Original Research

Prediction of CO_2 Emissions Based on the Analysis and Classification of Decoupling

Jianguo Zhou, Fengtao Guang*, Yameng Gao

Department of Economics and Management, North China Electric Power University,
689 Huadian Road, Baoding 071000, China

Abstract

This in-depth paper studies the issue of energy-related CO_2 emissions of China using sample data from 1980 to 2015. Due to the lack of official data, CO_2 emissions are first calculated by the recommended IPCC method. It shows that CO_2 emissions in China present an "S" type in shape. Then the Tapio decoupling index is applied to investigate the relationship between CO_2 emissions and economic growth. This suggests that weak decoupling is the main state during the study period and the decoupling trend is M-shaped. Moreover, the study years are divided into decoupling years and re-link years according to the decoupling relationship, and the ReliefF algorithm is proposed to verify the feasibility of the classification and judge the influencing weights of different driving factors. The ascending order is: actual GDP, urbanization rate, industrial structure, population, energy structure, and electricity consumption. Finally, a hybrid model of grey neural network model (GNNM) based on grey model (GM) and BP neural network (BPNN) is established to forecast CO_2 emissions. This demonstrates that the GNNM model has a better capacity for forecasting CO_2 emissions and capturing the non-linear and non-stationary characteristics of CO_2 emissions.

Keywords: CO_2 emissions, decoupling, ReliefF algorithm, GNNM model

Introduction

Today climate change, especially global warming, has seriously threatened the sustainable development of human beings. The endless emission of greenhouse gas, which gives priority to CO_2 emissions, is supposed to take the most responsibility for global warming. By 2050 CO_2 emissions from the energy field will be more than double today's figure if effective approaches to reduce them are not adopted [1]. So for the foreseeable future, statistics will paint a more sobering picture: frequent hurricanes, flood disasters, rising sea levels, and high temperatures.

China, as the largest developing country, has also overtaken America to become the largest country in respect to CO_2 emissions [2]. On the one hand, China is faced with increasing international pressure to reduce carbon emissions [3]; on the other hand, China is currently stepping into the critical process of industrial structure upgrading and energy structure optimization. Therefore, the development of a low-carbon economy and the establishment of low-carbon society are the inevitable choices and the focus of national strategies. On 30 June 2015 the Chinese government published Actions on Climate Change: China's Intended Nationally Determined Contributions (INDC) to control CO_2 emissions. INDC

*e-mail: Guangft@126.com

clearly indicates that CO_2 emissions will reach their peak around the year 2030 or much earlier than that, and CO_2 emissions will be reduced by 60-65% per unit of GDP in 2030 than in 2005. To accomplish the ambitious long-term goal, major progress should be made in building energy conservation and environmentally friendly society. Consequently, testing the relationship between CO_2 emissions and economic growth, studying the driving factors of CO_2 emissions, and forecasting CO_2 emissions for China are especially important and urgent.

There is abundant literature on the inseparable relationship between CO_2 emissions and economic growth in the past few years. Concerning research methods, the hypothesis of the Environmental Kuznets Curve (EKC) and the decoupling index have been extensively applied in the special issue of environmental quality and economic development. The EKC was first put forward by Panayotou [4], and the relationship between environmental quality and per capita income was regarded as an EKC. After that, the EKC has been expanded and adopted by many researchers. Schmalesee et al. [5] found that there was an inverted U type of EKC between CO_2 emissions and per capita income in developed countries. Moreover, some researchers [6-8] have replaced the per capita income with GDP and found that economic growth and CO_2 emissions did cause an inverted U type. However, Galeotti et al. [9] found that economic growth and CO_2 emissions were an N-shaped relationship. Due to the difference of research subjects, it is obvious that the relationship between carbon emissions and economic growth is uncertain. The decoupling index is a kind of useful technique to characterize the real-time dynamic index of the relationship between carbon emissions and economic growth. The decoupling model contains two kinds of models – the OECD decoupling index and Tapio decoupling index, both of which have their advantages and disadvantages for different applications. The Organization for Economic Cooperation and Development (OECD) implemented the concept of decoupling in 2002 for the first time in the report of the Indicator to measure Decoupling of Environmental Press from Economic Growth [10]. The OECD decoupling index is applied widely due to its simple operation, but it is too sensitive to the selection of the base year. The Tapio decoupling model is an elastic analysis first proposed by Tapio for his research on the volume of European transportation and decoupling status standards of CO_2 during 1970-2001 [11]. Compared with the OECD decoupling index, the Tapio decoupling index adopts the flexibility to study the decoupling relationship between CO_2 emissions and economic growth, which is not affected by changes of statistical dimensions [12].

Moreover, research on the driving factors of CO_2 emissions has also been given extensive attention. The approaches of these studies can be classified into four categories: decomposition analysis [13-14], bottom-up analysis [15-16], system optimization [17-18], and econometric approaches [19-20]. Decomposition analysis is the most commonly used method, which encompasses two main decomposition techniques: structural decomposition analysis (SDA) [13] and index decomposition analysis (IDA) [14]. Compared to SDA, IDA requires less data intensive and more diversified forms of decomposition than SDA. However, SDA is based on environmental input-output analysis [21], which is a widely-accepted method for quantifying sectoral CO_2 emissions considering both direct effects from the production process and indirect effects from the entire supply chain [22]. For the method of bottom-up analysis, Di Cosmo et al. [15] explored the effects of emission allowances on CO_2 emissions in the European Union, and Zhao et al. [16] examined the influence of technological progress on CO_2 emissions by the input-output model in Britain. The third method has been utilized to study the major influencing factors of CO_2 emissions for the European food industry [17], and explore the effect of renewable energy on CO_2 emissions of Japan [18]. The autoregressive distributed lag (ARDL) model [19] and vector autoregression (VAR) model [20] are the two common methods of econometric approaches. Yorucu [19] applied the ARDL model to investigate the influence of electricity consumption and foreign tourist arrivals on CO_2 emissions in Turkey's transport industry. Xu et al. [20] investigated how the main driving factors of CO_2 emissions have important significance of emission reduction in China using the VAR model.

Recently, the issue of CO_2 emissions forecasting has attracted considerable attention from many experts and scholars. Due to the complexity and diversity of the driving factors of CO_2 emissions, different kinds of approaches are proposed to deal with the prediction of CO_2. These approaches mainly fall into three kinds of categories: regressed based methods [23-25], scenario analysis [26], and artificial intelligence (AI) methods [27-29]. The regressed methods can be mainly divided into two types. The one, including GM (1, n) [23] model and logistic regression [24], adopts curve-fitting theory to forecast CO_2 emissions, which does not take into account the influencing factors of CO_2 emissions. Wang [23] divided the data of Chinese gross domestic product and carbon emissions from fossil energy consumption of 1953-2013 into 15 stages and used the non-linear grey model to quantify future Chinese carbon emissions from 2014 to 2020. Du et al. [24] verified that Chinese provincial carbon emissions can be classified into five categories by K-means clustering algorithm and forecasted provincial carbon emissions from 2011 to 2020 by the logistic model. The other is forecast CO_2 emissions based on the analysis of driving factors of CO_2 emissions. The IPAT model is one of the typical methods. Du et al. [25] modified the traditional IPAT model and forecasted the per capita CO_2 emissions of China on the basis of analyzing influencing factors of CO_2 emissions. Scenario analysis is becoming popular in forecasting CO_2 emissions, which generally depends on the government document and development plan. Moreover, this method is not used alone in the prediction period, but is combined with other methods. Zhang [26] examined the carbon emissions reduction potentials for China's power industry under different scenarios by the use of the long-range energy alternatives

planning (LEAP) model. With the development of artificial intelligence, the artificial methods have a number of applications within CO$_2$ emissions forecasting. Sun et al. [27] applied the least squares support vector machine to predict CO$_2$ emissions of three major industries, including the primary, secondary, and tertiary industries of China. Han et al. [28] proposed an energy and carbon emissions analysis and prediction approach based on an improved extreme machine (EM) integrated interpretative structural model (ISM). Li et al. [29] utilized the back propagation neural network (BPNN) to explore the carbon emissions reduction potential of China.

The contributions made by this paper are two-fold:

– From the previous literature, it can be concluded that the current studies only focus on the identification of the decoupling relationship between economic growth and CO$_2$ emissions. Therefore, further study is made by this paper. We classify the study years according to the decoupling relationship acquired by the Tapio decoupling index, and attempt to adopt the ReliefF algorithm to verify the feasibility of the classification. Moreover, the influencing weight of initially selected driving factors of CO$_2$ emissions can be obtained by the ReliefF algorithm. The influencing factors of CO$_2$ emissions whose influencing weight is greater than the threshold value will be regarded as the input variables.

– Establish a novel GM model and BPNN hybrid forecasting methodology to forecast CO$_2$ emissions. In our proposed methodology, the GM model is applied to address the problem of small sample data sets, and BPNN is introduced to handle the non-linear and non-stationary data of CO$_2$ emissions. And we also evaluate the forecasting performance of the single GM model and BPNN to demonstrate the higher prediction precision of the proposed GMMN.

Material and Methods

Influencing Factors Pre-Selection and Data Conversion

From the previous literature, it can be seen that the factors affecting CO$_2$ emissions mainly include GDP, energy structure, industrial structure, electricity consumption, urbanization rate, and population. To eliminate the impact of price, GDP is converted to 1980 prices by the index of GDP implicit deflator. The characteristics of Chinese energy are "lacking oil, little gas, rich coal," which shows that coal resources are the main energy source in China, and according to the government this structure will not change. Thus, coal accounted for the proportion of total primary energy consumption structure in this paper. Moreover, China began to speed up industrial restructuring and economic transformation since the implementation of the reform and opening up, and there has been a reduction in the proportion of primary industry and an increase in the proportion of secondary and tertiary industries. But the proportion of secondary industry is still the largest, so this study takes the proportion of secondary industry as industrial structure.

Taking into account how no official data of energy-related CO$_2$ emissions is scheduled in China for now, it is necessary to first calculate CO$_2$ emissions. In order to measure CO$_2$ accurately, the material balance method is utilized in this paper, which is based on the energy balance sheet. From the energy balance sheet, the energy consumption of all varieties can be acquired. To ensure the comparability of measurement results, carbon emission coefficients of various types of energy adopt the recommended value of IPCC [30]. Standard coal data are converted to values of CO$_2$ emissions by different conversion coefficients that are shown in Table 1. Through the numerical examples, it is noteworthy that various energies are assumed to achieve full burning.

Tapio Decoupling Index

Tapio Decoupling index can be calculated as the ratio of the percentage change of the accounting target to the percentage of the driving force from the base year to the target year. It can be expressed as the following formula.

$$R = \frac{\Delta C / C}{\Delta G / G} \tag{1}$$

…where r represents the decoupling elastic coefficient, C denotes the total amount of CO$_2$ emissions, and ΔC represents the change of CO$_2$ emissions between the base year and target year. G represents GDP and ΔG denotes the change of GDP in the same period. The eight decoupling states are presented in Table 2.

Based on the values of decoupling elastic coefficient, we need to divide up the study years. To simplify the criteria of characteristic classification and reduce the dimension of the classification data, there are only three decoupling states selected as the classification standard: negative decoupling, decoupling, and link.

Table 1. Conversion coefficients of different types of energy.

Energy	Conversion coefficient	Unit
Coal	Million ton	0.7143 kgce/kg
Coke	Million ton	0.9714 kgce/kg
Crude oil	Million ton	1.4286 kgce/kg
Gasoline	Million ton	1.4714 kgce/kg
Kerosene	Million ton	1.4714 kgce/kg
Diesel oil	Million ton	1.4571 kgce/kg
Fuel	Million ton	1.4286 kgce/kg
Natural Gas	Billion cubic meters	1.2721 kgce/m^3
Power	Billion kwh	0.1229 kgce/kwh

Table 2. Decoupling indicator criteria.

Classification standard	State	CO_2 emissions	Economic growth	Elasticity
Negative decoupling	Expansionary negative decoupling	>0	>0	r>1.2
	Strong negative decoupling	>0	<0	r<0
	Weak negative decoupling	<0	<0	0<r<0.8
Decoupling	Weak decoupling	>0	>0	0<r<0.8
	Strong decoupling	<0	>0	r<0
	Recession decoupling	<0	<0	r>1.2
Link	Growing link	>0	>0	0.8<r<1.2
	Recession link	<0	<0	0.8<r<1.2

ReliefF Algorithm

ReliefF algorithm is applied to identify, quantify, and construe the major influencing factors of CO_2 emissions, which is developed from the relief algorithm proposed by Kria. The relief algorithm is one of the widely used feature weight algorithms. According to the correlations between features and categories, different weights are assigned to each feature and the feature with a slighter weight greater than a certain threshold will be removed. With high efficiency, The relief algorithm has no requirement for data type, and is not sensitive to the relationship among features. But at the same time it is limited to the classification of two types of data. While the ReliefF algorithm can deal with multi-class and regression problems, it also can supplement the processed method under the conditions lacking data. Therefore, the ReliefF algorithm is recognized as the most efficient filter feature evaluation algorithm.

The specific steps of the ReliefF algorithm are:

Input: Take the initially selected influencing factors of CO_2 emissions as input dataset D. Each influencing factor denotes an eigenvalue (six in total – s = 6). Moreover, the data for each year represents a sample, and there 36 samples in total. Set the iteration times m to 100 and the number of near hits k to 1.

– **Step 1:** Normalization of the input dataset D by the following equation.

$$a = \frac{A - A_{min}}{A_{max} - A_{min}} \quad (2)$$

…where a denotes the normalization of the original data, and A_{max} and A_{min} denote the maximum and minimum of the original data, respectively

– **Step 2:** Set weight of each feature in the dataset to 0, namely, w(a) = 0, a = 1,2...s

– **Step 3:** Randomly select a sample R in dataset D

– **Step 4:** Find out the k-nearest samples of the same category and the different category in dataset D. $Hj(j = 1,2...k.)$ is similar to R, while $Mj(c)$ is not

– **Step 5:** Uprate the weight of each feature w(a) by the following – Eq. (3):

$$W(a) = W(a) \sum_{j=1}^{k} diff\left(a, R, H_j\right)/(mk)$$
$$+ \sum_{c \neq class(R)} \left[\frac{p(c)}{1 - p(class(R))} \sum_{j=1}^{k} diff(a, R, M_j(c)) \right] \bigg/ (mk)$$

$$(3)$$

…where $diff(a, R_1, R_2)$ represents the Euclidean distance between sample R_1 and sample R_2 on the basis of feature a, $p(c)$ is the probability of a sample belonging to the same class of c, and $Mj(c)$ denotes the j-th sample in the class c

– **Step 6:** Repeat step 5 for N = 40 times and figure out the average value of w(a)

– **Step 7:** Repeat steps 3 through 6 for m = 80 times, and output each mean weight for each feature

Output: Arrange the mean weight value of w(a) for each feature in descending order

Grey Neural Network Model

The Grey system theory, established by Deng in Ref. [31], is a methodology that concentrates on solving problems involving insufficient information and small samples. The method can be used for uncertain systems with partly known information by generating, mining, and extracting additional information from available data so that it can provide an accurate description of system behaviors and the hidden laws of evolution. The GM (1, n) model is researched on the most nature and adopted most widely in forecasting. While because of a lack of self-studying, self-adjusting, and self-organizing abilities, it is difficult to control prediction error and make timely responses to the new situation.

BP neural network (BPNN) is one of the back propagation algorithms of artificial neural network, and is also a way to imitate the behavior characteristics of animal neural networks for distributed parallel algorithm, which is for a mathematical model of information processing. With the better ability of nonlinear mapping and error feedback compensation, the model of BPNN can give relatively excellent prediction to target.

In order to take full advantage of similarities and complementarity between the GM model and BPNN and overcome the shortcoming of applying either model isolation, an integrated model (GNNM) combining the GM model and BPNN are proposed for forecasting carbon emissions.

The grey problem is a predictable problem targeting the development of the eigenvalue of the grey uncertain system behavior. Considering that the sequence x_t^1 obtained by Accumulated Generating Operation (AGO) of the original sequence $x_t^{(0)}(t = 0,1,2,...N-1)$ is exponentially distributed, a continuous function or differential equation can be employed in data simulation and prediction. To avoid confusion and awkward phrasing, the symbols of variable names have been redefined. The original sequence $x_t^{(0)}$ is replaced with x(t), the sequence $x_t^{(1)}$ obtained by first-order AGO is substituted by y(t), and the sequence of prediction results $x_t^{*(1)}$ is exchanged with z(t). Then the differential equation of GNNM model with n parameters can be expressed as follows.

$$\frac{dy_1}{dt} + ay_1 = b_1y_2 + b_2y_3 + \cdots + b_{n-1}y_n \quad (4)$$

...where $y_1, y_2,..., y_n$, denotes the system input parameters, y_1 is the system output parameter, and $a, b_1, b_2,..., b_{n-1}$ are coefficients of the differential equation.

The response equation of (4) is as follows:

$$z(t) = \left(y_1(0) - \frac{b_1}{a}y_2(t) - \frac{b_2}{a}y_3(t) - \cdots - \frac{b_{n-1}}{a}y_n(t)\right)e^{-at} + \frac{b_1}{a}y_2(t) + \frac{b_2}{a}y_3(t) + \cdots + \frac{b_{n-1}}{a}y_n(t)$$

$$(5)$$

Equation (5) can be transformed into:

$$z(t) = \left((y_1(0) - d) \cdot \frac{e^{-at}}{1+e^{-at}} + d \cdot \frac{1}{1+e^{-at}}\right) \cdot (1+e^{-at})$$
$$= \left((y_1(0) - d)\left(1 - \frac{1}{1+e^{-at}}\right) + d \cdot \frac{1}{1+e^{-at}}\right) \cdot (1+e^{-at})$$
$$= \left((y_1(0) - d) - y_1(0) \cdot \frac{1}{1+e^{-at}} + 2d \cdot \frac{1}{1+e^{-at}}\right) \cdot (1+e^{-at})$$

$$(6)$$

...where

$$d = \frac{b_1}{a}y_2(t) + \frac{b_2}{a}y_3(t) + \cdots + \frac{b_{n-1}}{a}y_n(t) \quad (7)$$

Then map the transformed Eq. (6) to an extended neural network, and a GNNM model with n input parameters and 1 output parameter will be established. Fig. 1 shows the network topology of the GNNM model.

In Fig. 1, t denotes the number of input parameters; $y_2(t),..., y_n(t)$ are network input parameters; $w_{21}, w_{22},..., w_{2n}, w_{31}, w_{32}, w_{3n}$, are network weights; y_1, is the network prediction value; and LA, LB, LC, LD represent the four layers of the neural network, respectively.
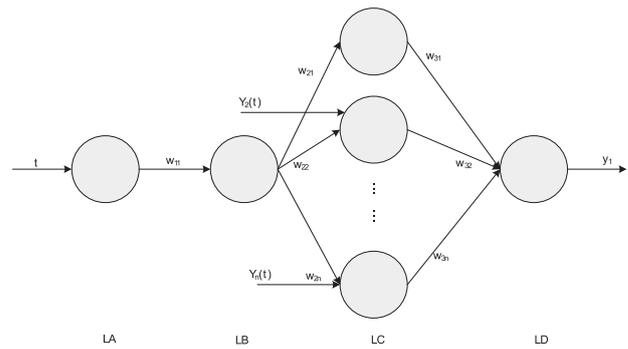

Fig. 1. Tetwork topology of the GNNM model.

Set $\frac{2b_1}{a} = u_1, \frac{2b_2}{a} = u_2, \cdots, \frac{2b_{n-1}}{a} = u_{n-1}$, then the network initial weights can be expressed as follows:

$$w_{11} = a, w_{21} = -y_1(0), w_{22} = u_1, w_{23} = u_2, \cdots, w_{2n} = u_{n-1}, w_{31} = w_{32} = \cdots = w_{3n} = 1 + e^{-at}$$

$$(8)$$

The threshold value of output node in the layer of LD is as follows:

$$\theta = (1 + e^{-at})(d - y_1(0)) \quad (9)$$

The specific learning process of GNNM is as follows:
- **Step 1**: Initialize the network structure based on the characteristics of training samples; set initialization parameters a and b, and calculate the value of u.
- **Step 2:** Calculate $w_{21}, w_{22},..., w_{2n}, w_{31}, w_{32}, w_{3n}$ based on the definition of network weight.
- **Step 3:** Calculate the output of different layers for each input sequence $(t, y(t))$, $t = 1, 2, 3,..., N$.

LA: $a = w_{11}t$ (10)

LB: $b = f(w_{11}t) = \frac{1}{1+e^{-w_{11}t}}$ (11)

LC: $c_1 = bw_{21}, c_2 = y_2(t)bw_{22}, c_3 = y_3(t)bw_3, \cdots, c_n = y_n(t)bw_{2n}$ (12)

LD: $d = w_{31}c_1 + w_{32}c_2 + \cdots + w_{3n}c_n - \theta_{y_1}$ (13)

- **Step 4:** Calculate the errors δ between predicted values and expected values, and adjust weights and threshold according to the calculated errors.

The error of LD: $\delta = d - y_1(t)$ (14)

The errors of LC:
$$\delta_1 = \delta(1 + e^{-w_{11}t}), \delta_2 = \delta(1 + e^{-w_{11}t}), \cdots, \delta_n = \delta(1 + e^{-w_{11}t})$$

$$(15)$$

The errors of LB:

$$\delta_{n+1} = \frac{1}{1 + e^{-w_{11}t}}\left(1 - \frac{1}{1 + e^{-w_{11}t}}\right)(w_{21}\delta_1$$
$$+ w_{22}\delta_2 + \cdots + w_{2n}\delta_n) \qquad (16)$$

Adjust weights from LB to LC.
$$w_{21} = -y_1(0), \cdots, w_{2n} = w_{2n} - u_{n-1}\delta_n b \qquad (17)$$

Adjust weights from LA to LB.
$$w_{11} = w_{11} + at\delta_{n+1} \qquad (18)$$

Adjust threshold
$$\theta = (1 + e^{-w_{11}t})\left(\frac{w_{22}}{2}y_2(t) + \frac{w_{23}}{2}y_3(t)\right.$$
$$\left. + \cdots + \frac{w_{2n}}{2}y_n(t) - y_1(0)\right) \qquad (19)$$

- **Step 5:** Determine whether the training is over; if not, return step 3.

Results and Discussion

An Overview of CO$_2$ Emissions

The total energy-related CO$_2$ emissions and economic activity of the Chinese economy (actual GDP) are presented in Fig. 2, which shows an S-type growing trend of CO$_2$ emissions and an upward of actual GDP during the study period. As shown in Fig. 2, total CO$_2$ emissions are embodied in China's energy consumption more than seven times, growing from 156,112.3 kt in 1980 to 1,113,700 kt in 2015. Actual GDP increased from 4,587.6 hundred million Yuan to 119,632.8 hundred million Yuan over the same period, with 69.66% annual increasing degree on average.

According to the developing trend of CO$_2$ emissions, it can be analyzed in three periods: 1980-99, 2000-07, and 2008-15. In the first period, both CO$_2$ emissions and actual GDP show no significant changes, having average annual growth rates of 6.66% and 25.18%, respectively. It is noteworthy that energy-related CO$_2$ emissions barely changed during the period of 1995-99. When entering the second period, CO$_2$ emissions grew explosively, with an
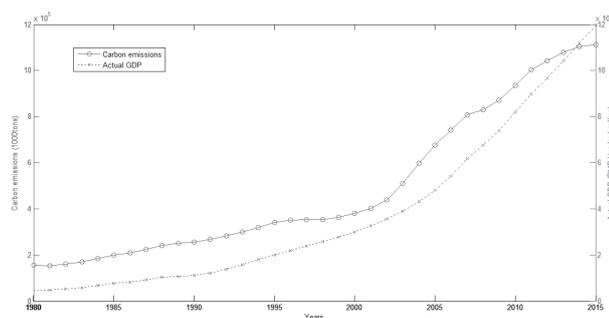

Fig. 2. Energy-related CO$_2$ emissions and actual GDP in China (1980-2015).

Table 3. Values of decoupling elasticity (1981-2015).

Year	$\Delta CO_2/CO_2$	$\Delta GDP/GDP$	R	Status
1981	-0.0137	0.0510	-0.2694	Strong decoupling
1982	0.0441	0.0900	0.4897	Weak decoupling
1983	0.0640	0.1080	0.5927	Weak decoupling
1984	0.0737	0.1520	0.4846	Weak decoupling
1985	0.0815	0.1350	0.6036	Weak decoupling
1986	0.0544	0.0878	0.6192	Weak decoupling
1987	0.0715	0.1169	0.6118	Weak decoupling
1988	0.0735	0.1123	0.6540	Weak decoupling
1989	0.0423	0.0419	1.0114	Growing Link
1990	0.0183	0.0391	0.4671	Weak decoupling
1991	0.0515	0.0929	0.5538	Weak decoupling
1992	0.0519	0.1422	0.3651	Weak decoupling
1993	0.0625	0.1387	0.4507	Weak decoupling
1994	0.0581	0.1305	0.4455	Weak decoupling
1995	0.0688	0.1095	0.6280	Weak decoupling
1996	0.0306	0.0993	0.3084	Weak decoupling
1997	0.0053	0.0923	0.0575	Weak decoupling
1998	0.0020	0.0784	0.0258	Weak decoupling
1999	0.0322	0.0767	0.4199	Weak decoupling
2000	0.0455	0.0849	0.5358	Weak decoupling
2001	0.0584	0.0834	0.7003	Weak decoupling
2002	0.0902	0.0913	0.9879	Growing Link
2003	0.1622	0.1004	1.6163	Expansionary negative decoupling
2004	0.1684	0.1011	1.6659	Expansionary negative decoupling
2005	0.1350	0.1140	1.1847	Growing link
2006	0.0960	0.1272	0.7549	Weak decoupling
2007	0.0872	0.1423	0.6126	Weak decoupling
2008	0.0294	0.0965	0.3049	Weak decoupling
2009	0.0484	0.0940	0.5148	Weak decoupling
2010	0.0730	0.1064	0.6859	Weak decoupling
2011	0.0732	0.0954	0.7675	Weak decoupling
2012	0.0390	0.0786	0.4964	Weak decoupling
2013	0.0367	0.0776	0.4736	Weak decoupling
2014	0.0213	0.0730	0.2923	Weak decoupling
2015	0.0099	0.0691	0.1425	Weak decoupling

average annual growth rate of 13.99%. Over the same period actual GDP was similar to that of CO_2 emissions, along with the average annual growth rate of 13.31%. With the entrance into its last period, the growth rate of CO_2 emissions began to decrease, having average annual growth rate of 4.26%. However, actual GDP shows a larger growth rate than in the other periods. In addition, it can be concluded from the scatter that there existed a non-linear relationship between energy-related CO_2 emissions and economic development. And this indicated that adopting the non-linear model proposed in this study for the prediction of CO_2 emissions is necessary and appropriate. Then on the base of these, further study was conducted to explore the specific relationship between CO_2 emissions and economic growth.

Decoupling Analysis

To further study the relationship between energy-related CO_2 emissions and actual GDP, Eq. (2) was utilized to measure decoupling elasticity. And according to the values of elasticity, we divided the annual statistics of China into different levels of decoupling. Table 3 shows the results of the decoupling state between CO_2 emissions and actual GDP. The decoupling trend is depicted in Fig. 3.

Considering that the economy of China does not experience negative growth during the study period, there are only four states of a decoupling relationship between economic growth and CO_2 emissions. Generally speaking, "weak decoupling'" was the main state between CO_2 emissions and actual GDP during the examined period, namely the growth rate of actual GDP was larger than CO_2 emissions. China had witnessed a clear tendency toward decoupling, moving from weak decoupling in 1982-88 with short-term volatility (growing link in 1989) to weak decoupling with short-term volatility (growing link in 2002 and 2005, expansionary negative decoupling in 2003 and 2004), followed by a period of weak decoupling from 2006 to 2015. Therefore, this illustrates that the decoupling trend showed an "M" type in shape.

These coincide with Fig. 3; for instance, the values from 1989 and 2002-05 were higher, and the values from 1981 were upside-down in Fig. 3. There are some possible reasons to explain the special phenomenon. For the year
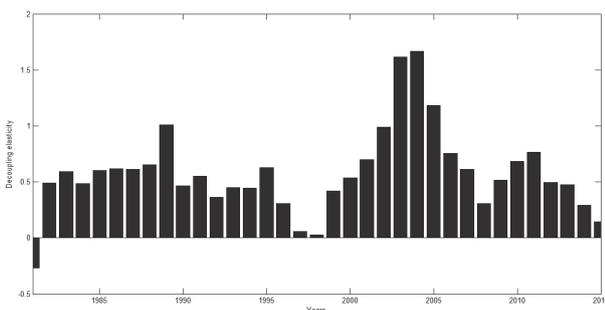
Table 4. Classification of the study years.

Type	Decoupling years	Re-link years
Time	1981-88,1990-2001, 2006-15	2002-05

1981, the coal consumption proportion is smaller than that of 1980, which leads to lower CO_2 emissions. But in 1989, the electricity consumption is larger than that of 1988. For the re-link relationship between CO_2 emissions and actual GDP during the period 2002-05, there was a paramount reason to explain it. Due to the lack of electricity, the power industry, which is the major source of China's man-made CO_2 emissions, received great support from the Chinese government and made significant progress. Limited by resource endowment, China's electricity energy formation takes coal electricity as the primary electricity energy. Thence the development of the power industry caused huge environmental pressure.

Influencing Weight Analysis

According to the results of decoupling analysis, we divided the study years into two types. One type was the years of decoupling, which contained two decoupling states: namely strong decoupling and weak decoupling. The other type was the years with decoupling states of growing link or expansionary negative decoupling. Table 4 illustrates the specific classification.

The influencing factors of CO_2 emissions affect the decoupling state between CO_2 emissions and actual GDP to a different degree. In order to probe into the effects of different driving factors on classifying study years, the ReliefF algorithm was used in further study. To simplify the analysis, "1" denotes the first type (decoupling years) and "2" represents the second type (re-link years). And the repeated sample numbers are set to 100 times. Taking into account that re-link samples are relatively smaller, the near hits is set to 1. Moreover, because the samples are randomly selected during the running time, different random numbers may lead to different results. Therefore, we ran the ReliefF algorithm 20 times and took the average to improve the reliability of the evaluated results.

Fig. 4 reflects the developing trend of influencing weights of different driving factors. It can be seen that there are significant differences among different driving
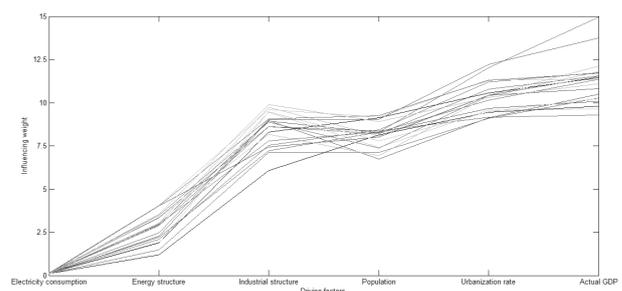

Fig. 3. The decoupling trend of China (1981-2015).

Fig. 4. The influencing weights of different driving factors.

Table 5. Mean values of influencing weight of each driving factor.

Driving factors	Electricity consumption	Energy structure	Population	Industrial structure	Urbanization rate	Actual GDP
Mean weight	0.1286	2.8044	8.1497	8.4748	10.3532	11.2461

factors on influencing weight. Among them, the impact of actual GDP on the decoupling relationship between CO_2 emissions and economic development is the highest, but the impact of electricity consumption is the lowest. In order to gain a complete and intuitive view of the influencing weight, the average values of influencing degree of each factor are shown in Table 5. It is manifest that the mean influencing weights are listed in an effective order as follows: actual GDP, urbanization rate, industrial structure, population, energy structure, and electricity consumption, which is consistent with the content presented in Fig. 4.

Results of CO_2 Forecasting

Fig. 5 depicts the flowchart of forecasting CO_2 emissions based on the grey neural network model, wherein the structure of GNNM is determined by the dimension of input data and output data during the construction of the GNNM model. For the selection of input data, the critical threshold is given to 1. Only when the influencing weight of driving factors is higher in the critical vale can it be considered as the input invariable. Considering that the influencing weight of electricity consumption is far less than 1, it is apparently discarded. Therefore, input data has five dimensions and output data is one dimension. On the basis of comprehensive analysis, a 1-1-6-1 structure for GNNM is set up to predict CO_2 emissions. Namely that the LA layer has one input vector, the LB layer has one input node, the LC layer has six input vectors, and the LD layer has one output vector. Moreover, the input of the LA layer is the time series. The normalization data of actual GDP, urbanization rate, industrial structure, population, and energy structure corresponds to the 2~5 input vectors of the LC layer, respectively. And the output of the proposed model is CO_2 emissions. In this study, the statistics in 36 ranging from 1980 to 2015 are collected to carry out this experiment. First take the first 30 years data as the training set to train the proposed model. Then take iterative 100

times as the end evolution, optimum parameters are outputted. In the end, the remaining data are used to test prediction performance.

To quantitatively measure the performance of the proposed model, four criteria – mean absolute error (MAE), mean absolute percentage error (MAPE), root mean square error (RMSE), and the coefficient of determination (R^2) – are adopted to calculate forecasting accuracy, $R^2 \in [0,1]$, wherein the overall fitting effect of the proposed model can be weighed by RMSE and R^2, where the average prediction capability at each sample point can be examined by MAE. MAPE is one important indicator that is immune to extreme values of data. The formulas of the above-mentioned criteria are:

$$MAE = \frac{1}{n}\sum_{t=1}^{n}|y_t - \widehat{y_t}| \tag{20}$$

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{y_t - \widehat{y_t}}{y_t}\right| \tag{21}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(y_t - \widehat{y_t})^2} \tag{22}$$

$$R^2 = 1 - \frac{\sum_{t=1}^{n}(y_t - \widehat{y_t})^2}{\sum_{t=1}^{n}(\overline{y_t} - \widehat{y_t})^2} \tag{23}$$

…where y_t is the measured value of CO_2 emissions in the i-th year, $\hat{y}_t$ denotes the predicted value of CO_2 emissions, and $\overline{y}_t$ is the mean value of the predicted value.

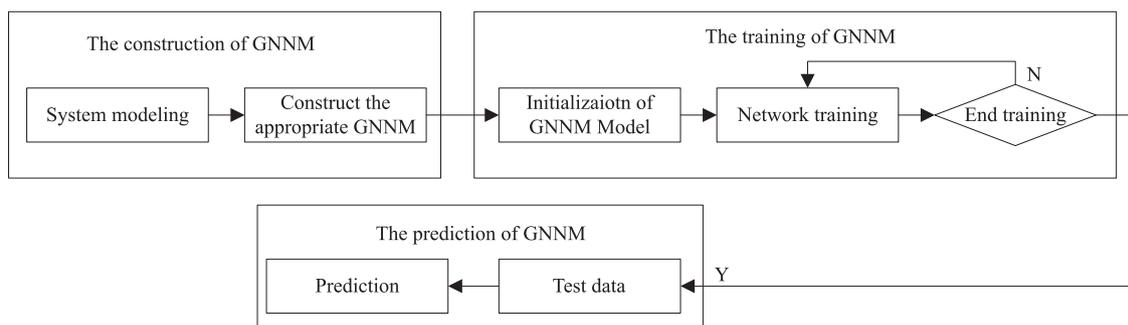To corroborate the outstanding performance of the proposed model, GM (1, 1) model and BPNN model are

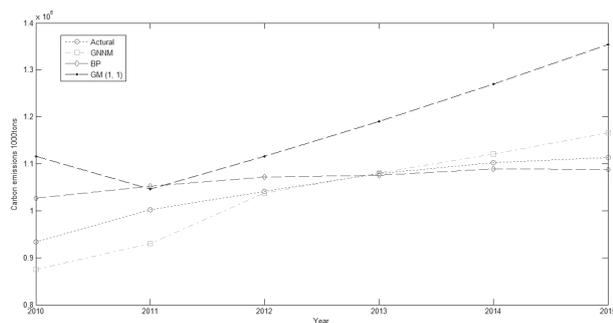Fig. 5. Flowchart of forecasting CO_2 emissions based on GNNM.

Fig. 6. Forecasting results of different models.

Table 6. Error analysis of CO_2 emissions.

Index	MAE (Mkt)	MAPE (%)	RMSE (Mkt)	R^2
GNNM	0.0344	3.39	0.0442	0.9994
GM (1, 1)	0.1365	13.01	0.1519	0.6161
BPNN	0.0361	3.63	0.0463	0.9776

also applied to predict CO_2 emissions for comparison. Fig. 6 shows the forecast results of different models. Comparing with other forecasting models, the GNNM model has a better predictive capability and obtains a more accurate forecasting result. Moreover, observing Table 6, it is apparent that the GNNM model exhibits better forecasting performance according to the assessment results of MAE, MAPE, RMSE, and R^2. By comparing all kinds of index values, it is suggested that the proposed model has an optimum evaluation value, followed by the BPNN model and GM (1, 1) model. There is a paramount reason for this phenomenon, namely that the proposed model combines the important advantages of the GM and BPNN models. On the one hand, the GM model is especially superior for small sample problems. On the other, the BPNN model has better capacity on handling the issues with the characteristics of non-linear and non-stationary. Considering that the data adopted in this paper is less and the CO_2 emissions of China exhibiting non-linear and non-stationary traits, it is appropriate by integrating the GM and BPNN models to forecast CO_2 emissions.

Conclusions

This in-depth paper studies the issue of CO_2 emissions of China using sample data from 1980 to 2015. First, according to the recommended method by IPCC, the amount of CO_2 emissions induced by consumption are estimated during the study period. Then the relationship between CO_2 emissions and economic development is investigated using the Tapio decoupling index. Moreover, depending on the decoupling state between CO_2 emissions and economic development, the study years are divided into decoupling years and re-link years, and the ReliefF algorithm is applied to judge and evaluate the influencing weight of each driving factor of CO_2 emissions. In the end, a novel hybrid model based on the GM and BPNN models is established to forecast CO_2 emissions.

According to the results discussed in this paper, our main conclusions include:

1. CO_2 emissions of China increased from 156,112 kt in 1980 to 1,113,700 kt in 2015, with an average annual growth of 17.04%. The CO_2 emissions curve shows an "S" shape, while the actual GDP of China maintains an upward trend overall.

2. Except for a few years, "weak decoupling" is the main state between CO_2 emissions and economic development during the examined period. The decoupling relationships between CO_2 emissions and economic development exhibit an "M-shaped" pattern.

3. The classification results of study years indicate that the number of decoupling years is far larger than the number of re-link years. The influencing weights of different driving factors of CO_2 emissions decreasing from the top to the bottom in the following order: actual GDP, urbanization rate, industrial structure, population, energy structure, and electricity consumption.

4. Due to the fact that the influencing weight of electricity consumption is lower than the critical threshold of 1, it is apparently aborted and the remaining driving factors are taken as the input variables of the GNNM model. Compared with a single GM model and single BPNN model, the hybrid model produces the lowest MAE, MAPE, and RMSE, and the highest R^2. That is to say, the GNNM model has the better capacity for forecasting CO_2 emissions and capturing the non-linear and non-stationary characteristics of CO_2 emissions.

References

1. DAI P., ZOU J., TIAN J., LIU T., ZHOU H. Integrated optimization of CO_2 emission mitigation in China power sector, Automation of Electric Power Systems, **37** (14), 1, **2013**.

2. BIROL F. World Energy Outlook 2013, International Energy Agency. Paris, **2013**.

3. YE B., JIANG J., MIAO L., YANG P., LI J., SHEN B. Feasibility study of a solar-powered electric vehicle charging station model, Energies, **8** (11), 13265, **2015**.

4. PANAYOTOU P. Empirical tests and policy analysis of environmental degradation at different stages of economic development, llo Working papers, 4, **1993**.

5. SCHMALENSEE R., Stoker T.M., Judson R.A. Word carbon dioxide emission: 1950-2050, Review of Economics & Statistics, **80**, 15, **1998**.

6. AROURI M.E.H., Youssef A.B., HENNI H.M., RAULT C. Energy consumption, economic growth and CO_2 emission in Middle East and North African countries, Energy Policy, **45**, 342, **2012**.

7. HAMIT-HAGGAR M. Greenhouse gas emission, energy consumption and economic growth: a panel co-integration analysis from Canadian industrial sector perspective, Energy Economics, **34**, 358, **2012**.

8. SABOORI B., SULAIMAN J., MOHD S. Economic growth and CO_2 emissions in Malaysia: a co-integration analysis of the environment Kuznets curve, Energy Policy, **51**,184, **2012**.

9. GALEOTTI M., LANZA A., PAULI F. Reassessing the environmental Kuznets curve for CO_2 emissions: a robustness exercise, Ecological Economics, **57** (1), 152, **2006**.

10. OECD. Sustainable development: Indicators to measure decoupling of environmental pressure from economic growth, Paris, **2002**.

11. TAOIO P. Towards a theory of decoupling: Degrees of decoupling in the EU and the case of road traffic in Finland between 1970 and 2001, Transport Policy, **12**, 137, **2005**.

12. WAN L., WANG Z.L., NG J. Measurement research on the decoupling effect of industries' carbon emissions – based on the equipment manufacturing industry in China, Energies, **9**, 21, **2016**.

13. DENG M.X., LI W., HU Y. Decomposing industrial energy-related CO_2 emissions in Yunnan province, China: Switching to low-carbon economic growth, Energies, **9**, 23, **2016**.

14. FREITAS L.C.D., KANEKO S. Decomposing the decoupling of CO_2 emissions and economic growth in Brazil, Ecological Economics, **70**, 1459, **2011**.

15. HYLAND M. Decomposing patterns of emission intensity in the EU and China: how much does trade matter, Journal of Environmental Planning and Management, **58**, 2176, **2015**.

16. ZHAO Y., ZHANG Z., WANG S., ZHANG Y., LIU Y. Linkage analysis of sectoral CO_2 emissions based on the hypothetical extraction method in South Africa, Journal of Cleaner Production, **103**, 916, **2015**.

17. BORTOLINI M., FACCIO M., FERRARI E., GAMBERI M., PILATIF. Fresh food sustainable distribution: cost, delivery time and carbon footprint three-objective optimization, Journal of Food Engineering, **174** (85), 56, **2016**.

18. HORI K., MATSUI T., HASUIKE T., FUKUI K., MACHIMURA T. Development and application of the renewable energy regional optimization utility tool for environmental sustainability, Renewable Energy, **93**, 548, **2016**.

19. YORUCU V. Growth impact of CO_2 emissions caused by tourist arrivals in Turkey: an econometric approach, International Journal of Climate Change Strategies and Management, **8** (1), 19, **2016**.

20. XU B., LIN B. What cause a surge in China's CO_2 emissions? A dynamic vector auto-regression analysis, Journal of Cleaner Production, **143**, 17, **2016**.

21. VOL N. Input-output economics, Oxford University Press, **185**, 15, **1986**.

22. CHEN L., YANG Z.F., CHEN B. Decomposition analysis of energy-related industrial CO_2 emissions in China, Energies, **6**, 2319, **2013**.

23. WANG Z.X., YE D.J. Forecasting Chinese carbon emissions from fossil energy consumption using non-linear grey multivariable models, Journal of Cleaner Production, **142**, 600, **2017**.

24. QIANG D.U., CHEN Q., YANG R. Forecast carbon emissions of provinces in China based on logistic model, Resources & Environment in the Yangtze Basin, **22** (2), 143, **2013**..

25. DU Q., WANG N., CHE L. Forecasting China's per capita carbon emissions under a new three-step economic development strategy, Journal of Resources & Ecology, **6** (5), 318, **2012**.

26. ZHANG Y., WANG C., WANG K., CHEN J. CO_2 emission scenario analysis for China's electricity sector based on LEAP software, Journal of Tsinghua University, **47** (3), 365, **2007**.

27. WEI SUN, MOHAN LIU Prediction and analysis of the three major industries and residential consumption CO_2 emissions based on least squares support vector machine in China, Journal of Cleaner Production, **122**, 144, **2016**.

28. HAN Y.M., ZHU Q.X., GENG Z.Q., XU Y. Energy and carbon emissions analysis and prediction of complex petrochemical systems based on an improved extreme learning machine integrated interpretative structural model, Applied Thermal Engineering, **115**, 280, **2017**.

29. LI J.Y., SHI J.F., LI J.C. Exploring reduction potential of carbon intensity based on back propagation neural network and scenario analysis: a case of Beijing, China, Energies **9**, 615, **2016**.

30. IEA. IPCC guidelines for national greenhouse gas inventories, Energy, 2, **2006**.

31. DENG J.L. Introduction to grey system theory, Journal of Grey System, **1** (1), 1, **1989**.