Forecasting China’s energy intensity by using an improved \(DVCGM\) (1, N) model considering the hysteresis effect

Zhaosu Meng and Xiaotong Liu  
Ocean University of China, Qingdao, China  
Kedong Yin  
School of Management Science and Engineering, Shandong University of Finance and Economics, Jinan, China and  
Marine Development Studies Institute, Ocean University of China, Qingdao, China  
Xuemei Li  
Ocean University of China, Qingdao, China, and  
Xinchang Guo  
School of Marxism, Ocean University of China, Qingdao, China

Abstract

Purpose – The purpose of this paper is to examine the effectiveness of an improved dummy variables control grey model (\(DVCGM\)) considering the hysteresis effect of government policies in China’s energy intensity (EI) forecasting.

Design/methodology/approach – Energy consumption is considered as an important driver of economic development. China has introduced policies those aim at the optimization of energy structure and EI. In this study, EI is forecasted by an improved \(DVCGM\), considering the hysteresis effect of energy-saving policies of the government. A nonlinear optimization method based on particle swarm optimization (PSO) algorithm is constructed to calculate the hysteresis parameter. A one-step rolling mechanism is applied to provide input data of the prediction model. Grey model (GM) (1, N), \(DVCGM\) (1, N) and ARIMA model are applied to test the accuracy of the improved \(DVCGM\) (1, N) model prediction.

Findings – The results show that the improved \(DVCGM\) provides reliable results and works well in simulation and predictions using multivariable data in small sample size and time-lag virtual variable. Accordingly, the improved \(DVCGM\) notes the hysteresis effect of government policies and significantly improves the prediction accuracy of China’s EI than the other three models.

Originality/value – This study estimates the EI considering the hysteresis effect of energy-saving policies in China by using an improved \(DVCGM\). The main contribution of this paper is to propose a model to estimate EI, considering the hysteresis effect of energy-saving policies and improve forecasting accuracy.

Keywords Grey system theory, Energy intensity forecasting, Grey prediction model, Hysteresis effect

Paper type Research paper
1. Introduction

Energy consumption is considered as an important driver of economic growth. Energy intensity (EI), which is the ratio of total energy consumption to gross domestic product (GDP) of a country or region in a certain period of time, measures the performance of energy utilization. According to the energy section of the National Bureau of Statistics (NBS), China’s largest energy consumption is 4,640 million tons of standard coal, and China has become the world largest energy consumer. Entering a new normal period, the government focuses on a crucial rebalancing and diversifying economy, with higher requirements for sustainable development. In the 13th Five-Year Plan for economic and social development, a total reduction of 15% has been set as the energy performance target. It is imperative to estimate and predict the EI to evaluate energy conservation and emission reduction. However, due to the complexity and dynamics of social economy, the implementation of energy conservation and emission reduction policies will not immediately reduce EI, and there is often a certain time lag. This inevitable time-lag should be taken into account to estimate and forecast the EI in China accurately.

Numerous research studies have studied the influencing factors and driving forces of EI through econometrics methods such as cointegration analysis, metrology and decomposition analysis and scenario analysis. Zhu et al. (2015) applied cointegration analysis method on the large and state-owned enterprises and found that energy-saving regulations in China are one of the most important factors in reducing aggregate EI. Karimu et al. (2017) studied the EI and convergence of Swedish industry by combining metrology and decomposition analytical methods. Ma and Yu (2017) used panel data model to discuss the driving factors which lead to EI decline. It is revealed that industrial structure, energy conservation regulations and EI are closely related. Tan et al. (2018) used index decomposition analysis and production decomposition analysis methods to analyze the factors which are related to the decline in the EI and pointed out that technology improvement effect is the most significant factor. In EI forecasting, Pao et al. (2012) used improved grey models (GMs) to predict China’s CO2 emissions, energy consumption and economic growth. Dong et al. (2018) estimated the driving force of regional EI in China and forecasted the potential of regional energy conservation with scenario analysis. Wu et al. (2018) used a new multivariable GM to predict energy consumption in Shandong Province.

The grey system theory is an interdisciplinary theory proposed by Deng (1982) and the grey prediction method applies well to small size data forecasting. As an important part of grey prediction theory, the GM (1, N) model is the basic model of multivariable grey system modeling approach. In recent years, numerous scholars have thoroughly discussed the model parameters optimization (Tien, 2005, 2010), the model accuracy improvement (Tien, 2011; Wang et al., 2016) and the expansion of the GM (Guo et al., 2013; Kose and Tasci, 2015, 2019; Ding et al., 2017).

Based on grey multivariable model with time-lagged system, Zhai et al. (1996) introduced the lag term into the GM (1, 2) model and determined the delay parameters with the goal of minimizing the modeling error. Hao (2011) used grey correlation analysis to determine time-lag period between variables and then on this basis to establish forecast model of GM (1, N). Zhang et al. (2015) constructed a time-delay multivariable discrete GM, DDGM (1, N) model, by introducing a time-delay control factor and solved the time-delay parameters by using the grey dimensionally expanding identification method, which obtained a good application effect. Ma and Yu (2017) used a novel time-delay multivariable GM to predict the natural gas consumption in China. Dang et al. (2017) constructed the discrete delay grey multivariable DDGMD (1, N) model by introducing the driving information control adjustment coefficient T and the action coefficient λ and solved the coefficient respectively by grey dimension expansion method and particle swarm optimization (PSO) algorithm. Xiong (2019) built a multivariable time-delay discrete MGM
(1, m, t) model and studied the mechanism of modeling and the process of modeling, and the calculation method of time delay is given. The hysteresis effect is discussed by an example to verify the validity of the model.

*GM* (1, N) model, in spite of successfully applying in various fields, sometimes ignores the influence of virtual variables such as policy on the main system in practical applications. Zhang (2016) considered the influence of dummy variables on system behavior variables and built a discrete multivariable prediction model based on dummy variables, which further expanded the application scope of the model. Ding *et al.* (2018) introduced the dummy variables into the *GM* (1, N) model, gave the concrete model construction method in mathematics and verified the effectiveness of the new model with cases.

Generally speaking, we have abundant literature discussing the influencing factors of EI and its forecasting. Furthermore, there are some practices targeting the time-lag phenomenon using grey theory. The existing literature has explored and studied the *GM* from multiple angles, but there is still room for the *GM* to expand in the combination of dummy variables and time-delay systems. As the common *GM* (1, N) model does not take into account the time delay between variables, and there are dummy variables in the system which are difficult to be measured by quantity, our optimized *DVCGM* (1, N) model, which is short for dummy variables control grey model of N variables, incorporates the above two features and improves prediction accuracy to a satisfying level. It is practical significant because the time delay of dummy variables, i.e. government policy, is often seen in real world but is difficult to measure. The optimized *DVCGM* (1, N) model takes these variables into consideration, solves the practical problem and expands the grey system theory and the grey prediction method system, which improves the accuracy of grey prediction model.

This paper’s main interest is to estimate and forecast EI by considering the influence of government policies and to test the accuracy of the improved *GM* through a comparison study. The main contribution of the study to the literature is to consider hysteresis effect and increase the forecast accuracy. It is imperative for optimizing energy structure, improving energy utilization efficiency and ensuring energy security. The results show that the improved *GM* produces better results than the other three conventional models. The rest of the paper is organized as follows. Section 2 briefly introduces grey theory and multivariable grey prediction models. A nonlinear optimization method based on PSO algorithm is constructed to calculate the hysteresis parameter in the improved model. In Section 3, EI is estimated by considering the hysteresis effect of energy-saving policies by an improved *DVCGM* (1, N) model. We compare the results with the other two *GMS* and one econometric model. The improved *DVCGM* (1, N) model has the best performance in the estimation comparison and is applied to forecast future EI. Section 4 is the conclusion of the study with the limitation and future path.

2. Methodology

The three grey prediction models used in this paper are interrelated and in a progressive order. The *GM* (1, N) model is a traditional multivariable grey prediction model. The *DVCGM* (1, N) model introduces virtual variables, taking into account the influence of policy and other factors on the basis of the *GM* (1, N) model. Time-delay parameter is introduced in the improved *DVCGM* (1, N) model, considering the hysteresis effects of historical variables, which further enriches the existing grey prediction theory.

2.1 *GM* (1, N) model

Grey prediction model can be regarded as two levels of work. At a lower level, the original sequence produces the sequence of generation by one accumulative generation (1-AGO), and
then it forms the sequence of mean generation of consecutive neighbors; similarly, the sequence of influencing factors generates the sequence by 1-AGO; Constructing B and Y matrix and calculating system parameters through ordinary least squares (OLS) regression. Once the system parameters a and b are determined, we can obtain the time response function (TRF) by solving the whitenization equation. At a higher level, the continuous differential equation with initial values is used as the reflection equation, and the discrete data are mapped to a manageable function, which is further restored to the TRF as the simulation basis. Typical procedures can be described briefly by the program in Figure 1.

**Definition 1.** The original sequence is

\[ X^{(0)}_1 = (x^{(0)}_1(1), x^{(0)}_1(2), \ldots, x^{(0)}_1(n)) \]

\( X^{(1)}_1 \) is the 1 - AGO sequence from \( X^{(0)}_1 \), and \( X^{(1)}_i \) is the sequences of the relevant factors, where

\[ X^{(1)}_i(k) = \sum_{m=1}^{k} x^{(0)}_i(m) \]

\( Z^{(1)}_1 \) is the sequence of mean generation of consecutive neighbors from \( X^{(1)}_1 \), where

\[ Z^{(1)}_1(k) = 0.5X^{(1)}_1(k-1) + 0.5X^{(1)}_1(k), \quad (k = 2, 3, \ldots, n) \]

**Definition 2.** Denote Eq (1) as the definition formula of GM (1, N) model:

\[ X^{(0)}_1(k) + aZ^{(1)}_1(k) = \sum_{i=2}^{n} b_i X^{(1)}_i(k) \]  

(1)

where \( k = 2, 3, n \), a is the development coefficient and \( b_1, b_2, \ldots, b_n \) are the grey input coefficients obtained by the least squares method. To determine these coefficients, the matrix B and \( Y_N \) are defined as follows:

\[
Y_N = \begin{bmatrix} x^{(0)}_1(2) \\ x^{(0)}_1(3) \\ \vdots \\ x^{(0)}_1(n) \end{bmatrix}, B = \begin{bmatrix} -z^{(1)}_1(2) & X^{(1)}_2(2) & \cdots & X^{(1)}_n(2) \\ -z^{(1)}_1(3) & X^{(1)}_2(3) & \cdots & X^{(1)}_n(3) \\ \vdots & \vdots & \ddots & \vdots \\ -z^{(1)}_1(n) & X^{(1)}_2(n) & \cdots & X^{(1)}_n(n) \end{bmatrix}
\]

The values of the coefficients a and \( b_1, b_2, \ldots, b_n \) can be determined by the following equation:

\[
\hat{P} = \begin{bmatrix} a \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = (B^T B)^{-1} B^T Y_n
\]

**Definition 3.** Let Eq (2) be defined as the differential equation (or) called grey reflection equation:

\[
\frac{dx^{(0)}_1}{dt} + ax^{(0)}_1(t) = \sum_{i=1}^{n} b_i x^{(1)}_i(t)
\]  

(2)
Figure 1. Principal diagram of $GM (1, N)$ model
After determining the coefficients of $a$ and $b_1, b_2, \cdots, b_n$, the differential equation of the $GM$ can be determined by Eq (2). The solution of the above differential equation is as follows:

$$\hat{X}_1^{(1)}(k + 1) = \left\{ X_1^{(0)} - \left(\frac{1}{a}\right) \sum_{i=2}^{n} \hat{b}_{i-1}X_i^{(1)}(k) \right\} \times e^{\hat{a}k} + \left(\frac{1}{a}\right) \sum_{i=2}^{n} \hat{b}_{i-1}X_i^{(1)}(k) \quad (3)$$

where $\hat{X}_1^{(1)}(k)$ is the prediction of the AGO of the original sequence. By considering that the estimation of the first element of the first AGO of a sequence is equal to the first element of the sequence, the following relation is determined: $\hat{X}_1^{(1)}(1) = X_1^{(0)}(1)$.

Finally, in order to predict the elements of the original sequence, the inverse accumulated generating operation should be performed. Therefore, the predicted values can be determined as follows:

$$\hat{x}_1^{(0)}(k) = \hat{x}_1^{(1)}(k) - \hat{x}_1^{(1)}(k - 1), k \geq 2,$$

where $\hat{x}_1^{(0)}(n)$ is an estimation of the original sequence, which is simulation values, $\hat{x}_1^{(0)}(n + 1), \hat{x}_1^{(0)}(n + 2), \cdots$ are predictive values.

2.2 DVCGM (1, N) model

Traditional $GM$ (1, N) model ignores the influence of virtual variables, which will inevitably lead to significant errors in practical applications. Therefore, it is necessary to construct a new multivariable predictive model with virtual variable control, based on the traditional $GM$ (1, N) model, i.e. the DVCGM (1, N) model. The modeling steps for DVCGM (1, N) can be illustrated in Figure 2.

Definition 4. The original sequence is

$$X_i^{(0)} = (x_i^{(0)}(1), x_i^{(0)}(2), \cdots, x_i^{(0)}(n)).$$

Virtual variable sequence is

$$D_j^{(0)} = (d_j^{(0)}(1), d_j^{(0)}(2), d_j^{(0)}(n)), \cdots, d_j^{(0)}(n) = 0 \text{ or } 1.$$

$X_i^{(1)}$ and $D_j^{(1)}$ are the 1-AGO sequence, $X_i^{(0)}$ is the behavior sequence of the system, $X_i^{(1)}(i = 2, \cdots, M)$ and $D_j^{(1)}(j = M + 1, \cdots, N)$ is the driving factor sequence. Then $GM$ (1, N) model with dummy variable can be expressed as:

$$x_1^{(0)}(k) + ax_1^{(1)}(k) = \sum_{i=2}^{M} b_i x_i^{(1)}(k) + \sum_{j=M+1}^{N} b_j d_j^{(1)}(k) \quad (4)$$

named DVCGM (1, N) model (dummy variables control grey model of $N$ variables). $Z_1^{(1)}$ is the sequence of mean generation of consecutive neighbors from $X_1^{(1)}$ where,

$$Z_1^{(1)}(k) = 0.5X_1^{(1)}(k - 1) + 0.5X_1^{(1)}(k), (k = 2, 3, \ldots, n)$$

where $\sum_{i=2}^{M} b_i x_i^{(1)}(k)$ is independent quantization variable driver, $\sum_{j=M+1}^{N} b_j d_j^{(1)}(k)$ is the dummy variable driver. If virtual variables are not considered or $b_j=0$, the model can be transformed to a traditional $GM$ (1, N) model.
Hysteresis effect of China’s EI forecasting

Figure 2. Principal diagram of DVCGM (1, N) mode
After determining the coefficients of \( a \) and \( bN \), the differential equation of the GM can be determined by Eq (5). The solution of the above differential equation is as follows:

\[
\begin{align*}
\dot{x}^{(1)}_i(t) &= e^{-at} \left\{ X^{(0)}_i - t \left[ \sum_{i=2}^{M} b_i x^{(1)}_i(0) + \sum_{j=M+1}^{N} b_j d^{(1)}_j(0) \right] 
+ \sum_{i=2}^{N} \int \left[ \sum_{i=2}^{M} b_i x^{(1)}_i(t) + \sum_{j=M+1}^{N} b_j d^{(1)}_j(t) \right] e^{at} dt \right\}
\end{align*}
\]

Finally, in order to predict the elements of the original sequence, the inverse accumulated generating operation should be performed. Therefore, the predicted values can be determined as follows:

\[
\tilde{x}^{(0)}_i(k) = \tilde{x}^{(1)}_i(k) - \tilde{x}^{(1)}_i(k-1), \quad k \geq 2,
\]

where \( \tilde{x}^{(0)}_i(n) \) is an estimation of the original sequence, which is simulation values, \( \tilde{x}^{(0)}_i(n+1), \tilde{x}^{(0)}_i(n+2), \cdots \) are predictive values.

### 2.3 Improved DVCGM (1, N) model

The classic multivariable grey prediction models, such as traditional GM (1, N) and DVCGM (1, N) models, can reflect the influences of current driving-variables on the present system.
behavior and innately ignore the hysteresis effect of historical variables. Therefore, an improved DVCGM (1, N) model is proposed, integrating these above prediction model.

2.3.1 Construction of the improved DVCGM (1, N) model. In this section, the hysteresis parameter \( \lambda_i \) is innovatively introduced into the DVCGM (1, N) model to improve the prediction accuracy. Supported by PSO algorithm, detailed process and algorithm can be described as following:

**Theorem 2.** Assuming \( X_i^{(0)}, X_i^{(1)}, D_j^{(0)}(k), D_j^{(1)}(k) \)

As mentioned in **Definition 4.** The parameter column of the model is

\[
\hat{b} = [a, b_2, \cdots, b_N]^T.
\]

The matrix B and \( Y_N \) are defined as follows:

\[
Y_N = \begin{bmatrix}
X_1^{(0)}(2) \\
X_1^{(0)}(3) \\
\vdots \\
X_1^{(0)}(n)
\end{bmatrix},
\]

\[
B = \begin{bmatrix}
-Z_1^{(1)}(2), \sum_{j=1}^{2} \lambda_2^{2-j} x_2^{(1)}(j), \cdots, \sum_{j=1}^{2} \lambda_M^{2-j} x_M^{(1)}(j), & \sum_{j=1}^{2} \lambda_{M+1}^{2-j} d_{M+1}^{(1)}(j) & \cdots & \sum_{j=1}^{2} \lambda_N^{2-j} d_N^{(1)}(j) \\
-Z_1^{(1)}(3), \sum_{j=1}^{3} \lambda_2^{3-j} x_2^{(1)}(j), \cdots, \sum_{j=1}^{3} \lambda_M^{3-j} x_M^{(1)}(j), & \sum_{j=1}^{3} \lambda_{M+1}^{3-j} d_{M+1}^{(1)}(j) & \cdots & \sum_{j=1}^{3} \lambda_N^{3-j} d_N^{(1)}(j) \\
\vdots & \vdots & \ddots & \vdots \\
-Z_1^{(1)}(n), \sum_{j=1}^{n} \lambda_2^{n-j} x_2^{(1)}(j), \cdots, \sum_{j=1}^{n} \lambda_M^{n-j} x_M^{(1)}(j), & \sum_{j=1}^{n} \lambda_{M+1}^{n-j} d_{M+1}^{(1)}(j) & \cdots & \sum_{j=1}^{n} \lambda_N^{n-j} d_N^{(1)}(j)
\end{bmatrix}
\]

The least square estimation of the parameter column satisfies the following requirements:

1. When \( n = N+1, \hat{b} = B^{-1}Y, |B| \neq 0; \)
2. When \( n > N+1, \hat{b} = (B^TB)^{-1}B^TY, \left| B^TB \right| \neq 0; \)
3. When \( n < N+1, \hat{b} = B^T(B^TB)^{-1}Y, \left| B^TB \right| \neq 0; \)

Proof: Substitute \( k = 2, 3, \ldots, n \) into the model, you can get

\[
x_1^{(1)}(2) = -ax_1^{(1)}(2) + \sum_{i=2}^{M} \sum_{j=1}^{k} b_i \lambda_i^{k-j} x_i^{(1)}(2) + \sum_{q=M+1}^{N} \sum_{j=1}^{k} b_q \lambda_q^{k-j} d_q^{(1)}(2)
\]

\[
x_1^{(1)}(3) = -ax_1^{(1)}(3) + \sum_{i=2}^{M} \sum_{j=1}^{k} b_i \lambda_i^{k-j} x_i^{(1)}(3) + \sum_{q=M+1}^{N} \sum_{j=1}^{k} b_q \lambda_q^{k-j} d_q^{(1)}(3)
\]

\[
x_1^{(1)}(n) = -ax_1^{(1)}(n) + \sum_{i=2}^{M} \sum_{j=1}^{k} b_i \lambda_i^{k-j} x_i^{(1)}(n) + \sum_{q=M+1}^{N} \sum_{j=1}^{k} b_q \lambda_q^{k-j} d_q^{(1)}(n)
\]

That is, by the least square method, \( Y = B\hat{b} \)
(1) When \( n = N+1 \), B has an inverse matrix, the equations have a unique solution, we can get \( \hat{b} = B^{-1}Y \), \(|B| \neq 0\);

(2) When \( n > N+1 \), B is column full rank, the full rank decomposition of B is \( B = DC \). Then the generalized inverse matrix of B can be obtained:

\[
B^+ = C^T \left( C C^T \right)^{-1} (D^T D)^{-1} D^T \beta = C^T \left( C C^T \right)^{-1} (D^T D)^{-1} D^T Y;
\]

Because B is a full rank matrix, C can be taken as a unit matrix, \( B = D \), so

\[
\hat{b} = C^T \left( C C^T \right)^{-1} (D^T D)^{-1} D^T Y = (B^T B)^{-1} B^T Y;
\]

(3) When \( n < N+1 \), B is a row full rank matrix, \( D \) can be taken as a unit matrix, \( B = C \), so

\[
\hat{b} = C^T \left( C C^T \right)^{-1} (D^T D)^{-1} D^T Y = B^T (B^T B)^{-1} Y;
\]

Definition 6. Let Eq (7) be defined as the differential equation of improved DVCGM (1, N) model:

\[
\frac{dx^{(1)}_1(t)}{dt} + ax^{(1)}_1(t) = \sum_{i=1}^{M} \int_{0}^{T} b_i \lambda^{(i)}_i - s x^{(1)}_i(s) ds + \sum_{q=M+1}^{N} \int_{0}^{Q} b_q \lambda^{(q)}_q - s d^{(1)}_q(s) ds \tag{7}
\]

Where \( k = 2, 3, n, a \) is the development coefficient and \( b_1, b_2, \ldots, b_n \) are the grey input coefficients, \( \lambda_i \) is the hysteresis parameter. PSO algorithm is used to determine the coefficients of \( a, b_1, b_2, \ldots, b_n \), and \( \lambda_i \). Then the differential equation of the GM can be determined by Eq (7). The solution of Eq (7) is as follows:

\[
x^{(1)}_1(t) = e^{-at} \left\{ X^{(0)}_1 - t \left[ \sum_{i=2}^{M} b_i x^{(1)}_i(0) + \sum_{q=M+1}^{N} b_q d^{(1)}_q(0) \right] \right.
\]

\[
+ \sum_{i=2}^{N} \int \left[ \sum_{j=2}^{M} \sum_{i=1}^{k} b_i \lambda^{(i)}_i x^{(1)}_i(t) + \sum_{q=M+1}^{N} \sum_{j=1}^{k} b_q \lambda^{(q)}_q d^{(1)}_q(t) \right] e^{at} dt \right\} \tag{8}
\]

When the range of the driving factor sequence is small, the driver term can be viewed as a grey constant, and then the approximate TRF sequence of the grey differential equation of the model is

\[
x^{(1)}_1(k) = \frac{1}{a} \left[ \sum_{i=2}^{M} \sum_{j=1}^{k} b_j \lambda^{(i)}_i x^{(1)}_i(t) + \sum_{q=M+1}^{N} \sum_{j=1}^{k} b_q \lambda^{(q)}_q d^{(1)}_q(t) \right]
\]

\[
- \frac{e^{-a(k-1)}}{a} \left[ \sum_{i=2}^{M} \sum_{j=1}^{k} b_j \lambda^{(i)}_i x^{(1)}_i(t) + \sum_{q=M+1}^{N} \sum_{j=1}^{k} b_q \lambda^{(q)}_q d^{(1)}_q(t) \right] + X^{(1)}_1 e^{-a(k-1)} \tag{9}
\]
Finally, in order to predict the elements of the original sequence, the inverse accumulated generating operation should be performed. The predicted values can be determined as follows:

\[ \hat{x}_1^{(0)}(k) = \hat{x}_1^{(1)}(k) - \hat{x}_1^{(1)}(k - 1), k \geq 2, \]

where \( \hat{x}_1^{(0)}(n) \) is an estimation of the original sequence, which is simulation values, \( \hat{x}_1^{(0)}(n + 1), \hat{x}_1^{(0)}(n + 2), \cdots \) are predictive values.

2.3.2 Estimation and optimization of hysteresis parameter in the improved DVCGM (1, N) model. The most important part of the improved DVCGM (1, N) model is estimating the time-lag parameter, which directly affects the accuracy of the model. However, the time-lag parameters must be determined in advance, followed by B and Y matrix construction and system parameters calculation through OLS. Once the system parameters a and b are determined, we can obtain the TRF and the simulation and prediction value of the model.

In this paper, a nonlinear optimization model is established by using the Least One Multiplication. Then, the time-lag parameter is determined. When the range of the driving factor sequence is small, Eq (9) is used as the TRF, \( \lambda_i \) can be solved by the following nonlinear programming model:

\[
\min_{\lambda_i} \sum_{k=2}^{n} \left[ \sum_{i=2}^{M} \sum_{j=1}^{k} b_i x_i^{k-j}(t) + \sum_{q=M+1}^{N} \sum_{j=1}^{k} b_q x_q^{k-j}(t) \right] - \frac{e^{-a(k-1)}}{a} \left[ \sum_{i=2}^{M} \sum_{j=1}^{k} b_i x_i^{k-j}(t) + \sum_{q=M+1}^{N} \sum_{j=1}^{k} b_q x_q^{k-j}(t) \right] + X_1^{(1)} e^{-a(k-1)}
\]

\[
\begin{align*}
\hat{x}_1^{(0)}(k) &= \hat{x}_1^{(1)}(k) - \hat{x}_1^{(1)}(k - 1), k \geq 2 \\
\hat{x}_1^{(1)}(k) &= \frac{1}{a} \left[ \sum_{i=2}^{M} \sum_{j=1}^{k} b_i x_i^{k-j}(t) + \sum_{q=M+1}^{N} \sum_{j=1}^{k} b_q x_q^{k-j}(t) \right] \\
\text{s.t.} \\
& \quad \frac{e^{-a(k-1)}}{a} \left[ \sum_{i=2}^{M} \sum_{j=1}^{k} b_i x_i^{k-j}(t) + \sum_{q=M+1}^{N} \sum_{j=1}^{k} b_q x_q^{k-j}(t) \right] + X_1^{(1)}(1) e^{-a(k-1)} \\
& \quad \hat{b} = [a, b_2, \cdots, b_N]^T \\
& \quad 0 < \lambda_i < 1, i = 2, 3, \ldots, M, M + 1, \ldots, N
\end{align*}
\]

The model takes the relationship between structural parameters as the constraint condition and minimizes the average simulation relative error of the system characteristic variables, which can improve accuracy to the greatest extent.

The above optimization problem can be solved by PSO (Kiran and Mustafa, 2017; Mason et al., 2018; Chen et al., 2018). According to Eq (10), a nonlinear optimization method based on PSO algorithm is constructed to obtain the hysteresis parameter. PSO sets a certain number of particles in feasible region to find the best location and can be used to seek optimal values of \( \lambda_i \). Denote \( \lambda_i \) in Eq (10) and construct the fitness function of each particle, according to Eq (12).
Obviously, the average simulation relative error of the system characteristic variable sequence varies depending on the number of lag periods. The value of lag period should be selected to make the average simulation relative error as small as possible. Therefore, the improved DVCGM (1, N) model can well describe the hysteresis effect between the system characteristic variables. Once the hysteresis parameter is determined, the structural parameters of the model are set accordingly, and the simulation and prediction results can be obtained according to Eq (9).

2.3.3 Modeling procedure. Detailed procedure of the improved DVCGM (1, N) model is illustrated as follows.

Step 1. Collect raw data and establish original sequence \(X^{(0)}_1\), \(X^{(1)}_1\) is the 1-AGO sequence from \(X^{(0)}_1\), \(X^{(1)}_1\) is the sequence of the relevant factors. \(Z^{(1)}_1\) is the sequence of mean generation of consecutive neighbors from \(X^{(1)}_1\). Then, determining virtual variable sequence

\[
D^{(0)}_j = \left( d^{(0)}_j(1), d^{(0)}_j(2), \cdots, d^{(0)}_j(n) \right), d^{(0)}_j(n) = 0 \text{ or } 1.
\]

\(D^{(1)}_j\) is also the sequence of the relevant factors.

Step 2. Solving delay parameters \(\lambda_i\) by PSO according (10), then constructing vector \(Y\) and matrix \(B\). Using the least square method, the values of the coefficients \(a\) and \(b_1, b_2, \cdots, b_n\) can be determined.

Step 3. After considering the lag effect, the differential equation of the grey model can be determined by Eq (7).

Step 4. The time response function of the new model is established to generate prediction data according to Eq (9).

3. Application

Forecasting EI can be considered as a grey problem, because EI is greatly affected by technological progress, population factor, industrial structure and so on. These factors influence EI through a dynamic and complicated mechanism. The uncertain impacts and limited number of data provides a good basis for grey theory application. There are four procedures in this part, including data collection, parameter estimation, result comparisons and future forecasts. Three competing models, namely GM (1, N), DVCGM (1, N) and ARIMA model, are employed to test the accuracy of the improved DVCGM (1, N) on EI forecasting.

The three GMs used in this paper are interrelated and are from basic to the advanced. The GM (1, N) model is a traditional grey multivariate prediction model. On its basis, DVCGM (1, N) model introduces dummy variables, taking into account the influence of policies and other factors. The improved DVCGM (1, N) model introduces time-lag parameters, which
further enriches the existing grey prediction theory by considering the hysteresis effect of policies. The above GMs are used to estimate China’s EI, and the optimal model is selected by comparing their performance, to predict China’s EI in the next five years. In addition, we use ARIMA model as a comparison of grey methods to show that the improved DVCGM (1, N) model is not only better than the traditional GM, but also better than the non-grey econometric model.

3.1 Variables selection and data collection
The indicators selected in this paper are as follows. EI is the ratio of total energy consumption to GDP. Population factor is the employed population. The ratio of the added value of industrial production to the added value of energy consumed by the industry is used as the substitution variable of technological progress (Yan, 2011). Industrial structure is measured by the ratio of output value of each industry to GDP. For the consistency of the statistical scope, we choose a time scale of 2001–2017, and all data are collected from China Statistical Yearbook. With small sample size (17 periods’ real measurement values) and insufficient information, this case fits well with the grey system.

First, we calculate the grey correlation between EI and population, technological progress and industrial structure. As shown in Table 1, technological progress has the highest correlation with EI. Therefore, we select technological progress as the driving variable, and EI is the system behavior variable. Then, we can build GM (1, N) model for EI forecast.

As shown in Figure 3, China’s EI increased rapidly from 2001, rising sharply to a peak of 1.4542 tons standard coal per 10,000 Yuan in 2005 and declined steadily afterward. In 2017, the EI has decreased by 62.67%, leaving only 0.5428 tons of standard coal per 10,000 Yuan. The 11th Five-Year Plan in 2006 is a kind of watershed for Chinese EI. The government introduced a series of strict policies of energy conservation and emission reduction, and

| Relative correlation     | Energy intensity |
|--------------------------|------------------|
| Population factors       | 0.6126           |
| Technological progress   | 0.9550           |
| Industrial structure     | 0.7652           |

Table 1. Relative correlation analysis

Figure 3. Trends of energy intensity (dash line) and technological progress (solid line) of China
China’s EI declined obviously. As a result, we need to put policy and the hysteresis effect of policy into consideration when estimate EI of China.

3.2 Simulation of energy intensity in China

Four models are applied in the simulation of EI of China. First, we use the $GM(1, N)$ model mentioned in Section 2.1 to predict EI. We select technological progress as the driving variable, EI as the system behavior variable and then establish the $GM(1, N)$ model. According to 2.1, the TRF is obtained as follows:

$$X^{(1)}_1(k) = 4.2429X^{(1)}_2(k) \times (1 - e^{-1.8419(k-1)}) + e^{-1.8419(k-1)}$$

The simulation results of $GM(1, N)$ model are illustrated in Table 3, and the relative errors and average relative errors can be calculated, as shown in Table 3.

Secondly, we build $DVCGM(1, N)$ model for EI estimate. We select technological progress as the driving variable, EI as the system behavior variable. Energy conservation policy ($P$) is introduced as a dummy variable. Before 2006, as the strict energy-saving policies had not been implemented, $P$ value is 0; after 2006, $P$ value is 1. According to Section 2.2, the TRF is obtained as follows:

$$X^{(1)}_1(k) = (4.2830X^{(1)}_2(k) - 0.0232d^{(1)}_3(k)) \times (1 - e^{-1.7372(k-1)}) + e^{-1.7372(k-1)}$$

The simulation results of $DVCGM(1, N)$ model are illustrated in Table 3, and the relative errors and average relative errors can be calculated, as shown in Table 3.

Thirdly, we use the improved $DVCGM(1, N)$ model to estimate EI. As discussed in Section 2.3, $DVCGM(1, N)$ model takes consideration of the hysteresis effect of dummy variables. We select technological progress as the driving variable, EI as the system behavior variable and energy-saving policy as virtual variable. Different from $DVCGM(1, N)$ model, we need to determine the delay parameters, along with the structure parameters and build the TRF to get the simulation results.

1) Determination of time-delay parameter.

As stated by Eq (10-12), a nonlinear optimization method based on PSO algorithm is constructed to obtain the hysteresis parameter. According to the optimization model shown in Eq (9), the average relative percentage errors (APE) of the model under different lag periods is calculated in Table 2. When the hysteresis parameter is 2, the average relative percentage errors (APE) of the model is the smallest (0.9754%). Therefore, the value of time lag is two years. In the view of the complexity of socio-economic ecology, adaptive adjustments are made according to the changes of policies. These energy-saving policies influence economic operation through a certain transmission mechanism and gradually affect the EI.

2) Determination of structural parameters.

The hysteresis parameter is substituted into the B matrix and Y matrix. According to Theorem 2, the matrix operation of least square regression is used to obtain the structural parameter values $b_1, b_2, B^T = (1.251669, 0.938994)$.

| Lag period | 0   | 1   | 2   | 3   | 4   | 5   |
|------------|-----|-----|-----|-----|-----|-----|
| APE(%)     | 1.1107 | 1.0023 | 0.9754 | 1.0667 | 1.0989 | 0.9836 |

Table 2. Average percentage errors (APEs) of the model under different lag periods.
(3) Calculate the simulated and predicted values. Putting the values of the estimated structural parameter and hysteresis parameter in Eq (7-9), we can obtain the optimal TRF as follows:

\[ X_{1}^{(1)}(k) = \frac{4.2830X_{2}^{(1)}(k) - 0.0232d_{3}^{(1)}(k)}{C_{0}} \times (1 - e^{-1.7372(k-1)}) d + e^{-1.7372(k-1)} \]

Then, the simulation results of improved DVCGM (1, N) model are illustrated in Table 3, and the relative percentage errors (PE) and average relative percentage errors (APE) can be calculated, as shown in Table 3.

Finally, we employ ARIMA (autoregressive composite moving average) model to test our results from the non-grey perspective. The estimation results of ARIMA model are shown in Table A1-4 in the Appendix. All the coefficients are statistically significant and the model is well fitted (R-squared is 0.850452). The ARIMA (2,1,1) model of time series is determined as follows:

\[ \Delta EI = -0.031 + 1.52\Delta EI_{t-1} - 0.94\Delta EI_{t-2} + \epsilon_{t-0.99} \epsilon_{t-1} \]

The performance of the ARIMA model and the calculated relative errors and average relative errors are presented in Table 3.

As illustrated in Table 3, there is great consistency between simulated values and real values for the improved DVCGM (1, N) model. For APE, which is the performance prediction index, its values of the improved DVCGM (1, N) is the smallest (1.76% in the in-sample periods and 1.92% in the out-sample periods) among all four models.
3.3 Comparison and discussion of the results

As shown in Figure 4a, the overall trends of the GM (1, N), DVCGM (1, N) and improved DVCGM (1, N) models and ARIMA model comply with the true curve to some extent. However, the performance evaluation is distinct (Figure 4b). The APE value of the improved DVCGM (1, N) model is the smallest with the minimum fluctuation, demonstrating the efficacy and reliability of the model. It also suggests that the improved DVCGM (1, N) model is better than the traditional grey multivariable models. The ARIMA model has the largest values of APE among these four competing models. GM (1, N) model obtains the second largest values of APE, suggesting DVCGM (1, N) is the suboptimal choice.

In view of the hysteresis effect of energy-saving policies, the improved DVCGM (1, N) model has a much lower error than the other three models. Therefore, we choose the improved DVCGM (1, N) model as the best model for predicting EI.

3.4 Forecasting the future energy intensity from 2018 to 2022

Based on the improved DVCGM (1, N) model, we can forecast the output value of China’s EI from 2018 to 2022 using data of technological progress. However, the data of technological progress from 2018 to 2022 is unknown and needs to be predicted in advance. The one-step rolling GM (1, 1) model is considered for the data prediction.

We used the GM (1, 1) model for in-sample simulation and out-of-sample prediction of technological progress. The results (Table 4) show that the in-sample simulation error is 2.97%, and the out-of-sample hind cast error is 2.11%, which indicates that the GM (1, 1) model gives satisfying results and can be used to predict technological progress from 2018 to 2022.

It is also noted that energy conservation policy (P) is introduced as a dummy variable. Before 2020, P value is 0; after 2020, P value is 1. It is because that the 13th Five-Year Plan is from 2016 to 2020, new performance target and new policies for the next stage are
formulating and will be effective after 2020. Therefore, the data we need to substitute into the improved $DVCGM$ $(1, N)$ model is shown in Table 5.

The predicted results of China’s EI are shown in Table 6. In addition, we also draw a line graph to make the results more iconic in Figure 5.

As illustrated in Figure 5, there is a downward trend of the EI in the next five years. By 2020, the EI is expected to decrease by 20% or more than it was in 2016. That is to say, during the 13th Five-Year Plan period (2016–2020), EI will drop by more than 15%, meeting the country's energy performance target. Therefore, government policies have a profound influence on EI. When formulating energy conservation and emission reduction policies, we should consider the hysteresis effect of the policies and make adjustments accordingly to achieve the goal.
4. Conclusions

Over the past 20 years, China is gradually shifting from a resource-intensive and energy-driven economy to a more sustained economy. EI in China has fallen almost continuously while China focuses on the industrial upgrading and promoting transformation of the economic structure. Energy-saving policies and regulations are introduced to help China reach its energy performance target. However, few studies have been carried out to consider the hysteresis effects of policies on the estimation of EI. Therefore, to address such a challenge problem, an improved grey multivariable model is designed to forecast China’s EI considering the hysteresis effect of government policies. To further improve its forecasting capability, a nonlinear optimization method based on PSO algorithm is constructed to calculate the hysteresis parameter. In addition, three conventional models, namely GM(1, N), DVCGM(1, N) and ARIMA models, are applied to test the accuracy of this improved DVCGM (1, N) model. The empirical results demonstrate that the proposed model considering the hysteresis effects of energy conservation policies performs best and matches well with the actual observations. Accordingly, this proposed model is used to forecast EI value from 2018 to 2022. The main conclusions are as follows:

(1) The improved DVCGM (1, N) model can solve the modeling problem of small sample systems with time-delay causality. A nonlinear optimization method based on PSO algorithm is constructed to calculate the hysteresis parameter. It overcomes the defects of traditional GMs and econometric models.

(2) GM (1, N), DVCGM (1, N) and ARIMA model are taken as comparative models. The accuracy of improved DVCGM (1, N) model was tested by the average relative percentage errors. The results show that the Improved DVCGM (1, N) model notes the hysteresis effect of government policies and significantly improves the prediction accuracy of China’s EI than the other three models. As suggested by APEs, the overall fitting in descending order is improved DVCGM (1, N) model, DVCGM (1, N), GM (1, N) and ARIMA model.

(3) China’s EI is greatly influenced by technological progress and is much of policy-driven. When formulating energy conservation and emission reduction policies, we should fully consider the hysteresis effect of the policies, so as to make adjustment of the relative policies and better achieve the national energy performance target.

A few caveats are appropriate. It is an interesting further path to work out the hysteresis parameter directly from the nonlinear programming model. Besides, the sustainability of the
hysteresis effect of policy is worth considering. Furthermore, population factors and industrial structure also have good correlation with EI. These will be investigated in our further studies.

References

Chen, K., Zhou, F. and Liu, A. (2018), “Chaotic dynamic weight particle swarm optimization for numerical function optimization”, Knowledge-Based Systems, Vol. 8 No. 10, pp. 23-40.

Dang, Y.G., Wei, L. and Ding, S. (2017), “Grey multivariable discrete delay model based on driver information control and its application”, (in Chinese with English abstract), Control and Decision making, Vol. 32 No. 9, pp. 1672-1680.

Deng, J.L. (1982), “Control problems of grey system”, Systems and Control Letters, Vol. 1 No. 5, pp. 288-294.

Ding, S., Dang, Y.G., Li, X.M., Wang, J.J. and Zhao, K. (2017), “Forecasting Chinese CO2 emissions from fuel combustion using a novel grey multivariable model”, Journal of Cleaner Production, Vol. 162 No. 3, pp. 1527-1538.

Ding, S., Dang, Y.G. and Xu, N. (2018), “Construction and application of GM (1, N) model based on virtual variable control”, (in Chinese with English abstract), Control and Decision making, Vol. 33 No. 2, pp. 309-315.

Dong, K., Sun, R., Hochman, G. and Li, H. (2018), “Energy intensity and energy conservation potential in China: a regional comparison perspective”, Energy, Vol. 155 No. 2, pp. 782-795.

Guo, H., Xiao, X. and Forrest, J. (2013), “A research on a comprehensive adaptive grey prediction model CAGM (1, N)”, Applied Mathematics and Computation, Vol. 225 No. 3, pp. 216-227.

Hao, Y., Wang, Y., Zhao, J. and Li, H. (2011), “Grey system model with time lag and application to simulation of karst spring discharge”, Grey Systems: Theory and Application, Vol. 1 No. 1, pp. 47-56.

Karimu, A., Brännlund, R. and Lundgren, T. (2017), “Energy intensity and convergence in Swedish industry: a combined econometric and decomposition analysis”, Energy Economics, Vol. 62 No. 5, pp. 347-356.

Kiran and Mustafa, S.. (2017), “Particle swarm optimization with a new update mechanism”, Applied Soft Computing, Vol. 60 No. 5, pp. 670-678.

Kose, E. and Tasci, L. (2015), “Prediction of the vertical displacement on the crest of Keban Dam”, Journal of Grey System, Vol. 27 No. 1, pp. 12-20.

Kose, E. and Tasci, L. (2019), “Geodetic deformation forecasting based on multi-variable grey prediction model and regression model”, Grey Systems: Theory and Application, Vol. 9 No. 4, pp. 464-471.

Ma, B. and Yu, Y. (2017), “Industrial structure, energy-saving regulations and energy intensity: evidence from Chinese cities”, Journal of Cleaner Production, Vol. 141 No. 3, pp. 1539-1547.

Mason, K., Duggan, J. and Howley, E. (2018), “A meta optimisation analysis of particle swarm optimisation velocity update equations for watershed management learning”, Applied Soft Computing, Vol. 62 No. 10, pp. 148-161.

Pao, H.T., Fu, H.C. and Tseng, C.L. (2012), “Forecasting of CO2 emissions, energy consumption and economic growth in China using an improved grey model”, Energy, Vol. 40 No. 1, pp. 400-409.

Tan, R. and Lin, B. (2018), “What factors lead to the decline of energy intensity in China’s energy intensive industries?”, Energy Economics, Vol. 71 No. 10, pp. 213-221.

Tien, T.L. (2005), “The indirect measurement of tensile strength of material by the grey prediction model GMC(1,N)”, Measurement Science and Technology, Vol. 16 No. 6, pp. 1322-1328.

Tien, T.L. (2010), “Forecasting CO2 output from gas furnace by grey prediction model IGMC(1,N)”, Journal of the Chinese Society of Mechanical Engineers, Vol. 31 No. 1, pp. 55-65.
Tien, T.L. (2011), “The indirect measurement of tensile strength by the new model FGMC(1,N)”, Measurement, Vol. 44 No. 10, pp. 1884-1897.

Wang, Z.X. and Hao, P. (2016), “An improved grey multivariable model for predicting industrial energy consumption in China”, Applied Mathematical Modeling, Vol. 40 No. 11, pp. 5745-5758.

Wu, L.F., Gao, X., Xiao, Y., Yang, Y. and Chen, X. (2018), “Using a novel multi-variable grey model to forecast the electricity consumption of Shandong Province in China”, Energy, Vol. 157 No. 10, pp. 327-335.

Xiong, P.P., Zhang, Y., Xing, Z. and Chen, F. (2019), “Multivariable time-delay discrete MGM (1, m, tau) model and its application”, (in Chinese with English abstract), Statistics and Decision, Vol. 35 No. 8, pp. 18-22.

Yan, Q.Y. and Liu, F. (2011), “Analysis and prediction of industrial energy consumption intensity in China from 2010 to 2030”, (in Chinese with English abstract), East China electric power, Vol. 39 No. 12, pp. 1858-1862.

Zhai, X. (1996), “GM (1, 2) model with time delay and its application”, Systems Engineering, Vol. 14 No. 6, pp. 56-62.

Zhang, K. (2015), “Time-delayed multivariable discrete grey model and its application”, Systems Engineering Theory and Practice, Vol. 35 No. 8, pp. 2092-2103.

Zhang, K. (2016), “Multivariate discrete grey model base on dummy drivers”, Grey Systems: Theory and Application, Vol. 6 No. 2, pp. 246-258.

Zhu, J.M. and Ruth, M. (2015), “Relocation or reallocation: impacts of differentiated energy saving regulation on manufacturing industries in China”, Ecological Economics, Vol. 110 No. 1, pp. 119-133.
## Appendix

### Null hypothesis: EI has a unit root
Exogenous: constant
Lag length: 1 (automatic - based on SIC, maxlag = 1)

| AuGmented Dickey–Fuller test statistic | \( t \)-statistic | Prob.* |
|----------------------------------------|--------------------|--------|
| Test critical values                   |                    |        |
| 1% level                               | \(-4.420595\)      | 0.0102 |
| 5% level                               | \(-3.259808\)      |        |
| 10% level                              | \(-2.771129\)      |        |

Table A1. Unit root test results of EI of ARIMA model

### Null hypothesis: DEI has a unit root
Exogenous: constant, linear trend
Lag length: 2 (automatic - based on SIC, maxlag = 2)

| AuGmented Dickey–Fuller test statistic | \( t \)-statistic | Prob.* |
|----------------------------------------|--------------------|--------|
| Test critical values                   |                    |        |
| 1% level                               | \(-5.277473\)      | 0.0102 |
| 5% level                               | \(-4.008157\)      |        |
| 10% level                              | \(-3.460791\)      |        |

Table A2. Unit root test results of DEI of ARIMA model

### Autocorrelation and partial correlation of ARIMA model

| Autocorrelation | Partial correlation | AC   | PAC   | Q-stat | Prob  |
|-----------------|---------------------|------|-------|--------|-------|
| .**** | .**** | 1 | 0.714 | 0.714 | 8.2763 | 0.004 |
| .* | .**** | 2 | 0.195 | -0.641 | 8.9492 | 0.011 |
| .** | . | 3 | -0.219 | 0.035 | 9.8859 | 0.020 |
| .*** | . | 4 | -0.366 | 0.010 | 12.796 | 0.012 |
| .** | .* | 5 | -0.297 | -0.066 | 14.940 | 0.011 |
| .* | . | 6 | -0.124 | 0.030 | 15.367 | 0.018 |
| . | .*** | 7 | -0.039 | -0.267 | 15.417 | 0.031 |
| . | .* | 8 | -0.024 | 0.139 | 15.439 | 0.051 |
| .* | .*** | 9 | -0.087 | -0.284 | 15.805 | 0.071 |
| .* | .* | 10 | -0.136 | 0.079 | 17.004 | 0.074 |
| .* | . | 11 | -0.100 | 0.033 | 17.982 | 0.082 |
| . | * | 12 | -0.017 | -0.188 | 18.037 | 0.115 |

Table A3. The autocorrelation and partial correlation of ARIMA model

Hysteresis effect of China’s EI forecasting
| Variable | Coefficient | Std. Error | t-statistic | Prob     |
|----------|-------------|------------|-------------|----------|
| C        | -0.031125   | 0.007696   | -4.044042   | 0.0037   |
| AR(1)    | 1.519400    | 0.137896   | 11.01843    | 0.0000   |
| AR(2)    | -0.936168   | 0.059890   | -15.63139   | 0.0000   |
| MA(1)    | -1.000000   | 39728.31   | -2.52E-05   | 1.0000   |
| SIGMASQ  | 0.000442    | 0.403940   | 0.001093    | 0.9992   |

Table A4. Estimation of ARIMA model

|                      |                |                |             |          |
|----------------------|----------------|----------------|-------------|----------|
| R-squared            | 0.850452       | Mean dependent var | -0.018404  |
| Adjusted R-squared   | 0.775678       | S.D. dependent var | 0.056562   |
| S.E. of regression   | 0.026789       | Akaike info criterion | -3.664240  |
| Sum squared resid    | 0.006741       | Schwarz criterion | -3.446952  |
| Log likelihood       | 28.81756       | Hannan-Quinn criterion | -3.708902  |
| F-statistic          | 11.37363       | Durbin–Watson stat | 1.064574   |
| Prob(F-statistic)    | 0.002202       |                |             |          |

Corresponding author
Xinchang Guo can be contacted at: guo201064@163.com