



# A novel grey prediction model with a feedforward neural network based on a carbon emission dynamic evolution system and its application

Weige Nie<sup>1,2</sup> · Ou Ao<sup>1,2</sup> · Huiming Duan<sup>1</sup>

Received: 25 July 2022 / Accepted: 6 October 2022 / Published online: 18 October 2022  
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

## Abstract

The objective and accurate prediction of carbon dioxide emissions holds great significance for improving governmental energy policies and plans. Therefore, starting from an evolutionary system of carbon emissions, this paper studies the evolution of the system, establishes a grey model of the system, and expands the modeling structure of this model. The modeling mechanism of the classical feedforward neural network model is organically combined with the function of the external influencing factors of carbon emissions, and the grey model of the carbon emission dynamic system is established with a neural network. Then, the properties of the model are studied, the parameters of the model are optimized, and the modeling steps are obtained. Finally, the validity of the model is analyzed by using the carbon emissions of Beijing from 2009 to 2018. The results of the four cases show that the simulation and prediction errors of the new model are all less than 10%, and case 1 shows the best results of 1.56% and 2.07%, respectively, which are used to predict the carbon dioxide emissions in the next 5 years in Beijing. The prediction results are in accordance with the actual trend, which indicates the effectiveness and feasibility of the model.

**Keywords** Grey prediction model · Carbon emissions · Neural network · Dynamic system

## Nomenclature

APE	The absolute percentage error
$f(t)$	The function of regulating external factors
$g(\cdot)$	The activation function
GMVM(1,1)	The improved grey multivariable Verhulst model
GM(1, N)	The first-order grey model with $N$ variables
GMC(1, N)	The convolution integral grey prediction model
$M(t)$	The carbon emission vector

MAPE	The mean absolute percentage error
NN-CDGM(1,N)	The neural network-carbon emission dynamic system grey model

## Introduction

With technological progress and population growth, a rapid rise in global carbon emissions is inevitable. Global carbon dioxide emissions increased by 40% between 2000 and 2019. According to a report (BP 2021) released by the British Petroleum Company, global carbon emissions have maintained continuous growth since 2013, reaching 34.36 billion tons in 2019, a record high. The increase in carbon dioxide in the atmosphere has led to a rise in surface temperatures, which directly leads to the frequent occurrence of extreme weather, such as drought, heat, and glacial melting. These changes have become a great challenge to human development. Reducing non-renewable energy consumption and controlling carbon dioxide emissions have become common goals of all countries worldwide.

Responsible Editor: Marcus Schulz

✉ Huiming Duan  
duanhm@cqupt.edu.cn

<sup>1</sup> School of Science, Chongqing University of Posts and Telecommunications, Chongqing 400065, China

<sup>2</sup> Key Laboratory of Intelligent Analysis Decision Complex System, Chongqing University of Posts and Telecommunications, Chongqing 400065, China

In 2020, with the outbreak of COVID-19, all countries in the world suffered an economic blow to varying degrees, and energy consumption and carbon dioxide emissions in all countries changed dramatically compared to levels before the COVID-19 pandemic, attracting the attention of governments and scholars around the world (Wang et al. 2022). China’s carbon emissions reached 9.899 billion tons, with an annual increase of 0.6% (China’s 2013–2020 carbon emissions are shown in Fig. 1). According to the statistics of the International Energy Agency (IEA, Greenhouse Gas Emissions from Energy 2021), in 2020, global carbon emissions mainly came from energy generation and heating, transportation, and manufacturing and construction, accounting for 43%, 26%, and 17%, respectively (Fig. 2(a)). In 2020, the Asia–Pacific region accounted for more than half of global carbon emissions, with a combined 52%, and China accounted for 30.7%, far more than other regions (Fig. 2(b)).

China is a developing country; therefore, its economic development still requires considerable coal energy consumption, and its carbon emissions are still growing. As the world’s largest developing country and the largest energy consumer and carbon dioxide emitter, China’s carbon emissions are closely monitored because they have a direct impact

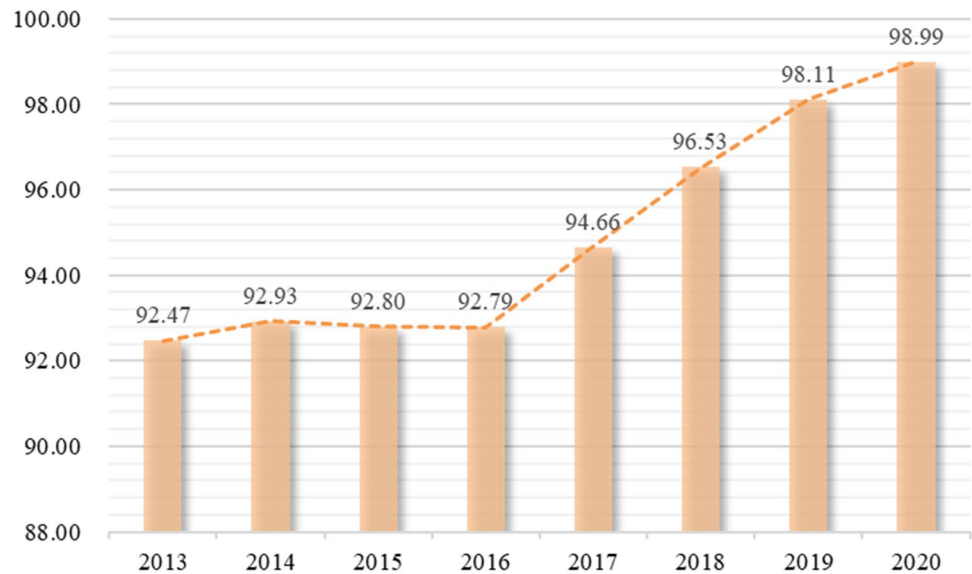
on global trends. Therefore, effectively predicting China’s carbon emissions and grasping the trend of carbon emissions in advance to adjust policies and effectively implement corresponding measures is of great significance. This information can provide a reliable basis for China’s energy conservation and emission reduction measures and serve as an important indicator for global carbon dioxide control.

At present, many scholars have studied the prediction methods of carbon dioxide emissions, including machine learning methods, statistical models, dynamic scenario simulation models, and grey prediction models.

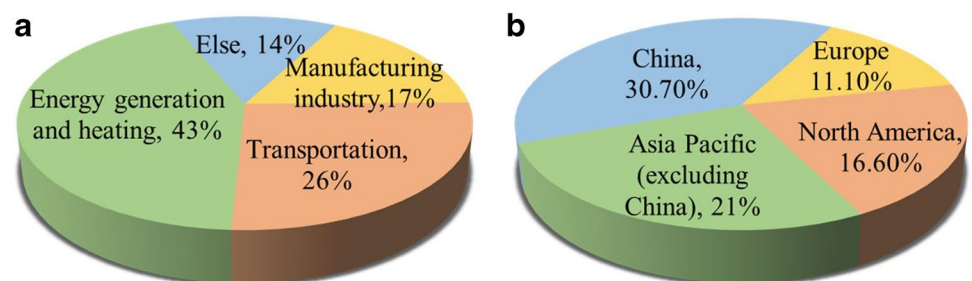
### Machine learning methods

Based on classical neural network models, such as generalized regression neural network prediction models (Heydari et al. 2019) and least squares support vector machine models (Sun and Liu 2016; Qiao et al. 2020), scholars have introduced optimization algorithms to improve model performance. In terms of the model structure, many scholars have established neural network models that are more suitable for predicting carbon emissions, such as the machine learning method optimized by Adam and proposed by Ashin Nishan

**Fig. 1** China’s carbon emissions from 2013 to 2020 (BP, Statistical Review of World Energy, 2019)



**Fig. 2** (a) Composition of global carbon emission sources in 2020; Fig. 2(b) Composition of global carbon emission regions in 2020 (IEA, Greenhouse Gas Emissions from Energy, 2021)



and Muhammed Ashiq (2020) and the meta-elastic network established by Nguyen et al. (2021).

## Statistical models

Scholars have used the Ridge regression decomposition method (Ma et al. 2019), cumulative Monte Carlo distribution curve (Robati et al. 2019), regression analysis and other methods (Xia et al. 2019; Modise et al. 2021; Boamah et al. 2021) to predict the carbon emissions of different countries or industries and obtained many excellent results.

## Dynamic scenario simulation model

Scholars have used dynamic scenario simulation models to analyze the future trend of carbon emissions, for example, by combining such a model with the Monte Carlo simulation method to construct a new dynamic scenario simulation model (Hu and Lv 2020), discussing the dynamic evolution trajectory of China's building carbon emissions, and then combining this model with specific data to provide effective information for the government (Huo et al. 2021a, b).

The above models have good effects in predicting carbon emissions, but they all have a common problem: the need for large quantities of high-quality data. Machine learning methods usually require large sample sizes to train models before making predictions, and there may be problems of excessive sample fitting or improper parameter setting. Statistical models and scenario simulation models rely heavily on sufficient and effective input data for accurate calculation, which also requires a considerable amount of data to obtain high-quality effective data. However, current carbon dioxide emissions mainly come from fossil fuels. The COVID-19 pandemic is likely to exacerbate changes in the global energy mix, and historical data of carbon dioxide emissions are inconsistent with the current situation. Due to the substantial reduction in reliable data, models requiring a substantial amount of data modeling are no longer applicable. Therefore, models that requires less data to obtain relatively effective results are attractive (Deng 2002).

The grey prediction model was proposed by Deng (2002). This kind of model has a simple structure and strong adaptability. It can better handle parameters and is suitable for systems with small samples and unclear information. Compared with other methods, the advantage of the grey prediction model is that it can produce accurate results for an uncertain system with scarce data. The grey prediction model weakens the randomness of the original data by generating the cumulative sequence of the original data to mine the hidden rules in the data sequence, and finally simulates and predicts the data through the restoration operation (Duan et al. 2020a; Duan and Pang 2021; Wang and Li 2020; Zeng et al. 2020; Yan et al. 2020). It is widely used in the prediction of energy

(Duan and Luo 2021; Duan and Luo 2020; Zhao and Wu 2020; Gao et al. 2022a), transportation (Duan et al. 2020b, Bezuglov and Comert 2016), the environment (Mao et al. 2020; Liu et al. 2020), and other industries (Gao et al. 2022b). The grey prediction model also plays an important role in predicting carbon emissions. The following is a description of the grey univariate model, the multivariate model and the combination of the grey prediction model and other models.

## Grey univariate prediction model

This kind of model ignores the influence of other related factors and only considers the change trend and characteristics of the only variable. Scholars have directly applied the existing grey univariate and discrete univariate prediction models to predict carbon emissions (Ikram et al. 2021; Qiao et al. 2021) and introduced rolling mechanisms and other methods based on the classical univariate model to optimize the prediction of carbon emissions (Zhou et al. 2021). The grey multivariate prediction model is a relatively grey univariate prediction model that considers more related factors and therefore has good adaptability. Scholars have generalized and optimized the grey multivariate prediction model based on the model structure (Cao et al. 2021), theoretical derivation (Duan et al. 2020a), and sequence optimization (Ye et al. 2021) and effectively applied it to predict carbon emissions.

Grey univariate and multivariate prediction models have certain advantages, while machine learning models have an excellent performance in parameter calculation and optimization. Therefore, scholars have combined the two models to establish a kind of combined model suitable for small sample data with optimized parameters and performance. Hu and Lv (2020) combined the grey model with the optimized vector machine algorithm to simulate the carbon emission transfer network among various industries. Huang et al. (2019) used grey correlation analysis to identify the strongly correlated factors of carbon emissions and then used the long short-term memory (LSTM) method to predict China's carbon emissions. Zhao et al. (2021) proposed a hybrid prediction model combining an improved grey model and a machine learning model to analyze the long-term trend and short-term fluctuation of carbon emissions. Combining the neural network model on the basis of the grey prediction model can indeed optimize the model performance, and this idea is practical and effective.

The aforementioned grey prediction models have yielded good prediction results, but some limitations exist. First, most carbon emission prediction applications do not consider the background of the carbon emission system; rather, they establish a model and apply it directly, ignoring the importance of background characteristics. Second, carbon dioxide emissions are affected by many factors, such as population, GDP,

and the use of energy. Grey univariate forecasting models have certain limitations in such complex systems, and the aforementioned multivariable or combined models basically choose population and GDP as influencing factors. However, energy and other factors also need to be considered to better describe the carbon emission system. Finally, the combination model of the grey prediction model and machine learning generally combines the two models based on the modeling data, and determining how to combine the two models based on the model structure requires further research.

Therefore, starting from the evolutionary dynamic system of carbon emissions, this paper studies the law of carbon emissions and establishes a grey prediction model based on the principle of grey difference information. At the same time, based on the structure of the model, the grey of model the carbon emission dynamic system is established with a neural network by combining the function of the relevant factors of the model with the neural network. The model is applied to predict carbon emissions in Beijing, China, and good results are obtained.

This paper makes the following three contributions:

- (1) Based on the evolutionary system of carbon emissions, the concepts of the dynamic evolutionary factor, evolutionary coefficient, and critical time point of carbon emissions are introduced, and the nonlinear grey prediction model of the dynamic system of carbon emissions is established by using the principle of grey difference information.
- (2) To further analyze the external factors of the dynamic system of carbon emissions, the modeling mechanism of the neural network is used to adjust the external factors, and the grey model of the carbon emission dynamic system is established with a neural network based on the structure of the model.
- (3) The new model is applied to predict carbon emissions, the effectiveness of the model is analyzed based on the number of modeling objects, and the carbon emissions in the next 5 years are predicted by using the most effective case. Then, relevant policy suggestions are proposed based on the trend of the results.

The remainder of this paper proceeds as follows: In the second section, the grey prediction model is first established by the dynamic method, and then, the grey model of the carbon emission dynamic system is established with a neural network by combining the grey prediction model with the neural network modeling mechanism. In the third section, data on Beijing's carbon emissions are used to conduct an effective analysis, the carbon emissions in the next 5 years are predicted, and relevant suggestions are proposed. The fourth section presents the conclusions.

## A carbon emission dynamic system grey model based on a neural network

System evolution is a dynamic process of occurrence, development, mutation, and complex change over time. Research on the evolutionary law of a system can help to better understand, control, and analyze the system and promote the development process of the system. Carbon emissions are related to macroeconomic factors such as the national economy and energy. To effectively predict carbon emissions, the grey prediction model of the carbon emission evolutionary dynamic system is established based on the carbon emission evolutionary system, and the grey model of the carbon emission dynamic system is established with a neural network by combining the modeling mechanism of the neural network with the grey prediction model.

## An evolutionary dynamic system model of carbon emissions

The carbon emission system is a dynamic system, and the evolutionary law of carbon emissions can be studied by dynamic methods. By introducing concepts such as the dynamic evolutionary factor of carbon emissions, evolutionary coefficient and critical time point and by using existing data, the dynamic system of carbon emissions can be established (Tian and Jin 2012). Based on the nonlinear influence, the nonlinear dynamic system model can be established while considering the influence of external factors, which enables policymakers to regulate and adjust external influences. external factors, which is convenient for policy to regulate and adjust external influences.

The evolutionary dynamic system model of carbon emissions is as follows:

$$\frac{dM(t)}{dt} = g(M(t)) + f(t), t \in I \quad (1)$$

where  $M(t)$  represents the carbon emission vector at time  $t$ ;  $\frac{dM(t)}{dt}$  represents the dynamic evolutionary factor of carbon emissions, that is, the rate of change in carbon emissions over time;  $g(M(t))$  is the composition function of  $M(t)$ ; and  $f(t)$  is the function of regulating external factors.

Considering that the change rate  $\frac{dM(t)}{dt}$  of the carbon emission factor at time  $t$  is proportional to the carbon emission  $M(t)$  in the economic period,  $\frac{dM(t)}{dt}$  is proportional to the share of the growth potential of carbon emissions  $1 - \frac{M(t)}{ck_M}$ . The growth potential of carbon emissions is affected by technological progress, emission reduction technologies, and other factors. At the same time, it is also related to external factors that affect carbon emissions, including energy consumption, the external factors of the system, and the regulatory function  $f(t)$  of regulatory behavior. Thus, the following dynamic

system model of the evolution of carbon emissions can be established:

$$\frac{dM(t)}{dt} = rM(t)\left(1 - \frac{M(t)}{ck_M}\right) + cf(t) \tag{2}$$

where  $r$  is the carbon emission coefficient of economic period  $I$ ;  $k_M$  is related to the time interval;  $c$  is the carbon emission evolutionary coefficient;  $M(t)$  is carbon emissions;  $t$  is time; and  $f(t)$  is the policy regulation affecting carbon emissions, which can be determined by external factors such as energy consumption and economic status.

From Eq. (2), the following equation can be obtained:

$$\frac{dM(t)}{dt} = rM(t) - \frac{rM^2(t)}{ck_M} + cf(t) \tag{3}$$

This system is a direct carbon emission system that enables the overall trend of carbon emissions to be analyzed, and  $f(t)$  is the external regulatory function. Therefore, we let  $r = a, -\frac{r}{ck_M} = b$ , and the control function is  $cf(t) = f(M_1(t); \omega)$ . Then, Eq. (3) becomes the following:

$$\frac{dM(t)}{dt} = aM(t) + bM^2(t) + f(M_1(t); \omega) \tag{4}$$

### Carbon emission dynamic system grey model

Set  $M^{(0)}$  as the original data sequence of carbon emissions:

$$M^{(0)}(k) = (m^{(0)}(1), m^{(0)}(2), \dots, m^{(0)}(n)) \tag{5}$$

$M^{(1)}$  is the 1-AGO sequence of  $M^{(0)}$ :

$$M^{(1)}(k) = (m^{(1)}(1), m^{(1)}(2), \dots, m^{(1)}(n)), \tag{6}$$

where  $m^{(1)}(k) = \sum_{i=1}^k m^{(0)}(i), k = 1, 2, \dots, n. Z^{(1)}$  is the adjacent mean-generating sequence of  $M^{(1)}$ :

$$Z^{(1)}(k) = (z^{(1)}(1), z^{(1)}(2), \dots, z^{(1)}(n)) \tag{7}$$

where  $z^{(1)}(k) = 0.5(m^{(1)}(k) + m^{(1)}(k - 1)), k = 2, 3, \dots, n.$

Set  $M_j^{(0)}$  as the sequence of the carbon emission control function:

$$M_j^{(0)}(k) = (m_j^{(0)}(1), m_j^{(0)}(2), \dots, m_j^{(0)}(n)) \tag{8}$$

$M_j^{(1)}$  is the 1-AGO sequence,

$$M_j^{(1)}(k) = (m_j^{(1)}(1), m_j^{(1)}(2), \dots, m_j^{(1)}(n)), \tag{9}$$

where  $m_j^{(1)}(k) = \sum_{i=1}^k m_j^{(0)}(i), k = 1, 2, \dots, n. Z^{(1)}$  is the adjacent mean-generating sequence of  $M^{(1)}$ :

Assuming that the carbon emissions at time  $t$  are  $M^{(0)}(t)$ , the total carbon emissions at time  $[0, t]$  can be expressed as follows:

$$M^{(0)}(t) = \frac{dM^{(1)}(t)}{dt}$$

Thus, Eq. (4) can be written as follows:

$$\frac{dM^{(1)}(t)}{dt} = aM^{(1)}(t) + b[M^{(1)}(t)]^2 + f(M_j^{(1)}(t); \omega) \tag{10}$$

Consider using the first-order difference form to replace the differential on the left-hand side of Eq. (4). Then, when  $t = k$ ,

$$\begin{aligned} \frac{dM^{(1)}(t)}{dt} \Big|_{t=k} &\approx \frac{\Delta M^{(1)}(t)}{\Delta t} \Big|_{t=k} = \frac{M^{(1)}(k) - M^{(1)}(k-1)}{k - (k-1)} = M^{(1)}(k) - M^{(1)}(k-1) = M^{(0)}(k); \\ \int_{k-1}^k M^{(1)}(t) &= \frac{1}{2}(M^{(1)}(k-1) + M^{(1)}(k)) = Z^{(1)}(k); \\ \int_{k-1}^k f(M_j^{(1)}(t); \omega) &= \frac{1}{2}(f(M_j^{(1)}(k); \omega) + f(M_j^{(1)}(k-1); \omega)) \end{aligned}$$

where  $M^{(0)}(k)$  can be regarded as the grey derivative sequence of  $M^{(1)}(t)$  at time  $[k - 1, k]$ . Therefore, the grey prediction model is defined as follows:

**Definition 1** Let  $M^{(0)}, M^{(1)}, Z^{(1)}$  as shown in the definition of Eqs. (5)–(7). Then,

$$M^{(0)}(k) = aZ^{(1)}(k) + b[Z^{(1)}(k)]^2 + \frac{1}{2}(f(M_j^{(1)}(k); \omega) + f(M_j^{(1)}(k-1); \omega)) \tag{11}$$

This is called the carbon emission dynamic system grey model (CDGM(1,N)).

If  $\varphi(M_j^{(1)}(k), M_j^{(1)}(k-1), \omega) = \frac{1}{2}(f(m_j^{(1)}(k); \omega) + f(m_j^{(1)}(k-1); \omega))$

Thus, Eq. (11) can be written as follows:

$$M^{(0)}(k) = aZ^{(1)}(k) + b[Z^{(1)}(k)]^2 + \varphi(M_j^{(1)}(k), M_j^{(1)}(k-1), \omega) \tag{12}$$

The whitening equation of the CDGM(1,N) model is as follows:

$$\frac{dM^{(1)}(t)}{dt} = aM^{(1)}(t) + b[M^{(1)}(t)]^2 + \varphi(M_j^{(1)}(k), M_j^{(1)}(k-1), \omega) \tag{13}$$

Let  $\varphi(M_j^{(1)}(k), M_j^{(1)}(k-1), \omega) = \lambda(k)$ . Based on Eq. (13),

$$\frac{dM^{(1)}(t)}{dt} = aM^{(1)}(t) + b[M^{(1)}(t)]^2 + \lambda(k) \tag{14}$$

and

$$\begin{aligned} \frac{dM^{(1)}(t)}{bdt} &= [M^{(1)}(t)]^2 + \frac{a}{b}M^{(1)}(t) + \lambda(k) \\ \frac{dM^{(1)}(t)}{bdt} &= \left[M^{(1)}(t)^2 + \frac{a}{2b}\right]^2 + \frac{1}{b}\lambda(k) - \frac{a^2}{4b^2} \end{aligned}$$

Let  $n = -\frac{a}{2b}, \omega(k) = \frac{1}{b}\lambda(k) - \frac{a^2}{4b^2}, k = 1, 2, \dots, n.$  Thus, we have the following:

$$\frac{dM^{(1)}(t)}{bdt} = [M^{(1)}(t)^2 - n]^2 + \omega(k) \tag{15}$$

Integrating both parts of Eq. (15):

$$\int dM^{(1)}(t) = \int \{ [M^{(1)}(t)^2 - n]^2 + \omega(k) \} bdt$$

and

$$\int \frac{dM^{(1)}(t)}{[M^{(1)}(t)^2 - n]^2 + \omega(k)} = \int bdt \tag{16}$$

Because  $\omega(k)$  is a constant, Eq. (16) is the Riccati equation. From the literature (Duan and Pang 2021), we can obtain the time response function of the CDGM(1,N) model:

(1) When  $\omega(k) < 0, C_1 = \frac{1}{2\sqrt{\omega(1)}} \ln \left| \frac{m^{(1)}(1)+n-\sqrt{\omega(1)}}{m^{(1)}(1)+n+\sqrt{\omega(1)}} \right| - b$

$$\hat{M}^{(1)}(k) = \frac{\sqrt{\omega(k)} - n \mp (\sqrt{\omega(k)} - n)e^{2\sqrt{\omega(k)}(bk+C_1)}}{1 \pm e^{2\sqrt{\omega(k)}(bk+C_1)}} \quad k = 1, 2, \dots, n \tag{17}$$

(2) When  $\omega(k) = 0, C_2 = -\frac{1}{m^{(1)}(1)+n} - b$ , and

$$\hat{M}^{(1)}(k) = -\frac{1}{bk + C_2} - n, \quad k = 1, 2, \dots, n \tag{18}$$

(3) When  $\omega(k) > 0, C_3 = \frac{1}{\sqrt{-\omega(1)}} \arctan \frac{m^{(1)}(1)+n}{\sqrt{-\omega(1)}} - b$ ,

$$\hat{M}^{(1)}(k) = \sqrt{-\omega(k)} \tan \left[ \sqrt{-\omega(k)}(bk + C_3) \right] - nk = 1, 2, \dots, n \tag{19}$$

(4) The reduction value is obtained as follows:

$$\hat{M}^{(0)}(k) = \hat{M}^{(1)}(k) - \hat{M}^{(1)}(k - 1), \quad k = 2, 3, \dots, n \tag{20}$$

### Neural network-carbon emission dynamic system grey model

In Eq. (4) of the evolutionary dynamic system of carbon emissions, the control function  $f(M_1(t); \omega)$  has many influencing factors, including energy consumption and GDP, so it can be used as a control function and redefined through neurons (Ma et al. 2021). Thus, it can be set as follows:

$$\varphi(M_j^{(1)}(k), M_j^{(1)}(k - 1), \theta) = \sum_{j=1}^l \beta_j g(m_j^{(1)}(k), m_j^{(1)}(k - 1), \omega_j, c_j) + c$$

where  $l$  indicates the number of sequences related to the control function, and  $m_j^{(1)}(k)$  is obtained from Eqs. (8)–(9). Thus, based on the CDGM(1,N) model, the following model can be defined:

**Definition 2** Let  $M^{(0)}, M^{(1)}, Z^{(1)}$  as shown in the definition of Eqs. (4)–(6). Then,

$$M^{(0)}(k) = aZ^{(1)}(k) + b[Z^{(1)}(k)]^2 + \sum_{j=1}^l \beta_j g(m_j^{(1)}(k), m_j^{(1)}(k - 1), \omega_j, c_j) + c \tag{21}$$

This is called the novel grey prediction model with a feed-forward neural network based on a carbon emission dynamic evolution system (NN-CDGM(1,N)).

where  $g(\cdot)$  is the activation function, which is the single hidden layer neural network. Consider the following hidden layer function:

$$g(m, w, c) = \frac{2}{1 + \exp(-w^T m + c)} - \alpha$$

$\alpha$  is the regulatory system, and the optimization is sought through a genetic algorithm. The specific methods can be found in the literature (Holland 1992).

Let  $W = (\omega_1, \omega_2, \dots, \omega_l), C = (c, c_1, c_2, \dots, c_l)$ . Based on Theorem 2,

$$\lambda(k) = \phi(M_j^{(1)}(k), M_j^{(1)}(k - 1), \omega) = \sum_{j=1}^l \beta_j g(m_j^{(1)}(k), m_j^{(1)}(k - 1), \omega_j, c_j) + c$$

Thus, the time response of the NN-CDGM(1,n) model is calculated by using Eqs. (16)–(19).

Next, the Levenberg–Marquardt (LM) algorithm (Holland 1992) is used to compute the neural grey system model. Let  $\theta = (\beta, W, C)^T, \beta = (\beta_1, \beta_2, \dots, \beta_l), W = (\omega_1, \omega_2, \dots, \omega_l), C = (c, c_1, c_2, \dots, c_l)$ . For convenience, the training error of each point is defined as follows:

$$e_k = M^{(0)}(k) - aZ^{(1)}(k) - b[Z^{(1)}(k)]^2 - \sum_{j=1}^l \beta_j g(m_j^{(1)}(k), m_j^{(1)}(k - 1), \omega_j, c_j) + c$$

Let  $\theta = (\beta, W, C)^T, \beta = (\beta_1, \beta_2, \dots, \beta_l), W = (\omega_1, \omega_2, \dots, \omega_l), C = (c, c_1, c_2, \dots, c_l)$ . The sum of squared errors can be expressed as follows:

$$E_D(a, b, \theta) = \sum_{k=2}^n e_k^2 = e^T e - (\Delta^2 E_D(a, b, \theta))^{-1} E_D(a, b, \theta)$$

where the first-order gradient can be defined as follows:

$$E_D(a, b, \theta) = J^T e$$

The second gradient can be given by a Taylor series:

$$E_D(a, b, \theta) = J^T J + S$$

Here, the  $S$  matrix is usually assumed to be the  $O$  matrix because its elements are the second partial derivatives of the error. The Jacobian formula is as follows:

$$J = [J(a), J(b), J(\beta), J(W), J(C)] \tag{22}$$

where

$$\begin{aligned}
 J(a) &= \begin{pmatrix} \frac{\partial e_2}{\partial a} \\ \frac{\partial e_3}{\partial a} \\ \vdots \\ \frac{\partial e_n}{\partial a} \end{pmatrix}_{(n-1) \times 1}, \quad J(b) = \begin{pmatrix} \frac{\partial e_2}{\partial b} \\ \frac{\partial e_3}{\partial b} \\ \vdots \\ \frac{\partial e_n}{\partial b} \end{pmatrix}_{(n-1) \times 1} \\
 J(\beta) &= \begin{pmatrix} \frac{\partial e_2}{\partial \beta_1} & \frac{\partial e_2}{\partial \beta_2} & \dots & \frac{\partial e_2}{\partial \beta_l} \\ \frac{\partial e_3}{\partial \beta_1} & \frac{\partial e_3}{\partial \beta_2} & \dots & \frac{\partial e_3}{\partial \beta_l} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial e_n}{\partial \beta_1} & \frac{\partial e_n}{\partial \beta_2} & \dots & \frac{\partial e_n}{\partial \beta_l} \end{pmatrix}_{(n-1) \times l} \\
 J(W) &= \begin{pmatrix} \frac{\partial e_2}{\partial \omega_1} & \frac{\partial e_2}{\partial \omega_2} & \dots & \frac{\partial e_2}{\partial \omega_{2l}} \\ \frac{\partial e_3}{\partial \omega_1} & \frac{\partial e_3}{\partial \omega_2} & \dots & \frac{\partial e_3}{\partial \omega_{2l}} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial e_n}{\partial \omega_1} & \frac{\partial e_n}{\partial \omega_2} & \dots & \frac{\partial e_n}{\partial \omega_{2l}} \end{pmatrix}_{(n-1) \times 2l} \\
 J(C) &= \begin{pmatrix} \frac{\partial e_2}{\partial c} & \frac{\partial e_2}{\partial c_1} & \dots & \frac{\partial e_2}{\partial c_l} \\ \frac{\partial e_3}{\partial c} & \frac{\partial e_3}{\partial c_1} & \dots & \frac{\partial e_3}{\partial c_l} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial e_n}{\partial c} & \frac{\partial e_n}{\partial c_1} & \dots & \frac{\partial e_n}{\partial c_l} \end{pmatrix}_{(n-1) \times (l+1)}
 \end{aligned} \tag{23}$$

The LM algorithm uses the error vector and Jacobian matrix to calculate the inverse matrix of  $J^T J$ , experimentally scales the identity matrix to avoid an inappropriate matrix, and finally updates the parameters as follows:

$$\begin{pmatrix} a \\ b \\ \theta \end{pmatrix}^{l+1} = \begin{pmatrix} a \\ b \\ \theta \end{pmatrix}^l - (J^T J + \mu I)^{-1} J^T e \tag{24}$$

where the coefficient  $\mu$  of the scaled identity matrix is updated with the change in  $E_D(a, b, \theta)$ , and the calculation scheme is updated in Martin and Mohammad (1994).

### The modeling steps of NN-CDGM(1,N)

To evaluate the validity and accuracy of the model, the APE, MAPE, MAE, RMSE, and  $R^2$  metrics are defined as follows. The evaluation criteria of MAPE are shown in Eq. (27) (Javed et al. 2020). The smaller the MAE and RMSE values are, the better the model performance is. The larger the value of  $R^2$  is, which is closer to 1, the better the model performance is.

$$APE = \left| \frac{x^{(0)}(i) - \hat{x}^{(0)}(i)}{x^{(0)}(i)} \right| \times 100\%, i = 1, 2, \dots, n \tag{25}$$

$$MAPE = \frac{1}{n-1} \left( \sum_{i=1}^n APE \right) \tag{26}$$

$$MAPE(\%) = \begin{cases} < 10 & \text{Highly accuracy forecast} \\ 10 - 20 & \text{Good forecast} \\ 20 - 50 & \text{Reasonable forecast} \\ > 50 & \text{Inaccurate forecast} \end{cases} \tag{27}$$

$$MAE = \frac{\sum_{i=1}^N |y_i - \hat{y}_i|}{N} \tag{28}$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2} \tag{29}$$

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y}_i)^2} \tag{30}$$

Based on the NN-CDGM(1,N) model established in sections “Neural network-carbon emission dynamic system grey model” and “The modeling steps of NN-CDGM(1,N),” the main prediction process of the model is introduced below.

**Step 1.** Input raw data  $M^{(0)}(k)$  as the main sequence;  $M_j^{(0)}(k) (j = 1, 2, 3, 4)$  are coal consumption, crude oil consumption, GDP, and population, respectively, and are the relevant sequences.

**Step 2.** The 1-AGO sequence  $M^{(1)}(k)$  of  $M^{(0)}(k)$ ,  $M_j^{(0)}(k) (j = 1, 2, 3, 4)$  and the mean value sequence  $z^{(1)}$  of  $M_j^{(1)}(k) (j = 1, 2, 3, 4)$  are calculated.

**Step 3.** By substituting the data obtained in Step 2 into Eqs. (22)–(24), parameters  $a, b, \beta_j, c_j, c$  can be obtained.

**Step 4.** The parameters obtained in Step 3 are substituted into Eqs. (17)–(20). The different time response functions are used to obtain simulated and reduced values.

**Step 5.** In step 4, the APE and MAPE values can be obtained, and then, the genetic algorithm is used to optimize the activation function to obtain the optimal  $\alpha$  value, where  $M^{(0)}$  is obtained when the lowest MAPE value is calculated.

The brief calculation process is shown in Fig. 3.

### Validity analysis

In this section, the validity of the model is first verified. The 10-year data on Beijing’s carbon emissions are divided into two types of modeling objects. For the first type, the first 8 data points are used for simulation, and the last 2 data points are used for prediction. For the second type, the first 7 data points are used for simulation, and the last 3 data points are used for prediction. At the same time, the validity of 9 data points is analyzed. The modeling objects are also divided into two types: For the first type, the first 7 data points are used for simulation, and the last 2 data points are used for prediction. For the second type, the first 6 data points are used for simulation, and the last 3 data points are used for prediction. Compared with the other five classic grey multi-variable models, all of them have good effects. Finally, the best model is selected to predict the carbon emissions of Beijing in the 2019–2023 period.

Because the carbon emission system is a complex dynamic system, many factors affect carbon emissions. In this section, four factors, including coal consumption, crude oil consumption, GDP, and population, are selected as influencing factors, and the NN-CDGM(1, N) model is used to analyze the data on Beijing’s carbon emissions. The specific data on Beijing’s carbon emissions, coal consumption (hundred thousand tons), crude oil consumption (hundred thousand tons), GDP (billions of dollars), and population (hundred thousand people) are given in Table 1.

**Efficiency analysis of the NN-CDGM(1,N) model**

Regarding the data in Table 1, carbon emissions are taken as the main sequence; coal consumption, crude oil consumption, GDP, and population are taken as the relevant sequences; and the modeling steps of the NN-CDGM(1,N) model in “The modeling steps of NN-CDGM(1,N)” are used to model and forecast carbon emissions. Data of different lengths are selected for modeling, and the calculated results are compared with the classic first-order multivariable grey prediction models: GM(1,N) (Liu and Lin 2011), GMVM(1,N) (Jiang et al. 2020), GMC(1,N) (Tien 2005), and GOM(1,N) (Wu et al. 2015).

Regarding the data from 2009 to 2018, 8 data points are used for modeling, and 2 data points are used for prediction (case 1); additionally, 7 data points are used for modeling, and 3 data points are used for prediction (case 2). Regarding

the data from 2010 to 2018, 7 data points are used for modeling, and 2 data points are used for prediction (case 3); additionally, 6 data points are used for modeling, and 3 data points are used for prediction (case 4).

**Case 1. 8 data points for modeling and 2 data points for prediction**

Eight data points from 2009 to 2016 are used for modeling, and the 2-year carbon emissions of Beijing from 2017 to 2018 are predicted. Finally, the modeling and prediction results are compared with those of the GM(1,N), GMVM(1,N), GOM(1,N), and GMC(1,N) models and analyzed. The specific results are shown in Table 3. The calculation results of the model parameters are as follows:

$$M^{(0)}(k) = -0.0391Z^{(1)}(k) + 2.265E - 05 [Z^{(1)}(k)]^2 + \sum_{j=1}^l \beta_j g(m_j^{(1)}(k), m_j^{(1)}(k-1), \omega_j, c_j) - 0.9883$$

where

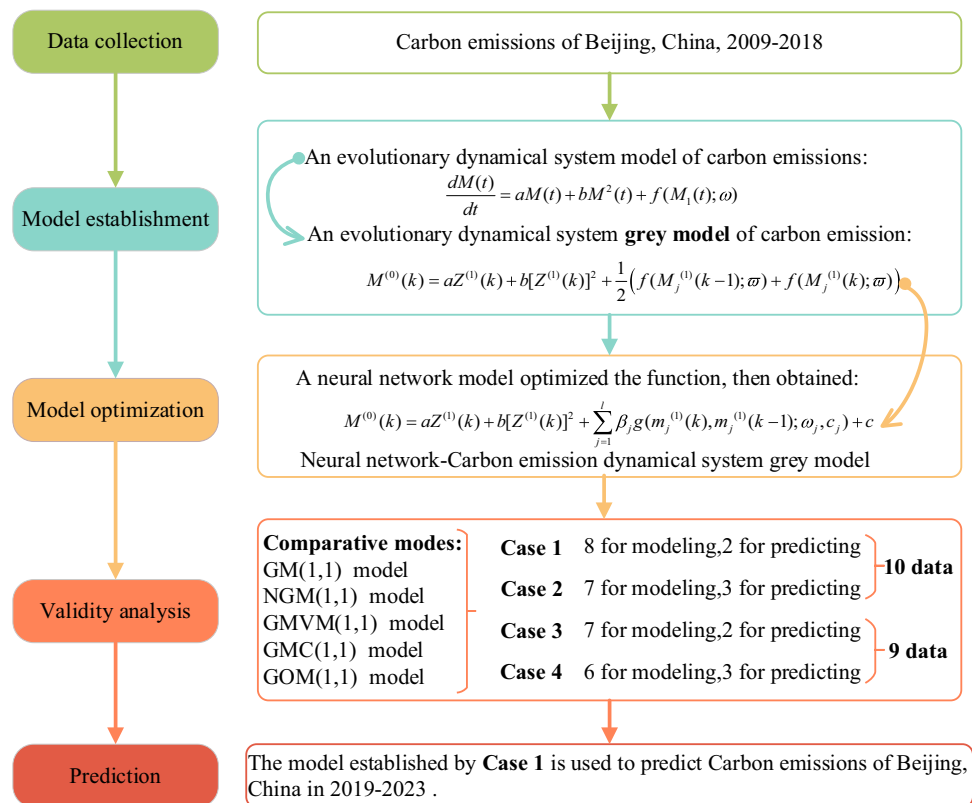
$$\beta = (-5.1840, -5.3631, -5.4956, -5.0651)$$

$$\omega = ((0.2613, 0.5678), (0.9615, 0.6130), (0.4976, 0.8070), (0.4278, 0.8808))$$

$$c = (0.0892, 0.3131, 0.6074, 0.3787)$$

Table 2 shows that all indexes of NN-CDGM(1,N) are the best, and the  $R^2$  value is very close to 1, indicating that

**Fig. 3** Flowchart of the NN-CDGM(1,N) model





**Table 1** Carbon emissions, coal consumption, crude oil consumption, GDP, and population of Beijing from 2009 to 2018

Year	Carbon emissions	Coal consumption	Crude oil consumption	GDP	Population
2009	100.40	266.47	116.29	129.01	186.00
2010	103.00	263.46	111.63	149.64	196.20
2011	94.40	236.60	110.51	171.89	202.40
2012	97.20	227.00	107.58	190.25	207.80
2013	93.40	201.92	87.09	211.35	212.50
2014	92.50	173.65	103.46	229.26	217.10
2015	92.17	116.52	99.15	247.79	218.80
2016	89.00	84.76	82.10	270.41	219.50
2017	85.00	49.05	89.25	298.83	219.40
2018	89.56	27.62	99.38	331.06	219.20

the simulated value closely matches the real value. In the simulation stage, the MAPE values of the NN-CDGM(1,N), GM(1,N), and GOM(1,N) models are all less than 5%, and the fitting effect is good. However, in the prediction stage, only the NN-CDGM(1,N) model has an accuracy below 5%. To further compare the differences between models, Fig. 4 is drawn based on the information in Table 2. The results of GMC(1,N) have negative values, which is inconsistent with reality. Thus, this model is not shown in Fig. 4.

Figure 4(a) clearly illustrates that the GMVM(1,N) model is always above the actual data and finally shows a sudden increasing trend with a large error. To compare the other models more clearly, GMVM(1,N) is removed, and Fig. 4(b) is shown. Clearly, the curve of the NN-CDGM(1,N) model is most consistent with the actual data curve. Although the MAPE values of the GM(1,N) and GOM(1,N) models are also very small, both models are clearly below the actual data at the front end of the data, and both show a sudden decreasing trend in the prediction stage, which is inconsistent with the actual curve trend. Therefore, the performance of the NN-CDGM(1,N) model is better in this case.

### Case 2. 7 data points for modeling and 3 data points for prediction

In this case, 7 data points from 2009 to 2015 are used to model and predict the carbon emissions of Beijing over the 3 years from 2016 to 2018. Finally, the modeling and prediction results are compared with those of the above comparison models and analyzed, and the specific results are shown in Table 3.

Table 3 shows that the MAPE value of GMC(1,N) is 1.4467%, which is the smallest among all the models, but its MAPE value is 39.1946%, showing a poor prediction effect. Its  $R^2$  is negative, and the simulated value does not match the real value. In general, the simulation effect of NN-CDGM(1,N), GM(1,N), and GOM(1,N) is better than that of the other models, but the overall error of NN-CDGM(1,N) is the smallest. To further compare the model

fitting and prediction effects, Table 3 is converted into Fig. 5. GMVM(1,N) yields negative values, which are not shown in the figure.

As illustrated in Fig. 5, the actual data show a relatively gentle fluctuation trend and eventually show a slight increase. The GMC(1,N) and NN-CDGM(1,N) models are most consistent with the actual data, especially in the simulation stage, when the GMC(1,N) model almost coincides with the actual data, but there is a sudden increase in the prediction stage, and the overall effect is not good. However, the NN-CDGM(1,N) model always fluctuates near the actual curve, and the final trend is consistent with the actual data. Thus, it is relatively superior overall.

### Case 3. 7 data points for modeling and 2 data points for prediction

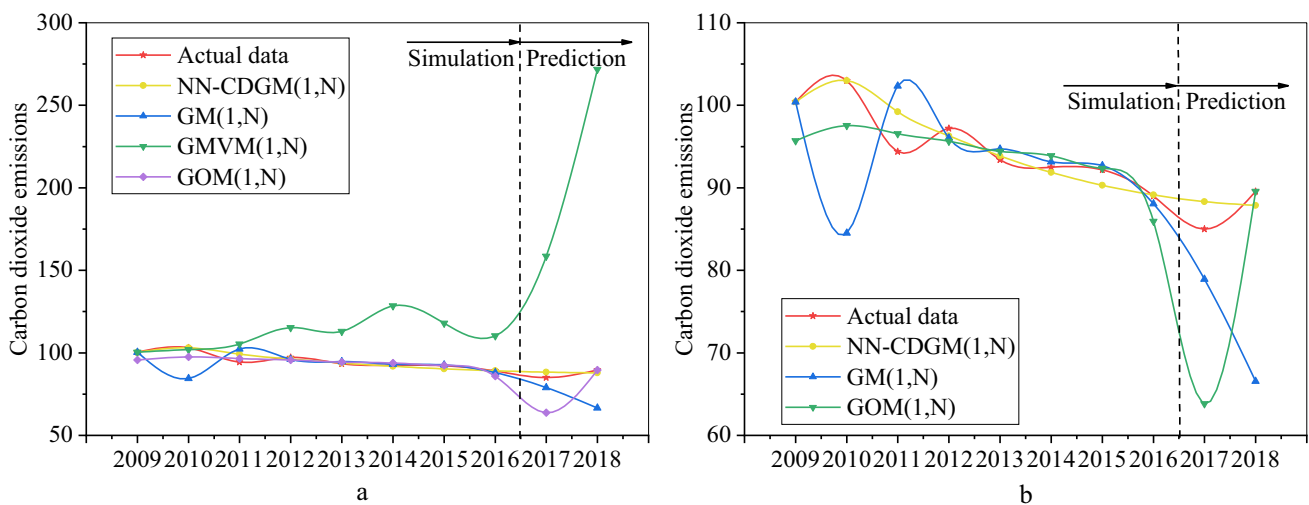
In this case, 7 data points from 2010 to 2016 are used to model and forecast the carbon emissions of Beijing for the 2 years from 2017 to 2018. Finally, the modeling and prediction results are compared with those of the above comparison models and analyzed, and the specific results are shown in Table 4.

Table 4 shows that for this set of data, the prediction errors of all models are above 5%, and the NN-CDGM(1,N) model has the best prediction effect. On the whole, the indexes of the NN-CDGM(1,N) model are also optimal, and the RMSE is significantly better than that of the comparison models. To observe the data trend more intuitively, Table 4 is plotted in Fig. 6. Similar to the figures above, GMC(1,N) yields negative values, which are not shown in this figure.

Figure 6(a) clearly shows that the GMVM(1,N) model has a large error with the actual data; therefore, it is removed from Fig. 3(b) for further comparison. In Fig. 6(b), NN-CDGM(1,N), GM(1,N), and GOM(1,N) all fluctuate near the actual data. However, the last two models show irregular fluctuations and are different from the trend of the actual data; thus, the prediction effect is not good. The curve of NN-CDGM(1,N) is always near the actual data curve. It is closest

**Table 2** Comparison of the results of multiple models based on 8 data points for modeling and 2 data points for prediction

Case1	Actual value	NN-CDGM(1,N)		GM(1,N)		GMVM(1,N)		GOM(1,N)		GMC(1,N)	
		Values	APE (%)	Values	APE (%)	Values	APE (%)	Values	APE (%)	Values	APE (%)
2009	100.40	100.40	0.00	100.40	0.00	100.40	0.00	95.70	4.68	100.40	0.00
2010	103.00	103.00	0.00	84.52	17.94	101.96	1.01	97.55	5.29	-10.63	110.32
2011	94.40	99.22	5.11	102.33	8.40	105.40	11.65	96.54	2.26	48.76	48.35
2012	97.20	96.29	0.93	96.05	1.18	115.21	18.52	95.63	1.61	48.90	49.69
2013	93.40	93.86	0.49	94.70	1.39	113.05	21.04	94.42	1.09	48.08	48.52
2014	92.50	91.87	0.68	93.13	0.68	128.45	38.87	93.87	1.48	47.46	48.69
2015	92.17	90.31	2.02	92.70	0.57	117.90	27.91	92.36	0.20	47.14	48.86
2016	89.00	89.13	0.14	88.07	1.05	110.30	23.93	85.93	3.45	45.66	48.70
MAPEs			<b>1.56</b>		4.46		20.42		2.87		57.59
2017	85.00	88.32	3.90	78.93	7.14	158.58	86.56	63.83	24.91	41.84	50.78
2018	89.56	87.87	1.89	66.55	25.70	271.82	203.50	89.56	0.00	36.11	59.68
MAPEp			<b>2.07</b>		16.42		145.03		12.45		55.23
MAE			<b>1.38</b>		6.00		38.85		4.07		48.29
RMSE			<b>2.05</b>		9.88		64.71		7.21		54.80
R <sup>2</sup>			<b>0.84</b>		-2.71		-158.27		-0.98		-113.25



**Fig. 4** Comparison of the results of each model based on 8 data points for modeling and 2 data points for prediction

to the actual data, and therefore, the simulation and prediction effects in relation to the comparison models are the best.

### Case 4. 6 data points for modeling and 3 data points for prediction

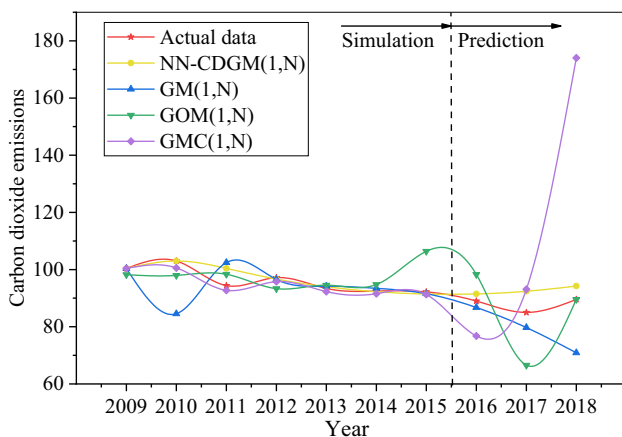
In this case, 6 data points from 2010 to 2015 are used for modeling, and the 3-year carbon emissions of Beijing from 2016 to 2018 are predicted. Finally, the modeling and prediction results are compared with those of the above comparison models and analyzed, and the specific results are shown in Table 5.

Table 5 shows that each evaluation index of the NN-CDGM(1,N) model is significantly better than that of the comparison model. The prediction results of the GMVM(1,N) models contain negative values, which is inconsistent with the actual situation. The NN-CDGM(1,N), GM(1,N) and GOM(1,N) models all have small simulation errors, but only the NN-CDGM(1,N) model has a prediction error of less than 10%, and its simulation error is also the smallest among all models. Figure 7 clearly compares the trends and fitting degrees of the models. The GMVM(1,N) model yields negative values, which are not shown in the figure.

Figure 7 illustrates that, except for the NN-CDGM(1,N) model, the other three models all show significant volatility.

**Table 3** Comparison of the results of multiple models based on 7 data points for modeling and 3 data points for prediction

Case 2	Actual value	NN-CDGM(1,N)		GM(1,N)		GMVM(1,N)		GOM(1,N)		GMC(1,N)	
		Values	APE (%)	Values	APE (%)	Values	APE (%)	Values	APE (%)	Values	APE (%)
2009	100.40	100.40	0.00	100.40	0.00	100.40	0.00	98.22	2.17	100.40	0.00
2010	103.00	103.00	0.00	84.53	17.93	99.56	3.34	97.95	4.90	100.59	2.34
2011	94.40	100.43	6.39	102.44	8.51	114.83	21.65	98.38	4.21	92.73	1.76
2012	97.20	96.66	0.55	96.58	0.64	121.27	24.77	93.34	3.97	95.80	1.44
2013	93.40	93.98	0.62	94.17	0.82	106.35	13.87	94.55	1.23	92.30	1.17
2014	92.50	92.26	0.26	93.42	1.00	101.55	9.78	94.77	2.46	91.55	1.03
2015	92.17	91.44	0.79	91.56	0.66	97.57	5.86	106.37	15.40	91.31	0.94
MAPEs			1.72		4.93		13.21		5.72		<b>1.45</b>
2016	89.00	91.50	2.80	86.76	2.51	-39.42	144.30	98.30	10.45	76.78	13.74
2017	85.00	92.42	8.73	79.72	6.21	-286.28	436.80	66.54	21.72	93.10	9.52
2018	89.56	94.26	5.25	70.90	20.83	-714.44	897.72	89.56	0.00	174.04	94.32
MAPEp			<b>5.60</b>		9.85		492.94		10.72		39.19
MAE			<b>2.27</b>		5.56		137.90		6.04		11.32
RMSE			<b>3.48</b>		8.88		283.20		8.34		27.14
R <sup>2</sup>			<b>0.54</b>		-2.00		-3049.83		-1.65		-27.01



**Fig. 5** Comparison of the results of each model based on 7 data points for modeling and 3 data points for prediction

Among them, the GMC(1,N) model has a trend that is similar to the actual data, but the range is large. The other two models show irregular and large fluctuations, which is inconsistent with the actual data trend; thus, the overall effect is poor. The NN-CDGM(1,N) model is consistent with the actual data trend, the curve is relatively close, and the effect is better.

**Case summary**

Based on the above four cases, the results obtained by the same model from different modeling objects vary. Due to the volatility of the original data, the NGM(1,N), GMVM(1,N), and GMC(1,N) models all show negative values in this

case, which is not in line with the actual situation, while the GM(1,N) and GOM(1,N) models show mutations in the simulation and prediction stages, respectively, making the overall effect poor.

However, the simulation and prediction accuracy of the four cases of the NN-CDGM(1,N) model are all approximately 5%, the models all perform well, and the results are close to the actual data, indicating that the performance of this model is relatively stable when modeling data of different lengths, and it is relatively less affected by the data length. The NN-CDGM(1,N) model is thus effective for predicting carbon emissions and has a good stability and adaptability.

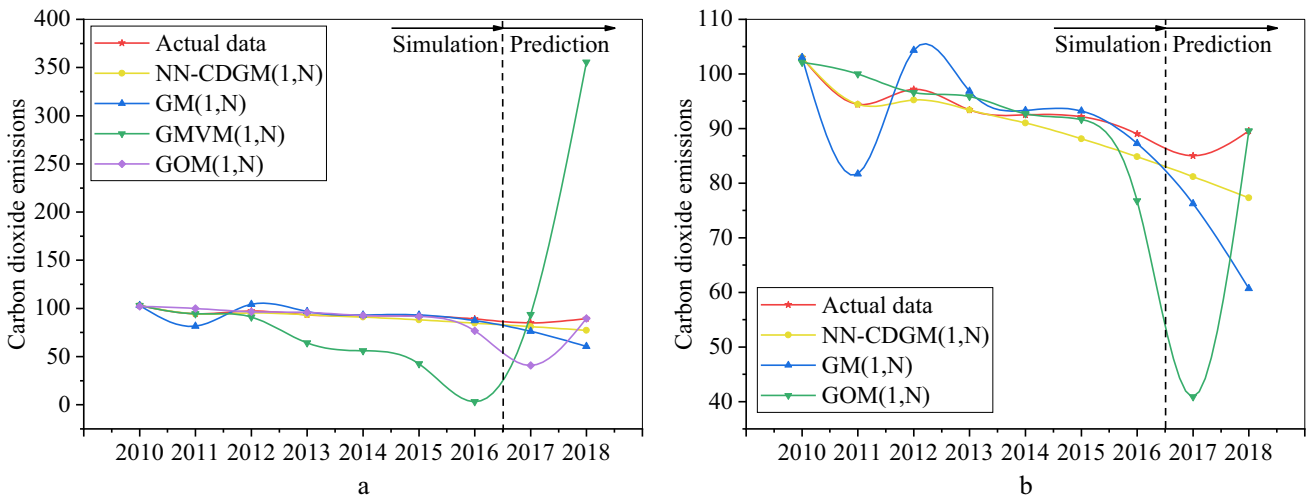
All the calculation results of the NN-CDGM(1,N) model in the four cases are shown in Table 6 below.

Table 6 shows that although the average error of the simulation and prediction of the four groups of results is less than 10% and all of them show a good performance, due to the different lengths of modeling data, the performance of the NN-CDGM(1,N) model also varies in each case. A further comparison shows that the evaluation indexes of case 1 are basically optimal, and the MAE and RMSE values are both small, indicating that the overall error is small, and the R<sup>2</sup> value is 0.84, close to the maximum value of 1, which means that the simulated value of the model matches the real value, and the model performance is good.

In conclusion, compared with multiple comparison models, the NN-CDGM(1,N) model proposed in this paper performs better in the four cases of predicting carbon emissions in Beijing, and the performance of modeling different lengths of data is relatively stable. In addition, by comparing the NN-CDGM(1,N) model among the four cases, the

**Table 4** Comparison of the results of multiple models based on 7 data points for modeling and 2 data points for prediction

Case 3	Actual value	NN-CDGM(1,N)		GM(1,N)		GMVM(1,N)		GOM(1,N)		GMC(1,N)	
		Values	APE (%)	Values	APE (%)	Values	APE (%)	Values	APE (%)	Values	APE (%)
2010	103.00	103.00	0.00	103.00	0.00	103.00	0.00	102.14	0.84	103.00	0.00
2011	94.40	94.40	0.00	81.69	13.47	94.34	0.06	99.99	5.92	-17.66	118.71
2012	97.20	95.24	2.02	104.35	7.36	90.93	6.46	96.59	0.63	46.99	51.66
2013	93.40	93.39	0.01	96.83	3.67	64.20	31.26	95.86	2.63	46.81	49.88
2014	92.50	91.00	1.62	93.30	0.86	56.01	39.45	92.70	0.22	45.98	50.29
2015	92.17	88.13	4.39	93.23	1.14	42.52	53.86	91.64	0.57	45.73	50.38
2016	89.00	84.84	4.68	87.29	1.92	3.24	96.36	76.76	13.75	44.56	49.94
MAPEs			<b>2.54</b>		4.74		37.91		4.09		61.81
2017	85.00	81.20	4.47	76.24	10.31	93.77	10.32	40.89	51.89	39.37	53.68
2018	89.56	77.29	13.70	60.73	32.19	355.59	297.04	89.56	0.00	30.57	65.86
MAPEp			<b>9.09</b>		21.25		153.68		25.94		59.77
MAE			<b>3.08</b>		7.16		53.58		7.40		50.10
RMSE			<b>4.77</b>		11.24		95.97		15.40		56.87
R <sup>2</sup>			<b>0.22</b>		-3.32		-314.31		-7.12		-109.72



**Fig. 6** Comparison of the results of each model based on 7 data points for modeling and 2 data points for prediction

overall performance of case 1 is the best. Therefore, the NN-CDGM(1,N) model calculated from case 1 is selected to predict the emissions in the next 5 years.

### Application

#### Forecast of carbon emissions in Beijing from 2019 to 2023

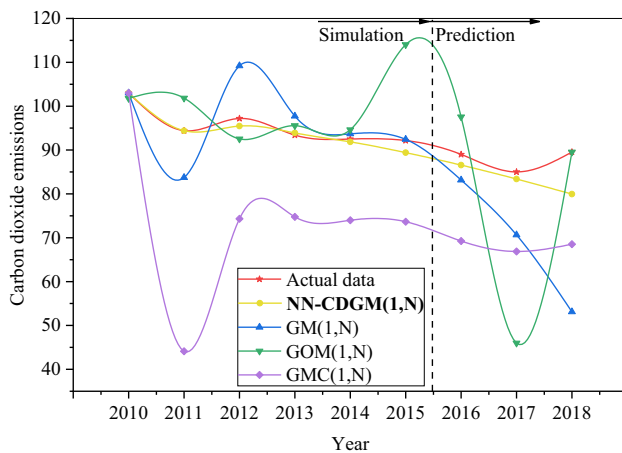
The above validity analysis shows that for the same set of data, there will be some differences in the results when different lengths of data are used for modeling (see Fig. 8 for details). Among the five comparison models, the simulation

performance of the GOM(1,N) and GMVM(1,N) models is better than the prediction performance. GM(1,N) and GMC(1,N) fluctuate greatly at the beginning of the simulation, and the predicted values in different cases are relatively scattered.

The actual data fluctuate between 90 and 105, but the simulated values of the comparison model range from 50~110 (GM(1,N) model) to -700 approx. 400 (GMVM(1,N) model). Although three cases of the NN-CDGM(1,N) model in Fig. 9 show some deviation from the actual results, the fluctuation is still controlled between 75 and 105, which is relatively stable. Table 6 and Fig. 9 show that when the NN-CDGM(1,N) model uses the same data, the model performance is not completely the same, but the difference is within 20, which is relatively stable.

**Table 5** Comparison of the results of multiple models based on 6 data points for modeling and 3 data points for prediction

Case 4	Actual value	NN-CDGM(1,N)		GM(1,N)		GMVM(1,N)		GOM(1,N)		GMC(1,N)	
		Values	APE (%)	Values	APE (%)	Values	APE (%)	Values	APE (%)	Values	APE (%)
2010	103.00	103.00	0.00	103.00	0.00	103.00	0.00	101.78	1.19	103.00	0.00
2011	94.40	94.40	0.00	83.72	11.31	94.38	0.02	101.84	7.88	44.11	53.28
2012	97.20	95.47	1.78	109.20	12.34	95.16	2.10	92.53	4.80	74.33	23.53
2013	93.40	93.91	0.54	97.73	4.64	74.64	20.09	95.56	2.31	74.76	19.95
2014	92.50	91.87	0.68	93.68	1.27	73.00	21.08	94.63	2.30	73.98	20.02
2015	92.17	89.41	3.00	92.44	0.29	67.72	26.53	114.03	23.71	73.64	20.10
MAPEs			<b>1.50</b>		5.97		13.96		8.44		27.38
2016	89.00	86.56	2.74	83.16	6.56	-4.79	105.39	97.55	9.61	69.26	22.17
2017	85.00	83.40	1.89	70.64	16.89	-43.36	151.01	45.99	45.89	66.88	21.32
2018	89.56	79.96	10.72	53.11	40.70	-96.48	207.73	89.56	0.00	68.52	23.50
MAPEp			<b>5.11</b>		21.38		154.71		18.50		22.33
MAE			<b>2.27</b>		5.56		137.90		6.04		11.32
RMSE			<b>3.48</b>		8.88		283.20		8.34		27.14
R <sup>2</sup>			<b>0.54</b>		-2.00		-3049.83		-1.65		-27.01

**Fig. 7** Comparison diagram of the results of each model based on 6 data points for modeling and 3 data points for prediction

In summary, the NN-CDGM(1,N) model yields different results for different modeling objects but is relatively stable and can effectively simulate and predict carbon emissions. In addition, the performance of the NN-CDGM(1,N) model in the four cases is observed and compared through Fig. 9. Clearly, the simulation and prediction results are the best when the data from 2009 to 2018 are selected, 8 data points are used for modeling, and 2 data points are used for prediction. Therefore, the modeling results of these data are selected to predict the carbon emissions of Beijing from 2019 to 2023, and the specific results are shown in Table 7.

The actual data from 2009 to 2018 in Fig. 9 show that the overall carbon emissions in Beijing have experienced a slight downwards trend with fluctuations, and the data in

2018 were in the early stage of rising fluctuations. The prediction results of the NN-CDGM(1,N) model for the next 5 years indicate that Beijing's carbon emissions will experience a slight upwards trend, which is consistent with the fluctuation trend of the actual data.

## Discussion and policy recommendations

Based on the relevant data used in the modeling (Table 1), over the past decade, Beijing's population growth has been small, coal and oil consumption has continued to decline, but the city's GDP has continued to increase, indicating that Beijing's policies and measures to optimize the industrial structure and support the clean energy transition have been very effective.

From the prediction results in Table 7, in the next 5 years, Beijing's carbon emissions will show a fluctuation trend, first declining and then rising, which is consistent with the previous fluctuation trend and conforms to the actual situation. In addition, carbon emissions are basically controlled below 90, although still a slight upwards trend is still apparent, and the growth rate is decreasing year by year and only increasing by less than 2% compared with 2018. This suggests that Beijing has entered the reduction platform period and needs to continue to implement effective measures to meet the existing conditions for further breakthroughs. We must continue to strictly control carbon emissions, further improve programs and measures, gradually introduce emission standards for industries, and better control the carbon emissions of industries and products with high levels of energy consumption and emissions to achieve the goal of continuously reducing carbon emissions.

**Table 6** Modeling and forecasting results of the NN-CDGM(1,N) model of carbon emissions in Beijing from 2009 to 2018

Year	Raw data	Case 1		Case 2		Case 3		Case 4	
		Values	APE (%)	Values	APE (%)	Values	APE (%)	Values	APE (%)
2009	100.40	100.40	0.00	100.40	0.00	103.00	0.00	103.00	0.00
2010	103.00	103.00	0.00	103.00	0.00	94.40	0.00	94.40	0.00
2011	94.40	99.22	5.11	100.43	6.39	95.24	2.02	95.47	1.78
2012	97.20	96.29	0.93	96.66	0.55	93.39	0.01	93.91	0.54
2013	93.40	93.86	0.49	93.98	0.62	91.00	1.62	91.87	0.68
2014	92.50	91.87	0.68	92.26	0.26	88.13	4.39	89.41	3.00
2015	92.17	90.31	2.02	91.44	0.79	84.84	4.68	86.56	2.74
2016	89.00	89.13	0.14	91.50	2.80	81.20	4.47	83.40	1.89
2017	85.00	88.32	3.90	92.42	8.73	77.29	13.70	79.96	10.72
2018	89.56	87.87	1.89	94.26	5.25				
MAPEs			1.56		1.72		2.54		<b>1.50</b>
MAPEp			<b>2.07</b>		5.60		9.09		5.11
MAE			<b>1.38</b>		2.27		3.08		2.14
RMSE			<b>2.05</b>		3.48		4.77		3.53
R <sup>2</sup>			<b>0.84</b>		0.54		0.22		0.57

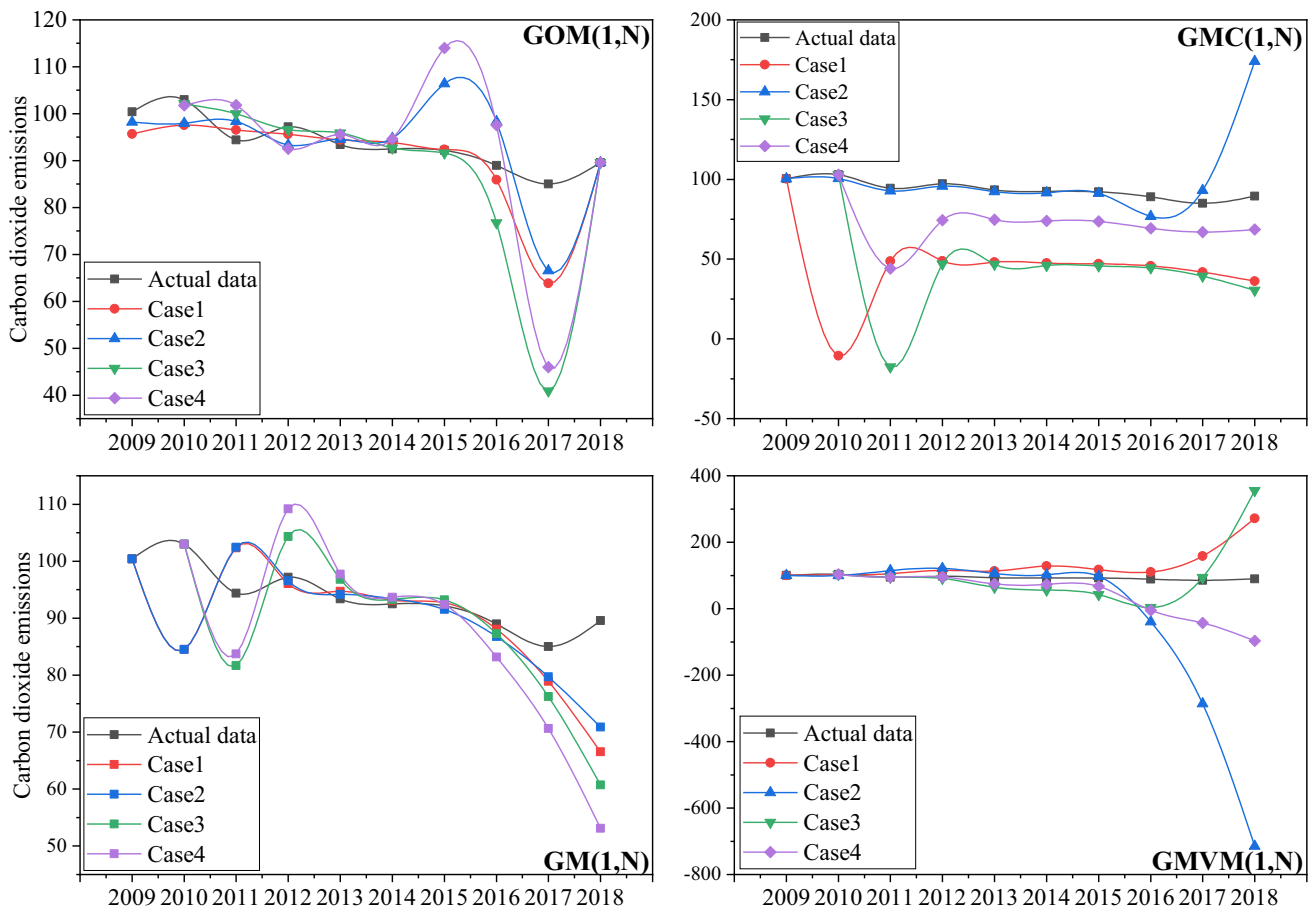


Fig. 8 Comparison of the modeling and prediction results of different lengths of data in the same model

### Conclusions

Carbon emissions affect the future development of the world. The scientific and effective prediction and analysis of carbon emission data are very important for the country and

the government to implement corresponding measures. The global energy mix has undergone major changes in recent years, and the COVID-19 pandemic has exacerbated the problem. Historical data on energy consumption and carbon emissions no longer apply to current realities.

First, the existing carbon emission prediction methods are mostly based on data and seldom consider the characteristics of the carbon emission system. Therefore, based on the background of the carbon emission dynamic system, this paper introduces a dynamic model with an external function. Second, considering the advantages of the grey prediction model in the application of small sample data, a dynamic grey prediction model of carbon emissions is established. Finally, the carbon emission system is affected by many factors, and the feedforward neural network model is introduced

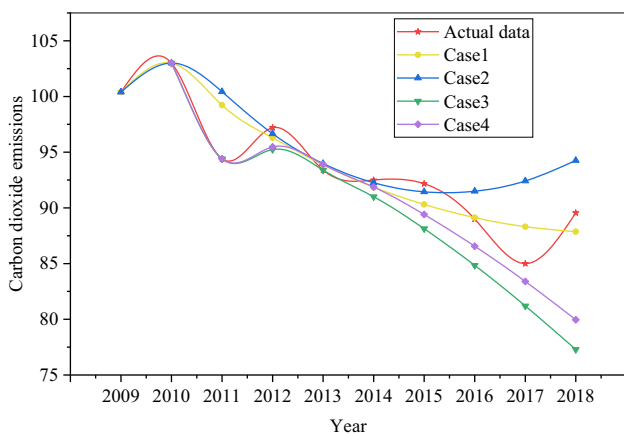


Fig. 9 Comparison of the results of the NN-CDGM(1,N) model with different lengths

Table 7 Prediction results of carbon emissions in Beijing from 2019 to 2023

Year	2019	2020	2021	2022	2023
Data	87.77	88.14	88.74	89.71	91.06

to calculate the function of other external factors in the calculation process.

This model is based on the dynamic system of the carbon emission system. Grey theory is used as the foundation, and the computational advantages of the neural network model are leveraged to yield results that are consistent with the actual change in carbon emissions. In the application process, the effectiveness of the model is verified by considering different modeling objects, and the results show that the NN-CDGM(1,N) model is basically better than other grey prediction models in the application of carbon emissions and that the model can effectively predict the carbon emissions of Beijing in the next 5 years.

This paper makes three major contributions: (1) the grey prediction model is established based on the carbon emission system; therefore, the scope of the grey prediction model is expanded from the aspect of the modeling structure. (2) In the process of modeling and calculation, the grey prediction model is organically combined with a neural network, which makes the model suitable for small sample data. (3) The NN-CDGM(1,N) model can effectively predict the carbon emissions of Beijing, China.

The model shows a good performance in this application, but some problems remain. In future research, the existing model can be further optimized based on the model structure and modeling mechanism, the selection of relevant factors, data pre-processing, and parameter optimization. The effectiveness of the model for predicting carbon emissions can be further verified through theoretical derivation and the use of additional data. The model features can be generalized, while retaining the good modeling properties, and applied to other domains. These questions need to be further discussed and studied in the future.

**Acknowledgements** The authors are grateful to the editors and the reviewers for their insightful comments and suggestions.

**Author contribution** The manuscript was reviewed and approved for publication by all authors. Weige Nie performed the experiments. Ou Ao analyzed the data. Weige Nie wrote the paper. Weige Nie, Ou Ao, and Huiming Duan reviewed and revised the paper.

**Funding** This work is supported by the Chongqing Natural Science Foundation of China (cstc2020jcyj-msxmX0649); National Natural Science Foundation of China (72171031); Social Science Planning and Doctoral Program of Chongqing, China (2020BS58).

**Data availability** The availability of data and materials on the base of personal request.

## Declarations

**Ethical approval** The manuscript was reviewed and ethical approved for publication by all authors.

**Consent to participate** The manuscript was reviewed and consents to participate by all authors.

**Consent to publish** The manuscript was reviewed and consents to publish by all authors.

**Conflict of interest** The authors declare no competing interests.

## References

- Ashin Nishan MK, Muhammed Ashiq V (2020) Role of energy use in the prediction of CO<sub>2</sub> emissions and economic growth in India: evidence from artificial neural networks (ANN). *Environ Sci Pollut Res* 27:23631–23642. <https://doi.org/10.1007/s11356-020-08675-7>
- Bezuglov A, Comert G (2016) Short-term freeway traffic parameter prediction: application of grey system theory models. *Expert Syst Appl* 62:284–292. <https://doi.org/10.1016/j.eswa.2016.06.032>
- Boamah KB, Du J, Adu D et al (2021) Predicting the carbon dioxide emission of China using a novel augmented hypo-variance brain storm optimisation and the impulse response function. *Environ Technol* 42:4342–4354. <https://doi.org/10.1080/09593330.2020.1758217>
- BP (2021) Statistical review of world energy, 70th edn. <https://www.bp.com/>. Accessed 11 Jan 2022
- Cao Y, Yin KD, Li XM, Zhai CC (2021) Forecasting CO<sub>2</sub> emissions from Chinese marine fleets using multivariable trend interaction grey model. *Appl Soft Comput J* 104. <https://doi.org/10.1016/j.asoc.2021.107220>
- Duan HM, Luo XL (2020) Grey optimization Verhulst model and its application in forecasting coal-related CO(2) emissions. *Environ Sci Pollut Res*. <https://doi.org/10.1007/s11356-020-09572-9>
- Duan HM, Luo XL (2021) A novel multivariable grey prediction model and application in forecasting coal consumption. *ISA Transaction*. <https://doi.org/10.1016/j.isatra.2021.03.024>
- Duan HM, Pang XY (2021) A multivariate grey prediction model based on energy logistic equation and its application in energy prediction in China. *Energy* 229. <https://doi.org/10.1016/j.energy.2021.120716>
- Duan HM, Wang D, Pang XY, Liu YM, Zeng SH (2020a) A novel forecasting approach based on multi-kernel nonlinear multivariable grey model: a case report. *J Clean Prod* 260. <https://doi.org/10.1016/j.jclepro.2020a.120929>
- Duan HM, Xiao XP, Long J, Liu YZ (2020b) Tensor alternating least squares grey model and its application to short-term traffic flows. *Appl Soft Comput J* 89. <https://doi.org/10.1016/j.asoc.2020b.106145>
- Deng JL (2002) Estimate and decision of grey system. Huazhong University of Science and Technology Press, Wuhan
- Gao MY, Yang HL, Xiao QZ, Mark G (2022a) A novel method for carbon emission forecasting based on Gompertz's law and fractional grey model: evidence from American industrial sector. *Renew Energy* 181:803–819. <https://doi.org/10.1016/j.renene.2021.09.072>
- Gao MY, Yang HL, Xiao QZ, Mark G (2022b) COVID-19 lockdowns and air quality: evidence from grey spatiotemporal forecasts. *Socioecon Plann Sci*. <https://doi.org/10.1016/j.seps.2022.101228>
- Heydari A, Garcia DA, Keynia F et al (2019) Renewable energies generation and carbon dioxide emission forecasting in microgrids and national grids using GRNN-GWO methodology. *Energy Procedia* 159:154–159. <https://doi.org/10.1016/j.egypro.2018.12.044>
- Holland JH (1992) Genetic Algorithms. *Sci Am* 267:66–72. <https://doi.org/10.1038/scientificamerican0792-66>
- Hu Y, Lv KJ (2020) Hybrid prediction model for the interindustry carbon emissions transfer network based on the grey model and



- general vector machine. IEEE Access. <https://doi.org/10.1109/ACCESS.2020.2968585>
- Huo TF, Ma YL, Cai WG et al (2021a) Will the urbanization process influence the peak of carbon emissions in the building sector? Dynamic Scenario Simulation Energy Build 232:110590. <https://doi.org/10.1016/j.enbuild.2020.110590>
- Huo TF, Xu LB, Feng W et al (2021b) Dynamic scenario simulations of carbon emission peak in China's city-scale urban residential building sector through 2050. Energy Policy 159:112612. <https://doi.org/10.1016/j.enpol.2021.112612>
- Huang YS, Shen L, Liu H (2019) Grey relational analysis, principal component analysis and forecasting of carbon emissions based on long short-term memory in China. J Clean Prod 209:415–423. <https://doi.org/10.1016/j.jclepro.2018.10.128>
- IEA, Greenhouse Gas Emissions from Energy (2021) <https://www.iea.org/data-and-statistics/data-product/co2-emissions-from-fuel-combustion>. Accessed 11 Jan 2022
- Ikram M, Sroufe R, Zhang QY et al (2021) Assessment and prediction of environmental sustainability: novel grey models comparative analysis of China vs the USA. Environ Sci Pollut Res 28:17891–17912. <https://doi.org/10.1007/s11356-020-11418-3>
- Javed SA, Ikram M, Tao L, Liu S (2020) Forecasting key indicators of China's inbound and outbound tourism: optimistic-pessimistic method. J Grey Syst. <https://doi.org/10.1108/GS-12-2019-0064>
- Jiang H, Kong P, Hu YC et al (2020) Forecasting China's CO<sub>2</sub> emissions by considering interaction of bilateral FDI using the improved grey multivariable Verhulst model, Environment. Dev Sustain. <https://doi.org/10.1007/s10668-019-00575-2>
- Liu SF, Lin Y (2011) Grey systems: theory and applications. Chapter 7. Beijing Science Press, p 169
- Liu C, Wu WZ, Xie WL, et al (2020) Application of a novel fractional grey prediction model with time power term to predict the electricity consumption of India and China. Chaos, Solitons & Fractals 141. <https://doi.org/10.1016/j.chaos.2020.110429>
- Mao SH, Kang YX, Zhang YH, Xiao XP, Zhu HM (2020) Fractional grey model based on non-singular exponential kernel and its application in the prediction of electronic waste precious metal content. ISA Trans 107. <https://doi.org/10.1016/j.isatra.2020.07.023>
- Martin TH, Mohammad BM (1994) Training feedforward networks with the marquardt algorithm. IEEE Trans Neural Networks 5:989–993. <https://doi.org/10.1109/72.329697>
- Modise RK, Mpofu K, Adenuga OT (2021) Energy and carbon emission efficiency prediction: applications in future transport manufacturing. Energies 14. <https://doi.org/10.3390/en14248466>
- Ma M, Ma X, Cai W (2019) Carbon-dioxide mitigation in the residential building sector: a household scale-based assessment. Energy Convers Manag 198. <https://doi.org/10.1016/j.enconman.2019.111915>
- Ma X, Xie M, Suykens JAK (2021) A novel neural grey system model with Bayesian regularization and its applications. Neurocomputing 456:61–75. <https://doi.org/10.1016/j.neucom.2021.05.048>
- Nguyen Q, Diaz-Rainey I, Kuruppuarachchi D (2021) Predicting corporate carbon footprints for climate finance risk analyses: a machine learning approach. Energy Econ 95. <https://doi.org/10.1016/j.eneco.2021.105129>
- Qiao WB, Lu HF, Zhou FZ et al (2020) A hybrid algorithm for carbon dioxide emissions forecasting based on improved lion swarm optimizer. J Clean Prod 244:1–37. <https://doi.org/10.1016/j.jclepro.2019.118612>
- Qiao ZR, Meng XM, Wu LF (2021) Forecasting carbon dioxide emissions in APEC member countries by a new cumulative grey model. Ecol Indic 125. <https://doi.org/10.1016/j.ecolind.2021.107593>
- Robati M, Daly D, Kokogiannakis G (2019) A method of uncertainty analysis for whole-life embodied carbon emissions (CO<sub>2-e</sub>) of building materials of a netzero energy building in Australia. J Clean Prod 225:541–553. <https://doi.org/10.1016/j.jclepro.2019.03.339>
- Sun W, Liu MH (2016) Prediction and analysis of the three major industries and residential consumption CO<sub>2</sub> emissions based on least squares support vector machine in China. J Clean Prod 122:144–153. <https://doi.org/10.1016/j.jclepro.2016.02.053>
- Tien TL (2005) The indirect measurement of tensile strength of material by the grey prediction model GMC(1, n). Meas Sci Technol 16:1322–1328. <https://doi.org/10.1088/0957-0233/16/6/013>
- Tian LX, Jin RL (2012) Theoretical exploration of carbon emissions dynamic evolutionary system and evolutionary scenario analysis. Energy 40:376–386. <https://doi.org/10.1016/j.energy.2012.01.052>
- Wang ZX, Li Q (2020) Forecasting the monthly iron ore import of China using a model combining empirical mode decomposition, non-linear autoregressive neural network, and autoregressive integrated moving average. Appl Soft Comput J. <https://doi.org/10.1016/j.asoc.2020.106475>
- Wu LF, Liu SF, Liu DL et al (2015) Modelling and forecasting CO<sub>2</sub> emissions in the BRICS (Brazil, Russia, India, China, and South Africa) countries using a novel multi-variable grey model. Energy 79:489–495. <https://doi.org/10.1016/j.energy.2014.11.052>
- Wang Q, Li SY, Zhang M, Li RR (2022) Impact of COVID-19 pandemic on oil consumption in the United States: a new estimation approach. Energy 239. <https://doi.org/10.1016/j.energy.2021.122280>
- Xia Y, Wang HJ, Liu WD (2019) The indirect carbon emission from household consumption in China between 1995–2009 and 2010–2030: a decomposition and prediction analysis. Comput Ind Eng 128:264–276. <https://doi.org/10.1016/j.cie.2018.12.031>
- Yan C, Wu LF, Liu LY, et al (2020) Fractional Hausdorff grey model and its properties, Chaos, Solitons Fractals 138. <https://doi.org/10.1016/j.chaos.2020.109915>
- Ye LL, Xie NM, Hu AQ (2021) A novel time-delay multivariate grey model for impact analysis of CO<sub>2</sub> emissions from China's transportation sectors. Appl Math Model 91:493–507. <https://doi.org/10.1016/j.apm.2020.09.045>
- Zeng B, Ma X, Shi JJ (2020) A new-structure grey Verhulst model for China's tight gas production forecasting. Appl Soft Comput J 96. <https://doi.org/10.1016/j.asoc.2020.106600>
- Zhao LT, Miao J, Shen Q, Chen XH (2021) A multi-factor integrated model for carbon price forecasting: market interaction promoting carbon emission reduction. Sci Total Environ 796. <https://doi.org/10.1016/j.scitotenv.2021.149110>
- Zhao HY, Wu LF (2020) Forecasting the non-renewable energy consumption by an adjacent accumulation grey model. J Clean Prod 275. <https://doi.org/10.1016/j.jclepro.2020.124113>
- Zhou WH, Zeng B, Wang JZ, et al (2021) Forecasting Chinese carbon emissions using a novel grey rolling prediction model. Chaos Solitons Fractals 147. <https://doi.org/10.1016/j.chaos.2021.110968>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.