grey model with the weighted accumulation, Soft Comput. 23 (2019) 12873–12881, https://doi.org/10.1007/s00500-019-03845-3.
[27] B. Zeng, H. Duan, Y. Bai, et al., Forecasting the output of shale gas in China using an unbiased grey model and weakening buffer operator, Energy 151 (2018) 238–249, https://doi.org/10.1016/j. energy.2018.03.045.
[28] L. Liu, Y. Chen, L. Wu, The damping accumulated grey model and its application, Commun. Nonlinear Sci. Numer. Simulat. 95 (2021), 105665, https://doi.org/10.1016/j. cnsns.2020.105665.
[29] Z. Jia, Z. Zhou, H. Zhang, et al., Forecast of coal consumption in Gansu Province based on Grey-Markov chain model, Energy 199 (2020), 117444, https://doi.org/10.1016/j.energy. 2020.117444.
[30] D. Zhou, A. Al-Durra, K. Zhang, et al., A robust prognostic indicator for renewable energy technologies: a novel error correction grey prediction model, IEEE Trans. Ind. Electron. 66 (12) (2019) 9312–9325. https://doi: 10.1109/TIE.2019.2893867.
[31] F. Van den Bergh, A.P. Engelbrecht, A study of particle swarm optimization particle trajectories, Inf. Sci. 176 (8) (2006) 937–971, https://doi.org/10.1016/j. ins.2005.02.003.
[32] G.W. Stewart, On the perturbation of pseudo-inverses, projections and linear least squares problems, SIAM Rev. 19 (4) (1977) 634–662, https://doi.org/10.1137/1019104.
[33] G.H. Golub, H. Zha, Perturbation analysis of the canonical correlations of matrix pairs, Lin. Algebra Appl. 210 (1994) 3–28, https://doi.org/10.1016/0024-3795(94)90463-4.
[34] L. Wu, N. Li, Y. Yang, Prediction of air quality indicators for the Beijing-Tianjin-Hebei region, J. Clean. Prod. 196 (2018) 682–687, https://doi.org/10.1016/j.jclepro.2018.06.068.
[35] B. Zeng, C. Li, Forecasting the natural gas demand in China using a self-adapting intelligent grey model, Energy 112 (2016) 810–825, https://doi.org/10.1016/j.energy.2016.06.090.
[36] W. Xie, C. Liu, W. Wu, et al., Continuous grey model with conformable fractional derivative, Chaos, Solitons & Fractals 139 (2020), 110285, https://doi.org/10.1016/j.chaos.2020.110285.
[37] Q. Shen, Q. Shi, T. Tang, et al., A novel weighted fractional GM(1,1) model and its applications, Complexity 2020 (2020) 1–20, https://doi.org/10.1155/2020/6570683.