A novel method for carbon emission forecasting based on EKC hypothesis and nonlinear multivariate grey model: evidence from transportation sector

Siyuan Huang1 · Xinping Xiao1 · Huan Guo2

Received: 26 October 2021 / Accepted: 3 April 2022 / Published online: 15 April 2022
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

Abstract
Greenhouse gas emissions have brought a serious challenge to the global environment and climate. Efficient and accurate prediction of carbon emissions is essential for the decision-making sectors to control growth and formulate policies. Firstly, considering the economic, demographic, and energy factors, a novel nonlinear multivariate grey model (ENGM(1,4)) based on environmental Kuznets curve (EKC) is proposed with respect to the data characteristics of the incomplete information of carbon emission of transportation sector. The model integrates the IPAT (“Influence = Population, Affluence, Technology”) equation and the extended stochastic impacts by regression on population, affluence, and technology model (STIRPAT). Secondly, the derivation method is used to solve the time response equation of the model and the quantum particle swarm optimization algorithm (QPSO) is designed to optimize the model parameters. Then, 18 years of carbon emission data from China, the USA, and Japan are selected as the validation set. Comparative analysis indicates that the prediction accuracy of the statistical models and the intelligent models depends on sufficient samples and complex variables, and has certain limitations in limited sample prediction. The calculation results show that the new model outperforms other models in various evaluation indicators, indicating that its prediction accuracy is higher. Finally, the projections show that in 2019–2025, the average increase in carbon emissions from the transport sector in China and the USA was 2.837% and 2.394%, respectively, while Japan shows a downward trend with an average decline rate of 1.2231%. The analyzed prediction results are consistent with current situation of the three countries and the transport sectors, demonstrating the high accuracy and reliability of the new model.

Keywords Carbon emission forecasting · Transportation sector · EKC · Grey model · ENGM(1,4) model

Introduction
This section is divided into two parts: the “Background and significance of the study” section explains the background and importance of this study in the light of the current international environmental situation and carbon emissions data. The “Motivation” section analyzes the current approaches to carbon emissions research and summarizes the advantages and disadvantages of each type of approach, explaining the motivation for this paper.

Background and significance of the study
Carbon dioxide emissions are a major contributor to the global greenhouse effect. There are studies showing that two-thirds of greenhouse gases come from the combustion of fossil fuels (Wu et al. 2015). Excessive carbon emissions will have a huge negative impact on the global ecological environment, economic development, and human life (Zhou et al. 2021; Rao and Yan 2020). The need to reduce energy production and carbon emissions in various industries has attracted growing attention from an increasing number of countries around the world (Saidi and Omri 2020). With the continuous acceleration of industrialization and urbanization, the demand for convenient transportation and freight
transportation in modern society continues to increase. Therefore, the energy consumption of the transportation sector is also increasing, becoming one of the industries with high carbon emissions (Li et al. 2021; Wang et al. 2022). Carbon emissions from the transportation sector are one of the major contributors to greenhouse gases. Data from the International Energy Agency (IEA) indicate that carbon emissions from the transportation sector have accounted for nearly 30% of the total carbon emissions (IEA 2021), and this proportion will gradually increase. In recent years, carbon emissions from the transportation has become a major concern for countries and related sectors in recent years (Tian et al. 2018), and energy conservation and emission reduction in the transportation sector will become the top priority of low-carbon city construction (Agbulut 2022). Therefore, it is crucial to analyze the factors influencing carbon emissions and adopt a scientific method to predict future carbon emissions efficiently and accurately. It can provide an important reference for the formulation of government energy saving and emission reduction measures, and the accuracy of the prediction results is of great significance to the concretization and refinement of government environmental protection policies (Sevgül and Agbulut 2021). Based on the forecast results, the government and relevant decision-making departments can anticipate and grasp the development trend of carbon emissions in advance, so as to reasonably optimize the transport structure to reduce energy consumption and formulate targeted carbon emission strategies to achieve sustainable development, which plays a key role in achieving the goal of “carbon neutrality” and “carbon peaking” as early as possible.

Motivation

Current carbon emission research methods can be roughly divided into three categories: statistical analysis models, intelligent prediction models, and grey prediction models, which mostly forecast carbon emissions according to time series. The prediction effect and accuracy of the first two types of models usually depend on sufficient samples and complex variables, and they have the disadvantage of low robustness. Firstly, the current international energy structure is changing rapidly, the reliability of short-term carbon emission forecasting with long data spans is low, and the phased carbon emission data are mostly limited samples with dispersion and uncertainty. The grey model is proven to have high accuracy and validity in such data forecasting. Secondly, in view of the characteristic that carbon emissions are also influenced by economic, demographic, and energy use, the EKC hypothesis is widely used in studying the influencing factors of carbon emissions (Bilgili et al. 2016; Bibi and Jamil 2021; Bilgili et al. 2022a). In summary, considering the characteristics of carbon emission data and many influencing factors, statistical analysis models usually cannot adapt to the complex and changeable research background due to their simple structure, and the prediction accuracy is not stable enough due to the influence of the sample size. Although nonlinear intelligent prediction models have high prediction performance, they often have defects such as over-fitting, falling into local optima, and poor generalization ability, and show certain limitations in limited sample prediction based on the research background of this paper (Xiao et al. 2021; Gao et al. 2022). Therefore, apart from the characteristics of carbon emission data, various influencing factors must be considered. A short-term prediction model based on the EKC hypothesis, which is consistent with the intrinsic mechanism of carbon emission change, should be established. The model parameters should be optimized by using intelligent algorithms so as to improve the accuracy of the model, which is also the focus of this study.

Next, this paper reviews the common methods of carbon emission prediction and related research status.

Literature review

In recent years, due to the pressure of global energy conservation, emission reduction, and environmental protection, countries have strengthened their research on carbon emissions. In this section, by reviewing relevant literature and combining the research background of this paper, the main research methods are determined: Based on the EKC assumption, a multivariate grey model is established to study the relationship between carbon emissions and other influencing factors and to predict carbon emissions. This section is divided into four parts: the “Research progress on carbon emission forecasting” section introduces the relevant methods and research progress of carbon emission prediction in recent years. The “Grey prediction model” section mainly introduces the application of grey prediction model in carbon emission. The “Theories related to the factors influencing carbon emissions” section introduces the theoretical research progress of carbon emission influencing factors. The “Contribution and organization” section introduces the research contribution of this paper and the organization of the whole text.

Research progress on carbon emission forecasting

The three types of methods commonly used for carbon emission prediction are summarized in Table 1, namely statistical analysis models, intelligent forecasting models, and grey forecasting models. In aspects of statistical analysis models, Yin et al. (2017) first calculated the carbon dioxide content in flue gas using fuel characteristic coefficients and then predicted carbon emissions in Guangdong Province using
a multiple linear regression model. Fang et al. (2018) proposed an improved Gaussian regression method based on the hyper parameters of covariance function in the Gaussian processes regression for predicting carbon emissions in China, the USA, and Japan. Sutthichaimethee and Ariyasajjakorn (2018) developed the integrated moving average with explanatory variables (ARIMAX) model for carbon emission projections in the industrial sector in Thailand. Hosseini et al. (2019) used time series-based multiple linear regression (MLR) and polynomial regression (MPR) to predict Iran’s carbon emissions from 2017 to 2030. The biggest shortcoming of the statistical model is that it requires sufficient sample input to accurately estimate the model parameters and requires a high level of raw data (Pao et al. 2012). In aspects of intelligent predictive models, Sun et al. (2017) used the particle swarm optimization (PSO) algorithm to optimize the extreme learning machine and used it to predict the carbon emissions of Hebei Province, China. Wen and Cao (2020) optimized the parameters in a support vector machine model through an improved chicken swarm optimizer and genetic algorithm to predict China’s commercial sector’s CO2 emissions. The disadvantage is that these two types of models require large sample input in practical applications. Moreover, the prediction results depend on complex parameters and variables, and the robustness is insufficient. The next section will focus on the

| Author | Model type | Model | Application |
|--------|------------|-------|-------------|
| Yin et al. (2017) | Statistical analysis model | Multiple linear regression analysis | China’s CO₂ emissions |
| Fang et al. (2018) | Statistical analysis model | Improved Gaussian processes regression method | China, the USA, and Japan’s CO₂ emissions |
| Sutthichaimethee and Ariyasajjakorn (2018) | Statistical analysis model | ARIMAX model | Industrial CO₂ emissions in Thailand |
| Hosseini et al. (2019) | Statistical analysis model | Time series model based on MLP and MPR | Iran’s CO₂ emissions |
| Sun et al. (2017) | Nonlinear intelligent model | PSO-ELM model | CO₂ emissions in Hubei, China |
| Wen and Cao (2020) | Nonlinear intelligent model | SVM based on ICSO algorithm | Residential energy related CO₂ emissions |
| Qiao et al. (2020) | Nonlinear intelligent model | Improved lion swarm optimizer | CO₂ emissions from 14 countries |
| Leerbeck et al. (2020) | Nonlinear intelligent model | A novel machine learning | CO₂ emission intensity in power grids |
| Wen and Yuan (2020) | Nonlinear intelligent model | Neural network based on random forest and PSO | CO₂ emissions in China’s commercial department |
| Ren and Long (2021) | Nonlinear intelligent model | FLN and CSO algorithm | CO₂ emissions of Guangdong, China |
| Ding et al. (2020) | Grey forecasting method | A novel DGPM(1, N) model | Chinese energy-related CO₂ emissions |
| Zhou et al. (2021) | Grey forecasting method | A novel grey rolling GRPM(1,1) model | China’s CO₂ emissions |
| Gao et al. (2021) | Grey forecasting method | FAGGM(1,1) model | American industrial sector’s CO₂ emissions |
| Guo et al. (2021) | Grey forecasting method | Exponential cumulative grey model | BRICS countries’ CO₂ emissions |
| Ofosu-Adarkwa et al. (2020) | Grey forecasting method | Hybrid Verhulst-GM(1,N) model | CO₂ emissions of China’s cement industry |
| Xie et al. (2021) | Grey forecasting method | RWGM(1,N) model based on LASSO regression | CO₂ emissions in European |
| Duan et al. (2020) | Grey forecasting method | GMC(1,N) model | CO₂ emissions of Chongqing, China |
| Ye et al. (2021) | Grey forecasting method | A time-delay multivariate ATGM (1, N) model | CO₂ emissions from China’s transportation sectors |
| Cao et al. (2021) | Grey forecasting method | TIGM(1, N) model | CO₂ emissions from Chinese marine fleets |
characteristics and applications of grey prediction models in carbon emissions.  

**Grey prediction model**  
At present, the international ecological environment is changing rapidly. The significant changes that have occurred and are occurring in the global energy structure, which make carbon emissions show distinct development trends at different stages. On the one hand, this characteristic makes the prediction based on historical data spanning too long a time less reliable (Wang et al. 2020). On the other hand, in practical applications, carbon emission data series are usually in years, possessing discreteness and uncertainty. Therefore, a grey model that shows high accuracy and effectiveness for incomplete information and sparse sample prediction is considered (Deng 1982; Xie and Liu 2009). A dataset with few samples is chosen to train the model and then perform short-term prediction. In terms of univariate grey prediction models, Ding et al. (2020) designed a new discrete grey prediction model (DGPM(1,N)) to predict China’s carbon emissions from 2016 to 2020. Zhou et al. (2021) proposed a novel rolling grey prediction model (GRP(1,1)) to dynamically predict carbon emissions in China. Gao et al. (2021) developed a new fractional order cumulative grey model (FAGGM(1,1)) based on the Gompertz differential equation for the study of carbon emissions in the USA. Guo et al. (2021) constructed an exponential-order cumulative grey model (ECGM(1,1)) to predict the carbon emissions of BRICS (Brazil, Russia, India, China, South Africa) countries, and adjusted the weight of information by adjusting the size of the parameters. In terms of multivariate grey prediction models, Ofosu-Adarkwa et al. (2020) used a hybrid Verhulst-GM(1,N) model to predict carbon emissions in China’s cement industry. Xie et al. (2021) used a novel robust reweighted multivariate grey model (RWGM(1,N)) to study the European carbon emissions forecast. Duan et al. (2020) developed a new multi-kernel nonlinear multivariate grey model with convolution integral (GMC(1,N)) based on a Gaussian vector basis kernel function and a global polynomial kernel function to predict Chongqing’s carbon emissions from 2016 to 2020. Ye et al. (2021) developed an accumulative time-delay multivariate grey model (ATGM(1,N)) for predicting carbon emissions in the transportation sector in China and demonstrated that the model predicts better than other comparative models. Cao et al. (2021) developed a multivariate trend interaction grey model (TIGM(1,N)) for carbon emission prediction, taking into account the inherent trends of carbon emissions in the context of the marine fleet and the interaction between its influencing factors. The above grey model optimizes and improves the traditional grey model from different perspectives, which is well adapted to the carbon emission data characteristics. In addition to the choice of an appropriate prediction model, adequate consideration of the factors influencing carbon emissions also occupies an important place in carbon emissions research. Next, the theoretical research progress on the influencing factors of carbon emissions will be introduced.  

**Theories related to the factors influencing carbon emissions**  
With the global energy restructuring in recent years, factors such as economic development, population migration, and energy use have influenced carbon emissions. In practical applications, apart from considering characteristics of carbon emission data, the influence of various complex factors should be integrated. A prediction model that conforms to the inherent development law of carbon emissions should be established. In the study of the factors influencing carbon emissions, economist Kuznets proposed the environmental Kuznets curve (EKC) in 1955 that can be used to reflect the relationship between environmental and economic variables (Kuznets 1995). Grossman and Krueger proposed the EKC theory in 1991 (Grossman and Krueger 1991), which specifically explores the relationship between carbon emissions and GDP. They validated this theory in 1995 through different environmental indicators (Grossman and Krueger 1995). The EKC curve reflects the impact of GDP on carbon emissions to a certain extent by revealing the inverted “U”-shaped relationship between economic development and environmental change (Bilgili et al. 2021b, a, 2022a, b). The EKC-related theory is then widely used in carbon emission studies. Dong et al. (2018) used the EKC theory to explore the dynamic causal relationship between per capita carbon emissions, per capita GDP, and per capita fossil fuel consumption in China. Brown et al. (2020) studied the relationship between carbon emissions and investor remittances in Jamaica using a modified asymmetric EKC. Bilgili et al. (2021b) established the EKC model through quantile simulation, and studied the relationship between environmental pollution and energy research and development in 13 developed countries. There are many other studies of this type (Aslam et al. 2021; Kharbach and Chfadi 2017; Altıntaş and Kassouri 2020; Ali et al. 2017). Based on the related research on the EKC, American ecologists Ehrlich and Holdren proposed the IPAT equation in the 1970s (Ehrlich and Holdren 1970), which is widely used to study the relationship between the environment (carbon emissions) and the economy, population, and technology (Chontanawat 2018; Wang et al. 2011). However, the single-factor variation characteristic of IPAT cannot adapt to the complex social environment (Lin et al. 2017). To overcome this shortcoming, York et al. (2003) improved it to obtain the extended STIRPAT model, which converted the original IPAT equation into exponential form and enhanced the flexibility of the model. Subsequently, many scholars
used the STIRPAT model in carbon emission studies (Wang et al. 2017; Liu and Xiao 2018; Zhang and Zhao 2019). The above model studies the influence of economic, energy, and population factors on carbon emissions. The above models analyze the implications of factors such as economy, energy, and population on carbon emissions, and explain the internal mechanism of carbon emissions changes. These models can make predictions under the premise of comprehensively considering various influencing factors, with high reliability and prediction accuracy.

In this study, starting from the characteristics of carbon emission data and internal development law and considering that carbon emission is influenced by various factors such as economy, population, and energy use, the population factor in the IPAT equation and the extended STIRPAT model is introduced. A new nonlinear multivariate grey model (ENGM(1,N)) based on the assumption of EKC hypothesis is constructed. The new model is a coupling of multiple methods with multiple features that can predict carbon emissions with high accuracy and flexibility. It considers complex influencing factors, which provides new ideas and methods for short-term prediction of carbon emissions. Definitions of abbreviations appearing in the text are listed in Table 2.

### Contribution and organization

The main contributions of this study are outlined below.

1. A new type of multivariate grey prediction model is established on the basis of the EKC theory, combined with the IPAT equation and the extended STIRPAT model. The model considers the impact of economic development, population, and energy use on carbon emissions.

2. The time response equation of the new model is obtained by the derivation method. The parameter values are obtained by the least squares method. The quantum particle swarm algorithm is designed to optimize the model parameters to improve the accuracy of the model.

3. Data on carbon emissions from the transportation sector in China, the USA, and Japan are selected for model validation. The data from these three countries have increasing, fluctuating, and decreasing characteristics, so the adaptability of the new model to different types of data is depicted. The new model is applied to the prediction of carbon emissions in these three countries for the next 7 years.

The remainder of this paper is organized as follows. The structure diagram of the full text is shown in Fig. 1. The

### Table 2 Definition of abbreviations

| Abbreviation | Description |
|--------------|-------------|
| EKC          | Environmental Kuznets curve |
| IPAT         | Influence = population, affluence, technology |
| STIRPAT      | Stochastic impacts by regression on population, affluence, and technology model |
| IEA          | International Energy Agency |
| IBRD         | The World Bank |
| BRICS        | Brazil, Russia, India, China, South Africa |
| R&D          | Research and development |
| ENGM(1,N)    | Nonlinear multivariate grey model based on EKC |
| DGPM(1,N)    | Multivariate discrete grey prediction model |
| GRPM(1,1)    | Rolling grey prediction model |
| FAGGM(1,1)   | Fractional order cumulative grey model |
| ECGM(1,1)    | Exponential-order cumulative grey model |
| GM(1,1)      | Grey model |
| GMC(1,N)     | Multivariate grey model with convolution integral |
| RWGM(1,N)    | Robust reweighted multivariate grey model |
| ATGM(1,N)    | Accumulative time-delay multivariate grey model |
| TIGM(1,N)    | Multivariate trend interaction grey model |
| ARIMAX       | Integrated moving average with explanatory variables |
| ARIMA        | Autoregressive integrated moving average model |
| MLR          | Multiple linear regression |
| MPR          | Multiple polynomial regression |
| PSO          | Particle swarm optimization |
| QPSO         | Quantum particle swarm optimization |
| CSO          | Chicken swarm optimization |
| ICSO         | Improved chicken swarm optimization |
| FLN          | Fast learning network |
| SVR          | Support vector regression |
| ANN          | Artificial neural network |

“Model building” section introduces the process of model building. The “Model solving” section describes the model solving process, which mainly includes estimating parameters using least squares method, solving the whitening differential equation using the derivative method, and optimizing the model parameters using the quantum particle swarm algorithm. The “Model verification” section selects carbon emission data from the transportation sector in three countries, China, the USA, and Japan, as a validation set to verify the validity of the model. The “Application” section predicts the carbon emissions of the transportation sector in the three countries of China, the USA, and Japan from 2019 to 2025, and analyzes the prediction results. The main conclusions of this study are summarized in the “Conclusion” section.
Model building

This section is divided into two parts: the “Carbon emission prediction model for transportation sector” section introduces the process of establishing the differential equation for carbon emission prediction based on the EKC assumption. The “Theoretical modeling of the ENGM(1,N)” section derives the ENGM(1,4) model through the differential information principle.

Carbon emission prediction model for transportation sector

The EKC proposed by economist Kuznets in 1955 revealed an “inverted U-shaped” relationship between economic development and environmental damage (Ali et al. 2017). The EKC hypothesis proposed by Grossman and Krueger (1991) states that there is a certain degree of linear and nonlinear correlation between carbon emissions and economic factors in carbon emission studies; the expression of which can be found in Eq. (1):

\[ CO_{2,t} = f(Y, Y^2, RE, FO) \]  

where \( Y^2 \) and \( Y \) denote economic development; \( RE \) and \( FO \) denote renewable energy and fossil fuel, respectively, representing primary energy and secondary energy, which can be replaced by total energy consumption (\( EC \)) (Le and Van 2020). Thus, Eq. (2) is obtained. The EKC hypothesis indicates that there is a relationship between carbon emissions, the economy, and energy sources.

\[ CO_{2,t} = f(Y^2, Y, EC) \]  

Based on related research on the EKC, Kharbach and Chfadi (2017) tested the significant differences between \( GDP \), \( GDP^2 \), and road diesel consumption (\( RSD \)) and road traffic carbon emissions (\( RCO_{2,t} \)). They verified the differences to get Eq. (3):

\[ RCO_{2,t} = a_0 + a_1 GDP_t^2 + a_2 GDP_t + a_3 RSD_t + \mu_t \]  

Further extending Eq. (3) to the carbon emissions of the transportation sector (\( TCO_{2,t} \)), the Eq. (4) can be obtained:
\[ TCO_{2g} = \alpha_0 + \alpha_1 GDP^2 + \alpha_2 GDP_t + \alpha_3 EC_t + \mu_t \tag{4} \]

In Eq. (4), \( TCO_{2g} \) represents the carbon emissions of the transportation sector, \( EC_t \) represents energy consumption, and the constant term \( \mu_t \) is used to reduce the error caused by random disturbance factors.

On the basis of the EKC and in addition to considering economic and energy factors, Ehrlich and Holdren (1970) further introduced demographic factors to establish a constant expression for the IPAT equation, as shown in Eq. (5) (Dogan and Inglesi-Lotz 2020):

\[ I = P \times A \times T \tag{5} \]

To enhance the flexibility of the model based on the IPAT equation, York et al. (2003) introduced exponents to establish an extended STIRPAT model, as shown in Eq. (6):

\[ I_t = aP_{A}^{\beta}A_{t}^{\alpha}T_{t}^{\gamma}e_{i} \tag{6} \]

where \( I \) can be interpreted as the rate of environmental change (in carbon emission studies as the level of carbon emissions), \( P \) is the population size, \( A \) is regarded as the degree of affluence, and \( T \) is the environmental damage caused by production or technology (usually expressed in carbon emission studies as the consumption of nuclear energy representing the technology) (Liu and Xiao 2018).

The IPAT equation and the extended STIRPAT model consider the effect of demographic factors on carbon emissions. The significant effect of demographic factors on carbon emissions has been tested in recent years in carbon emission-related studies (Chontanawat 2018). Therefore, the population factor \( P_t \) is introduced on the basis of Eq. (4) to obtain the relationship between transportation sector carbon emissions and economy, energy, and population, as given in Eq. (7):

\[ TCO_{2g} = \alpha_0 + \alpha_1 GDP^2 + \alpha_2 GDP_t + \alpha_3 EC_t + \alpha_4 P_t + \mu_t \tag{7} \]

Assuming that the carbon emissions of the transportation sector at time \( t \) is \( X_{1}^{(t)}(t) \), then the total carbon emissions in the time period \([1, t]\) can be expressed as follows:

\[ X_{1}^{(t)}(t) = \int_{0}^{t} X_{1}^{(u)}(u) du \quad \text{or} \quad \frac{dX_{1}^{(t)}(t)}{dt} = X_{1}^{(0)}(u) dX_{1}^{(t)}(u) = \frac{dX_{1}^{(t)}(t)}{dt} \tag{8} \]

Combining Eq. (7) and Eq. (8), we can obtain:

\[ \frac{dX_{1}^{(t)}(t)}{dt} = X_{1}^{(0)}(u) = TCO_{2g} = \alpha_0 + \alpha_1 GDP^2 + \alpha_2 GDP_t + \alpha_3 EC_t + \alpha_4 P_t + \mu_t \tag{9} \]

When the research stage is short, the national energy structure changes are ignored. The increment of carbon emissions \( \Delta X_{1}^{(t)}(t) \) and the total energy consumption \( TEC_t \) can be considered as an approximate linear relationship:

\[ \Delta X_{1}^{(t)}(t) = X_{1}^{(t)}(t) - X_{1}^{(t)}(1) = \gamma EC_t \tag{10} \]

then:

\[ EC_t = \frac{1}{\gamma} X_{1}^{(t)}(t) - \frac{1}{\gamma} X_{1}^{(t)}(1) \tag{11} \]

Data from IEA shows that the total energy consumption in the transport sector accounts for approximately 29% of the total overall energy consumption (IEA 2021). According to Eqs. (10) and (11), the total energy consumption of the transportation sector \( (TEC) \) can be used to represent the total energy consumption \( (EC) \):

\[ EC_t = (1 - p)EC_t + TEC_t = (1 - p) \left( \frac{1}{\gamma} X_{1}^{(t)}(t) - \frac{1}{\gamma} X_{1}^{(t)}(1) \right) + TEC_t = \gamma_0 + \gamma_1 X_{1}^{(t)}(1) + TEC_t \tag{12} \]

where \( p \) is the proportion of total energy consumption in the transportation sector to total energy consumption, \( \gamma_0 = (p - 1)\frac{1}{\gamma} X_{1}^{(1)}(1) \), \( \gamma_1 = \frac{1-p}{\gamma} \). Bringing Eq. (12) into Eq. (9), we get:

\[ \frac{dX_{1}^{(t)}(t)}{dt} = \alpha_0 + \alpha_1 GDP^2 + \alpha_2 GDP_t + \alpha_3 \left( \gamma_0 + \gamma_1 X_{1}^{(t)}(1) + TEC_t \right) + \alpha_4 P_t + \mu_t \tag{13} \]

Equation (13) reveals the relationship between the growth rate of carbon emissions in the transportation sector and GDP, energy consumption in the transportation sector, and population. \( X_2, X_3, \) and \( X_4 \) are used to denote GDP, transportation sector energy consumption, and population values over a period of time. Considering that economic development is influenced by various internal and external factors and has the characteristics of instability, the index of \( GDP^2 \) variable is extended from 2 to \( r \) to enhance the model flexibility. The optimal \( r \) value can be obtained by an optimization algorithm to improve the prediction accuracy of the model. Let \( a = -a_2 \gamma_1, b = \alpha_1, c_i = \alpha_i (i = 2, 3, 4), d = a_0 + \mu_t \), and simplify Eq. (13) to obtain:

\[ \frac{dX_{1}^{(t)}(t)}{dt} + aX_{1}^{(t)}(t) = b X_{2}^{(1)}(t) + \sum_{i=2}^{N} c_i X_{1}^{(1)}(t) + d, N = 4 \tag{14} \]
That is, Eq. (14) is a predictive differential equation about the cumulative carbon emissions $X_{1}^{(1)}(t)$ of the transportation sector within $[1, t]$.

**Theoretical modeling of the ENGM(1, N)**

Equation (14) is a first-order nonlinear differential equation containing multiple variables, whereas $X_{1}^{(1)}(t)$ is a first-order continuous derivable and contains the infinite information. In carbon emissions research, carbon emissions are usually measured in years, which is a discrete time series with limited uncertain information or so-called grey information. According to the differential information principle (Deng 1982), Eq. (14) is discretized and transformed to obtain a new nonlinear multivariate grey model based on the EKC assumption.

Suppose that the data series of annual carbon emissions from transportation sector is $X_{1}^{(1)} = (X_{1}^{(0)}(1), X_{1}^{(0)}(2), \ldots, X_{1}^{(0)}(t))$. Then, according to Eq. (8), $X_{1}^{(1)}(t)$ is the total carbon emission during the period of $[1, t]$, which can be expressed as follows:

$$X_{1}^{(1)}(t) = \sum_{k=1}^{t} X_{1}^{(0)}(t), t = 1, 2, \ldots, n \tag{15}$$

As for the cumulative carbon emissions $X_{1}^{(1)}(t)$, the rate of its changing during the time period $[k-1, k]$ can be approximated as follows (Deng 1982):

$$\frac{dX_{1}^{(1)}(t)}{dt} \mid_{(k-1)} \approx \frac{X_{1}^{(1)}(k) - X_{1}^{(1)}(k-1)}{k - (k-1)} = X_{1}^{(0)}(k) \tag{16}$$

It was worth noticing that $X_{1}^{(0)}(k)$ is usually be regarded as the grey derivative and can provide the necessary information on time-dependent function $X_{1}^{(1)}(k)$. Namely, $X_{1}^{(0)}(k)$ can be regarded as the grey derivative function of $X_{1}^{(1)}(k)$ and take the place of $\frac{dX_{1}^{(1)}(t)}{dt}$ during $[k-1, k]$. Besides, the background value of the grey derivative $z_{1}^{(1)}(k)$ can be formalized as follows:

$$X_{1}^{(1)}(t) \mid_{[k-1, k]} \approx pX_{1}^{(1)}(k) + (1 - p)X_{1}^{(1)}(k) = z_{1}^{(1)}(k) \tag{17}$$

where $p \in [0, 1]$. To simplify $z_{1}^{(1)}(k)$, $p$ is equal to 0.5 in general. In grey modeling, the background value $z_{1}^{(1)}(k)$ can usually take the place of the value of $X_{1}^{(1)}(k)$ during the period $[k-1, k], i = 1, 2, \ldots, n$. Hence, the following grey model can be defined.

**Definition 1.** Equation (18) is named as the ENGM(1,4) model. The whitenization differential equation of the ENGM(1,4) model can be expressed as Eq. (19):

$$X_{1}^{(0)}(k) + az_{1}^{(1)}(k) = b \left( z_{2}^{(1)}(k) \right) + \sum_{i=2}^{N} c_{i}z_{i}^{(1)}(k) + d, N = 4 \tag{18}$$

$$\frac{dX_{1}^{(1)}(t)}{dt} + aX_{1}^{(1)}(t) = b \left( X_{2}^{(1)}(t) \right) + \sum_{i=2}^{N} c_{i}X_{i}^{(1)}(t) + d, N = 4 \tag{19}$$

In Eq. (18), $a$ is the development coefficient that can reflect the development trend of $X_{1}^{(0)}$ and $X_{1}^{(1)}$, $b$ and $c_{i}$ are the control coefficients, $i = 2, 3, 4$, $d$ is a constant term, and $r$ is the power exponent.

**Model solving**

This section is divided into three parts: The “Parameter estimation based on least squares method” section introduces the process of parameter estimation using the least squares method. The solution of the whitenization differential equation by the derivative method is discussed in the “Derivative method for solving whitlenization differential equation” section. The quantum particle swarm algorithm is designed in the “Parameter optimization” section to optimize the exponent $r$ thus improving the model accuracy.

**Parameter estimation based on least squares method**

The parameters $a$, $b$, $c_{2}$, $c_{3}$, $c_{4}$, $\ldots$, $c_{N}$, $d$ and the power exponent $r$ in the model affect the prediction effect and accuracy of the model. To facilitate the calculation, the power exponent is first considered as a known parameter and the expression containing the parameter $r$ for $a$, $b$, $c_{2}$, $c_{3}$, $c_{4}$, $\ldots$, $c_{N}$, $d$ is derived using the least squares method.

**Theorem 1** Let $P = (a, b, c_{2}, c_{3}, \ldots, c_{N}, d)^{T}$ be the sequence of parameters in the equation,
The original grey model can be written as follows:

\[ X_1^{(0)}(k) = -a z_1^{(1)}(k) + b \left( z_2^{(1)}(k) \right)^r + \sum_{i=2}^{N} c_i z_i^{(1)}(k) + d \]

Perform 1-AGO processing on the original sequence \( X_1^{(0)} \) to get:

\[
\begin{align*}
X_1^{(0)}(2) &= -az_1^{(1)}(2) + b \left( z_2^{(1)}(2) \right)^r + \sum_{i=2}^{N} c_i z_i^{(1)}(2) + d \\
X_1^{(0)}(3) &= -az_1^{(1)}(3) + b \left( z_2^{(1)}(3) \right)^r + \sum_{i=2}^{N} c_i z_i^{(1)}(3) + d \\
&\vdots \\
X_1^{(0)}(n) &= -az_1^{(1)}(n) + b \left( z_2^{(1)}(n) \right)^r + \sum_{i=2}^{N} c_i z_i^{(1)}(n) + d 
\end{align*}
\]  

The matrix form of Eq. (20) can be written as \( Y = BP \), where \( Y \) is the original sequence, \( B \) is the mixed matrix processed by 1-AGO on the original sequence, and \( P \) is the parameter sequence. The error sequence can be written as \( e = Y - BP \).

\[
s = e^T e = (Y - BP)^T (Y - BP)
\]

From this, the expression of the parameter is obtained as follows:

\[
\begin{align*}
a &= b_1 z_1^{(1)}(k) + c_1 z_1^{(1)}(k) + d z_1^{(1)}(k) \sum_{i=2}^{N} c_i z_i^{(1)}(k) - X_1^{(0)}(k) z_1^{(1)}(k) \\
b &= b_2 z_2^{(1)}(k) + c_2 z_2^{(1)}(k) - z_2^{(1)}(k) \sum_{i=2}^{N} c_i z_i^{(1)}(k) - d z_2^{(1)}(k) \\
c_i &= b_3 z_3^{(1)}(k) + c_3 z_3^{(1)}(k) - b z_3^{(1)}(k) \sum_{i=2}^{N} c_i z_i^{(1)}(k)
\end{align*}
\]  

(22)

According to Eq. (22), if the original sequence \( X_1^{(0)} \) is known, then the estimated value of the parameter sequence \( P = (a, b, c_2, c_3, \ldots, c_N, d) \) can be obtained by the least squares method. The prediction in the validation and application is mainly based on the time response equation corresponding to the whitening differential equation, and there is an unknown index \( r \) in it. The model corresponding to the whitening differential equation is a multivariate grey model, which brings some difficulties to the solution of the equation. Hence, the derivation method applicable to the multivariate grey model is considered to solve the equation.

**Derivative method for solving whitening differential equation**

The grey model is

\[
X_1^{(0)}(k) + az_1^{(1)}(k) = b \left( z_2^{(1)}(k) \right)^r + \sum_{i=2}^{N} c_i z_i^{(1)}(k) + d.
\]

\( X_1^{(0)}(k) \) is the original sequence and \( z_i^{(1)}(k) \) is the first-order cumulative sequence. According to Eq. (16) the approximate representation of \( \frac{dx}{dk} \) can be written as follows:

\[
\begin{align*}
\frac{dx}{dt} &= 2X_1^{(0)}(k)z_1^{(1)}(k) + 2az_1^{(1)}(k)z_2^{(1)}(k) + 2b z_1^{(1)}(k) \left( z_2^{(1)}(k) \right)^2 \sum_{i=2}^{N} c_i z_i^{(1)}(k) - 2z_1^{(1)}(k) \sum_{i=2}^{N} c_i z_i^{(1)}(k) - 2dz_1^{(1)}(k) \\
\frac{dx}{dt} &= -2X_1^{(0)}(k)z_2^{(1)}(k) - 2az_1^{(1)}(k)z_2^{(1)}(k) + 2b \left( z_2^{(1)}(k) \right)^2 \sum_{i=2}^{N} c_i z_i^{(1)}(k) + 2dz_1^{(1)}(k) \\
\frac{dx}{dt} &= -2X_1^{(0)}(k)z_3^{(1)}(k) - 2az_1^{(1)}(k)z_3^{(1)}(k) + 2b \left( z_3^{(1)}(k) \right)^2 \sum_{i=2}^{N} c_i z_i^{(1)}(k) + 2dz_1^{(1)}(k) \\
\frac{dx}{dt} &= -2X_1^{(0)}(k) - 2az_1^{(1)}(k) + 2b \left( z_1^{(1)}(k) \right)^2 \sum_{i=2}^{N} c_i z_i^{(1)}(k) + 2dz_1^{(1)}(k)
\end{align*}
\]
\[ \frac{dX_i^{(1)}(k)}{dk} = X_i^{(0)}(k) \]

From the above equation, we get \( X_i^{(1)}(k) - X_i^{(1)}(k - 1) = X_i^{(0)}(k), i = 1, 2, \ldots, n \), that is \( X_i^{(1)}(k) = X_i^{(1)}(k - 1) + X_i^{(0)}(k), i = 1, 2, \ldots, n \).

Then,
\[
z_i^{(1)}(k) = 0.5(X_i^{(0)}(k) + X_i^{(1)}(k - 1)) \\
= 0.5(X_i^{(0)}(k) + X_i^{(1)}(k - 1) + X_i^{(1)}(k - 1)) = 0.5X_i^{(0)}(k) + X_i^{(1)}(k - 1)
\]

Substituting the above formula into Eq. (18) can obtain:
\[
X_i^{(0)}(k) + a\left[0.5X_i^{(0)}(k) + X_i^{(1)}(k - 1)\right] \\
= b\left[0.5X_i^{(0)}(k) + X_i^{(1)}(k - 1)\right] + \sum_{j=2}^{n}\left[0.5X_j^{(0)}(k) + X_j^{(1)}(k - 1)\right] + d
\]

that is:
\[
(1 + 0.5a)X_i^{(0)}(k) + aX_i^{(1)}(k - 1) = (1 + 0.5b)\left[0.5X_i^{(0)}(k) + X_i^{(1)}(k - 1)\right] + 0.5\sum_{j=2}^{n}X_j^{(0)}(k) + \sum_{j=2}^{n}\left[0.5X_j^{(0)}(k) + X_j^{(1)}(k - 1)\right] + d, N = 4
\]

If \( 1 + 0.5a \neq 0 \), then:
\[
X_i^{(0)}(k) = -\frac{a}{1 + 0.5a}X_i^{(1)}(k - 1) + \frac{b}{1 + 0.5a}\left[0.5X_i^{(0)}(k) + X_i^{(1)}(k - 1)\right] + 0.5\sum_{j=2}^{n}X_j^{(0)}(k) + \sum_{j=2}^{n}\left[0.5X_j^{(0)}(k) + X_j^{(1)}(k - 1)\right] + \frac{d}{1 + 0.5a}
\]

Equation (23) is the time response equation of the whitened differential equation corresponding to the grey model ENGM(1, N). The time response equation of the whitened differential equation corresponding to the ENGM(1, 4) model can then be written as Eq. (24):
\[
X_i^{(0)}(k) = -\frac{a}{1 + 0.5a}X_i^{(1)}(k - 1) + \frac{b}{1 + 0.5a}\left[0.5X_i^{(0)}(k) + X_i^{(1)}(k - 1)\right] + \frac{c_1}{1 + 0.5a}\left[0.5X_1^{(0)}(k) + X_1^{(1)}(k - 1)\right] + \frac{c_2}{1 + 0.5a}\left[0.5X_2^{(0)}(k) + X_2^{(1)}(k - 1)\right] + \frac{c_3}{1 + 0.5a}\left[0.5X_3^{(0)}(k) + X_3^{(1)}(k - 1)\right] + \frac{c_4}{1 + 0.5a}\left[0.5X_4^{(0)}(k) + X_4^{(1)}(k - 1)\right] + \frac{d}{1 + 0.5a}
\]

Parameter optimization

The particle swarm algorithm is a global random search algorithm based on swarm intelligence proposed by Kennedy and Eberhart, inspired by the results of artificial life research, by simulating the migration and swarming behavior of birds in the foraging process (Kennedy and Eberhart 1995). The traditional particle swarm algorithm is easy to lead to the local optimum, so the quantum particle swarm algorithm has achieved optimization and improvement in the search mechanism. To overcome the uncertainty of quantum mechanics, the wave function is used to represent the particle shape, and the Schrödinger equation is employed to obtain the probability density function of each individual at any position. The Monte-Carlo stochastic simulation is introduced to obtain the position equations of particles in the problem space (Xia et al. 2019), which can improve the global search capability of the population and the accuracy of the algorithm. Quantum particle swarm algorithm is widely used in various fields (Feng et al. 2017; Li et al. 2019; Song et al. 2020; Turgut 2016).

According to the least squares estimation, all other parameters of the model can be represented by the power exponent \( r \). The value of \( r \) largely affects the prediction results and accuracy of the model. To obtain the optimal parameter combination, the average absolute percentage error (MAPE) of the model is selected to measure the estimation accuracy of the model. The smaller the MAPE, the smaller the unexplainable error of the model and the higher the accuracy of the model. The power index \( r \) is the optimal object, \( MAPE(r) \) is the optimal function (Eq. (25)), and the nonlinear optimization model is constructed as shown in Eq. (26) by combining the parameter constraints in the above time response equation (Fig. 2).

\[
MAPE(r) = \frac{1}{n - 1} \sum_{k=2}^{n} \left| \frac{X_i^{(0)}(k) - X_i^{(1)}(k)}{X_i^{(1)}(k)} \right| \times 100\%
\]

\[
X_i^{(0)}(k) = -\frac{a}{1 + 0.5a}X_i^{(1)}(k - 1) + \frac{b}{1 + 0.5a}\left[0.5X_i^{(0)}(k) + X_i^{(1)}(k - 1)\right] + \frac{c_1}{1 + 0.5a}\left[0.5X_1^{(0)}(k) + X_1^{(1)}(k - 1)\right] + \frac{c_2}{1 + 0.5a}\left[0.5X_2^{(0)}(k) + X_2^{(1)}(k - 1)\right] + \frac{c_3}{1 + 0.5a}\left[0.5X_3^{(0)}(k) + X_3^{(1)}(k - 1)\right] + \frac{c_4}{1 + 0.5a}\left[0.5X_4^{(0)}(k) + X_4^{(1)}(k - 1)\right] + \frac{d}{1 + 0.5a}
\]

Model verification

China, the USA, and Japan are the top three countries in the world in terms of GDP as well as among the top five countries in terms of carbon emissions. This section validates the accuracy and validity of the ENGM(1, 4) model using the transportation sector carbon emissions dataset for these three countries for the 18-year period of 2001–2018. The transportation sector carbon emissions and transportation sector energy consumption data are from the IEA (International Energy Agency) (IEA 2021), and the GDP and population data from the World Bank (IBRD 2021). The 13 years of
data from 2001 to 2013 are used to fit to build the model, and the 5 years of data from 2014 to 2018 are used to forecast to verify the accuracy of the model. The different transportation status, economic development, and energy use of these three countries constitute three carbon emission trends with different characteristics, as shown in Fig. 3. The transportation sectors of these three countries show increasing, fluctuating, and decreasing carbon emissions during the 18 years from 2001 to 2018, respectively. The validation of the model with these three datasets with different characteristics shows that the model can adapt to different types of data series with high accuracy and flexibility.

We compare the results with GM(1,4), GMC(1,4), SVR (support vector regression), ANN (artificial neural network), and ARIMA (autoregressive integrated moving average model). The indicators for evaluating the predictive effect are shown in Table 3. In the calculation result table, the numbers in bold are the optimal values.
Fig. 3  Comparison of carbon emissions data from China, the USA, and Japan

Table 3  Metrics for evaluating effectiveness of the models

| Name                                           | Abbreviation | Formulation                                                                 |
|------------------------------------------------|--------------|-----------------------------------------------------------------------------|
| The mean absolute simulation percentage error  | MAPE<sub>SIM</sub> | $\frac{1}{n-1}\left(\sum_{i=2}^{n} \left| \frac{x^{(i)} - \hat{x}^{(i)}}{x^{(i)}} \right| \right) \times 100\%$ |
| The mean absolute prediction percentage error  | MAPE<sub>PRE</sub> | $\frac{1}{n} \left(\sum_{i=1}^{n} \left| \frac{x^{(i)} - \hat{x}^{(i)}}{x^{(i)}} \right| \right) \times 100\%$ |
| The total mean absolute percentage error       | MAPE<sub>TOT</sub> | $\frac{1}{n-1}\left(\sum_{i=2}^{n} \left| \frac{x^{(i)} - \hat{x}^{(i)}}{x^{(i)}} \right| \right) \times 100\%$ |
| Mean absolute percentage error                 | MAE          | $\frac{1}{n-1}\sum_{i=2}^{n} \left| x^{(i)} - \hat{x}^{(i)} \right|$                                      |
| Root mean square error                         | RMSE         | $\sqrt{\frac{1}{n-1}\sum_{i=2}^{n} \left( x^{(i)} - \hat{x}^{(i)} \right)^2}$               |
| Standard deviation                              | STD          | $\sqrt{\frac{1}{n-1}\sum_{i=2}^{n} \left( \frac{x^{(i)} - \hat{x}^{(i)}}{x^{(i)}} \right)^2 - \text{MAPE}^2}$ |
| The absolute percentage error                  | APE          | $\left( \frac{x^{(i)} - \hat{x}^{(i)}}{x^{(i)}} \right) \times 100\%$                           |
Validation case 1: China

The dataset of the first verification case includes China’s transportation sector carbon emissions, GDP, energy consumption, and population data from 2001 to 2018. Carbon emissions show an increasing trend. The parameter value $r = 3$ of the ENGM(1,4) model is obtained by the quantum particle swarm algorithm. The fitted and predicted values of each model and the MAPE values are shown in Table 4. Other comparison indices are shown in Table 7. Table 4 shows that the new models MAPESIM and MAPETOT are the lowest among all models with $0.4465\%$ and $1.8963\%$, respectively. MAPEPRE is $5.3759\%$, which is the second lowest among all models, indicating that the overall fit and prediction of the model is relatively good. The three types of MAPE values of the GM(1,4) model are also small, all within $3\%$, indicating that the GM(1,4) model is also suitable for the background of this research. The three types of MAPE values of the GMC(1,4) model all exceed $10\%$, indicating that the fitting and prediction effects are not ideal. The MAPESIM of both ANN and SVR are less than $5\%$, which is a good fit. However, the MAPEPRE is more than $10\%$, which is a large prediction error due to the phenomenon of “overfitting.” The fitting error of the ARIMA model is as high as $20\%$, indicating that the model is not suitable for the background of this research.

To further demonstrate the approximation effect of fitting and prediction, the fitted and predicted values of each model are transformed into curve trend plots and APE comparison plots, as shown in Fig. 4. In the trend plot (Fig. 4a), the ENGM(1,4) and GM(1,4) models fit the original series more closely. The ENGM(1,4) model is slightly higher than the original value in the prediction part, and the GM(1,4) model shows a slight fluctuation at the beginning of the initial few time points. Overall, these two models seem to have a higher degree of approximation to the original value. The GMC(1,4) model fits the original values better at the initial time point, and then shows large fluctuations and deviates from the original values. SVR and ANN models are closer to the original sequence in the fitting part. The predicted value is significantly lower than the original value though, and the prediction error is large. The ARIMA model shows a large fluctuation at the first time point. Then, the fitted and predicted values are slightly lower than the original series overall. Figure 4b shows the APE value of each model. Of note, the APE of the ARIMA model at the first point in time is as high as $200\%$, which is excluded for the esthetics and readability of the image. The ENGM(1,4) model has an APE value of less than $5\%$ at most of the remaining time points, excluding the APE value of more than $10\%$ at time point in 2018. The overall fitting and forecasting errors are small.
In summary, although there are individual indicators that are non-optimal, the ENGM(1,4) model can fit and predict the carbon emissions of China’s transportation sector most effectively in terms of fitting and predicting the overall situation.

Validation case 2: the USA

The dataset of the second verification case includes the transportation sector carbon emissions, GDP, energy consumption, and population data of the USA from 2001 to 2018. Carbon emissions show a fluctuating trend. The parameter value $r = 0.1840$ of the ENGM(1,4) model is obtained by the quantum particle swarm algorithm. The fitted and predicted values of each model and the MAPE values are shown in Table 5. Other comparison indices are shown in Table 7. Table 5 shows that the MAPE$_{SIM}$ and MAPE$_{TOT}$ of the new model are the lowest among all models, with 0.2367% and 1.0495%, respectively, which are less than 2%. This outcome implies that the fitting and overall accuracy are higher. The MAPE$_{PRE}$ of the GM(1,4) model is 0.2275%, which is the lowest among all models and nearly 3% lower than the 3.0004% of the new model. The new model is inferior to the GM(1,4) model in terms of predictive ability. From the other comparison indices in Table 7, the MAE, RMSE, and STD of the new model are the lowest, so the ENGM(1,4) model seems to be the overall best. In addition, the fitting accuracy of the SVR and ANN models is high, but the prediction
results are not ideal. The phenomenon of “overfitting” occurs. The MAPESIM of the GMC(1,4) model is as high as 50.4846%, and the performance of other indicators is not ideal, indicating that the GMC(1,4) model is not suitable for the research background.

To further demonstrate the approximation effect of fitting and prediction, the fitted and predicted values of each model are transformed into curve trend plots and APE comparison plots, as shown in Fig. 5. From the trend plot (Fig. 5a), except for the GMC(1,4) model, which deviates from the original values to a large extent, the other five models are able to stick to the original series to different degrees. The GM(1,4) model shows significant fluctuations at the first few time points, deviating from the original values, and then performs better in the subsequent fitting and prediction. SVR, ANN, and ARIMA are all closer to the original series in the fitting part, but SVR is much higher than the original value in the prediction part. ANN and ARIMA are lower than the original value. Compared with the other models, the ENGM(1,4) model has the highest approximation to the original series, with only slight fluctuations in the prediction part. The APE comparison chart (Fig. 5b) illustrates that the APE value performance is consistent with the trend plot. The APE value of the new model is less than 10% at most time points, especially in the fitting part, which is around 1%. The overall performance is optimal compared with other models.

In summary, even if there are individual indicators that are non-optimal, from the perspective of fitting and predicting the overall situation, the ENGM(1,4) model can most effectively fit and predict the carbon emissions of the US transportation sector.

Validation case 3: Japan

The dataset of the third verification case includes Japan’s transportation sector carbon emissions, GDP, energy consumption, and population data from 2001 to 2018. Carbon emissions show a decreasing trend. The parameter value $r = 0.9630$ of the ENGM(1,4) model is obtained by the quantum particle swarm algorithm. The fitted and predicted values of each model and the MAPE values are shown in Table 6. Other comparison indices are shown in Table 7. Table 6 shows that the MAPE$_{SIM}$ and MAPE$_{TOT}$ of the new model are the lowest among all models, with 0.4967% and 1.5404%, respectively, which are less than 2%. This outcome indicates that the overall accuracy is higher. The MAPE$_{PRE}$ of the GM(1,4) model is 0.6025%, which is the lowest among all the models and nearly 4% lower than the
4.0454% of the new model. The predictive ability of the new model is not as good as that of the GM(1,4) model. From the other comparison indices in Table 7, it seems that the new model has the lowest MAE and RMSE among all the models except for the STD, which is lower than that of the ANN. Overall, the ENGM (1,4) model is the best. The MAPE$_{SIM}$ of the SVR and ANN models are 0.6216% and 0.8156%, respectively, both of which are less than 1%, indicating that these two models fit better, and their MAPE$_{PRE}$ are 7.3684% and 4.7280%, respectively, both of which are less than 10%, indicating that these two models are also better adapted to the context of this study, but their prediction effects and overall accuracy are not as good as the new model. All three MAPE values of the GMC(1,4) model are around 25% and the rest of the indicators performed poorly, indicating that GMC(1,4) has poor accuracy in this study context.

To further demonstrate the approximation effect of fitting and prediction, the fitted and predicted values of each model are transformed into curve trend plots and APE comparison plots, as shown in Fig. 6. As can be seen from the Fig. 6a trend plot, the prediction and fitting trends of GMC(1,4) are significantly lower than the original series with larger errors, and the remaining five models are able to fit the original series to varying degrees. The GM(1,4) and ARIMA models generally coincide with the original sequence, except for the large fluctuations at the previous several time points that deviate from the original value, indicating that the errors of these two models are relatively small. The SVR and ANN models are very close to the original series in the fitting part, but both are significantly higher than the true values in the prediction part, indicating that these two models are not good enough for prediction. Compared to the other models, ENGM(1,4) approximates the original series the most, with only slight fluctuations in the prediction part, which is the closest to the original value as a whole. From the comparison of APE values in Fig. 6b, the APE values of ENGM(1,4) are the lowest among all models in the fitting part, and the APE values in the prediction part are also within 10%, so the APE performance of ENGM(1,4) seems to be the best overall.

In summary, even if individual indicators are non-optimal, from the perspective of fitting and predicting the overall situation, the ENGM(1,4) model can most effectively fit and predict the carbon emissions of Japan’s transportation sector.

| Year | Actual values | ENGM(1,4) | GM(1,4) | GMC(1,4) | SVR | ANN | ARIMA |
|------|---------------|----------|---------|----------|-----|-----|-------|
|      | Simulated data | Simulated data | Simulated data | Simulated data | Simulated data | Simulated data | Simulated data |
| 2001 | 263           | 263      | 263     | 263      | 261 | 265 | 199   |
| 2002 | 259           | 260      | 224     | 113      | 260 | 261 | 263   |
| 2003 | 256           | 256      | 282     | 192      | 258 | 258 | 255   |
| 2004 | 256           | 257      | 264     | 198      | 256 | 254 | 256   |
| 2005 | 254           | 256      | 255     | 198      | 253 | 251 | 256   |
| 2006 | 250           | 252      | 250     | 195      | 248 | 247 | 254   |
| 2007 | 243           | 245      | 243     | 190      | 241 | 242 | 250   |
| 2008 | 234           | 235      | 234     | 184      | 235 | 234 | 243   |
| 2009 | 228           | 228      | 228     | 178      | 230 | 227 | 233   |
| 2010 | 229           | 230      | 229     | 177      | 227 | 223 | 227   |
| 2011 | 223           | 224      | 223     | 176      | 225 | 222 | 228   |
| 2012 | 221           | 222      | 221     | 173      | 222 | 221 | 222   |
| 2013 | 218           | 220      | 219     | 170      | 220 | 221 | 220   |
| MAPE$_{SIM}$ | 0.4967      | 2.3464   | 25.0597 | 0.6216   | 0.8156 | 3.2016 |
| Prediction |                   |          |         |          |     |     |       |
| 2014 | 212           | 217      | 213     | 165      | 219 | 220 | 218   |
| 2015 | 211           | 209      | 213     | 161      | 220 | 219 | 214   |
| 2016 | 209           | 218      | 210     | 160      | 223 | 219 | 210   |
| 2017 | 207           | 198      | 208     | 159      | 227 | 218 | 206   |
| 2018 | 205           | 223      | 206     | 157      | 232 | 217 | 202   |
| MAPE$_{PRE}$ | 4.0454      | 0.6025   | 23.1431 | 7.3684   | 4.7280 | 1.2414 |
| MAPE$_{TOT}$ | 1.5404      | 1.8335   | 24.4960 | 2.4957   | 1.9024 | 2.6571 |
Fig. 5 Fitted and predicted trends and APE values for each model in case 2

| Country | Index | ENGM(1,4) | GM(1,4) | GMC(1,4) | SVR | ANN | ARIMA |
|---------|-------|-----------|---------|----------|-----|-----|-------|
| China   | MAE   | 15.6352   | 9.4103  | 95.9102  | 64.7824 | 44.4606 | 70.5802 |
|         | RMSE  | 34.9310   | 19.5100 | 129.5526 | 121.9970 | 63.6885 | 159.6783 |
|         | STD   | 0.0335    | 0.0375  | 0.1058   | 0.1110  | 0.0474  | 0.5571 |
| The USA | MAE   | 17.9124   | 31.3363 | 660.0481 | 69.3164 | 45.6668 | 75.1263 |
|         | RMSE  | 29.4537   | 73.7406 | 841.3790 | 121.4052 | 68.9790 | 98.1774 |
|         | STD   | 0.0136    | 0.0381  | 0.2914   | 0.0577  | 0.0297  | 0.0364 |
| Japan   | MAE   | 3.2985    | 4.6182  | 56.9676  | 5.3140  | 4.1557  | 6.5597 |
|         | RMSE  | 5.5335    | 10.8091 | 61.2977  | 9.0988  | 5.6492  | 15.6659 |
|         | STD   | 0.0218    | 0.0377  | 0.0800   | 0.0361  | 0.0188  | 0.0539 |
Analysis of results

The three validation cases mentioned above are analyzed. By comparing the fitting and prediction results and the values of each statistical indicator and parameter, the following conclusions are drawn.

(1) The parameter values of the ENGM(1,4) model for the three validation cases are shown in Table 8.

| Parameters | $r$ | $a$ | $b$ | $c_2$ | $c_3$ | $c_4$ | $d$ |
|------------|-----|-----|-----|-------|-------|-------|-----|
| China      | 3   | 42.8068 | 5.84e-15 | −0.0203 | 0.1296 | 2.13e-05 | 323.2469 |
| The USA    | 0.1840 | −31.7518 | 290.4235 | 0.0293 | −0.1106 | 0.2562 | −606.5978 |
| Japan      | 0.9629 | −9.0313 | 0.0086 | −0.0041 | −0.0271 | 0.0006 | 237.9455 |

$r$ and $c_2$ are the indices and coefficients of multiple variables of GDP, where $r$ can reflect the impact of GDP on changes in carbon emissions. The value of $r$ in the case of
China is as high as 3, and the value of $r$ in the cases of the USA and Japan is less than 1. Based on the data characteristics and economic development of the three countries, China’s GDP has increased by approximately 4.6 times since its accession to the World Trade Organization in 2001 to 2011. The implementation of the “12th Five-Year” and “13th Five-Year” plans, which were implemented from 2011 to 2018, resulted in great efforts to develop economic policies that have allowed GDP to continue to maintain rapid growth. The economic development has brought the transportation industry to take off. Therefore, the index $r$ in China’s case is the largest; that is, GDP changes have a high degree of impact on carbon emissions. Compared with China, the USA and Japan, which have already entered the ranks of developed countries, have shown steady economic growth in the past 20 years, so they have relatively small $r$.

The coefficient $a$ can reflect the characteristics of the original series data to some extent. $a$ is 42.8068, $r$ - 31.7518, and $r$ - 9.0313 in the three cases of China, the USA, and Japan, respectively, which show increasing, fluctuating, and decreasing characteristics of carbon emissions in the transportation sector in the last 20 years. The increase in China from 2001 to 2018 was as high as 250.3788%, while the USA showed a decreasing trend from 2005 to 2014, with a decrease of 9.2920%; and Japan decreased by 28.2926% from 2001 to 2018. These indicate that the positive and negative values of $a$ can represent the increase or decrease of the original series and that the size of $a$ can reflect the increase or decrease.

The coefficients $b$, $c_3$, and $c_4$ can reflect the impact of GDP, energy consumption in the transportation sector, and population on the carbon emissions of various countries to a certain extent. If the coefficient is positive, it is positively correlated, and if the coefficient is negative, it is negatively correlated.

The constant term $d$ represents the degree of influence of random disturbance factors on the model, which can somewhat balance the error and improve the accuracy of the model.

(2) Comparing the ENGM(1,4) model with the other five models, it can be seen by the three types of MAPE indicators that in these three national transportation sector carbon emission forecasts show that there are cases in which the comparison indicators of other models are better than the new model. Overall though, the new model has the best fitting and forecasting effect. The APE values at each time point are lower, implying that the new model has less error and higher validity. The three categories of MAE, RMSE, and STD of each model are compared in Table 7. The indicators of the new model are optimal or suboptimal compared with those of other models. The ENGM(1,4) model thus has high robustness in carbon emission prediction. The GM(1,4) model also achieves ideal results in fitting and prediction, indicating that it is also suitable for the background of this research. SVR and ANN models have high accuracy in fitting, but the phenomenon of “overfitting” makes their prediction errors large. The GMC(1,4) model has poor indicators and calculation results in the verification case, so the model is not suitable for the background of this research.

(3) The three application cases of China, the USA, and Japan in the “Validation case 1: China” to “Validation case 3: Japan” sections respectively apply the ENGM(1,4) model to three types of data series with different characteristics: increasing, fluctuating, and decreasing. The model is built with data from 2001 to 2013 to predict carbon emissions from the transportation sector in each country for the next 5 years from 2014 to 2019. The trend plots of the fitted and predicted values and each statistical indicator can illustrate the flexibility and validity of the new model with high accuracy. The new model can be adapted to the research context of short-term carbon emission prediction with data characterized by incomplete information.

**Application**

**Further prediction**

In the fifth part, by comparing it with six other five models, it is proved that the ENGM(1,4) model can better fit and predict the carbon emissions of the transportation sectors of China, the USA, and Japan. In this section, the new model is used to predict the carbon emissions of the transportation sector in these three countries for the next 7 years. The corresponding analysis and recommendations are given based on the prediction results. It should be noted that carbon emissions are energy data, and there is usually a time delay in the statistical process. The latest carbon emissions data for these three countries from the International Energy Agency (IEA) are currently available as of 2018.

The multivariate grey model relies on the values of other influencing factors in the forecast. As such, the values of the three influencing factors of GDP ($X_2$), total energy consumption in the transportation sector ($X_3$), and population ($X_4$) for the three countries are predicted for the next 7 years using the GM(1,1) model based on the data of the last 10 years from 2009 to 2018 before the prediction. The results are shown in Table 9. Based on the data of the three countries for the 18 years (2001–2018), the ENGM(1,4) model is used to predict the value of carbon emissions in the transportation sector for the 7 years (2019–2025) for each country. The predicted values and the rate of change are obtained as shown in
Table 10. The values of the model parameters are provided in Table 11, and the trend graphs of the fitted and predicted values compared with the real values are depicted in Fig. 7. The results of the projections for the three countries according to Table 10 and Fig. 7 are analyzed as follows.

(1) China’s carbon emissions from the transportation sector show a steady upward trend from 2019 to 2025, with an average growth rate of 5.3692%. The slope of the change curve in 2019–2020 is significantly lower than in 2021–2025, due to the impact of the COVID-19 and a significant decrease in the transportation travel rate as a result of the government’s closure management measures. In the “14th Five-Year Plan,” China specifies the goal of peaking carbon in 2030, in which transportation carbon emissions account for 10% of the country’s total carbon emissions. At present, China’s new energy industry is in a stage of vigorous development, and the transportation sector’s energy conservation and emission reduction policies are gradually being implemented. Owing to the wide coverage of the policy and the long period of time, related technical research is still in its infancy. According to the “Report on China’s Car Ownership Data,” the number of private and public travel vehicles continues to increase, so low-carbon travel-related measures are effective yet slow. CICC research data points out that China’s overall carbon emission growth rate in the past 20 years is 5.6%, which is higher than the average growth rate predicted in this study. By comparing the slope of the graph, it can be seen that the future growth rate of carbon emissions in the transportation sector will slow down to a certain extent. This partly explains the effectiveness of

| Country | Year | X1 (GDP ($100 million)) | X2 (Energy consumption (k toe)) | X3 | X4 (Population (10 thousand)) | MAPESIM |
|---------|------|------------------------|-------------------------------|----|-----------------------------|---------|
| China   | 2019 | 150,710                | 365,820                       | 139,990 |                          | 4.5466% |
|         | 2020 | 164,090                | 389,420                       | 140,700 |                          | 2.2507% |
|         | 2021 | 178,650                | 414,530                       | 141,420 |                          | 0.0227% |
|         | 2022 | 194,500                | 441,270                       | 142,140 |                          |         |
|         | 2023 | 211,760                | 469,740                       | 142,860 |                          |         |
|         | 2024 | 230,550                | 500,040                       | 143,590 |                          |         |
|         | 2025 | 251,010                | 532,290                       | 144,330 |                          |         |

Table 10. The values of the model parameters are provided

| Parameter | r    | a      | b      | c2   | c3   | c4   | d        |
|-----------|------|--------|--------|------|------|------|----------|
| China     | 0.0002 | 2.4938 | -1.447563.33 | 0.0045 | 0.0039 | 0.0037 | 1,450,337.77 |
| US        | 1.2647 | 9.5767 | -5.6874e-06  | -0.0108 | 0.0321 | -0.0442 | 1715.0059  |
| Japan     | 2.3764 | 6.6212 | -7.6142e-13 | 3.8825e-05 | 0.0194 | -0.0006 | 270.9546   |
transportation reform and the reliability of the conclusions of this paper.

(2) The overall carbon emissions of the US transportation sector show a fluctuating upward trend, with an average rate of change of 0.9392%. The characteristics of changes from 2019 to 2021 are related to the COVID-19. The US epidemic monitoring and control policies lagged behind those of China, and related measures were gradually adopted in the second half of 2020. Therefore, the carbon emissions of its transportation sector were on the rise in 2019–2020, declined in 2020–2021, and then will rise in volatility. The slow growth rate and low overall rate of change in carbon emissions from the US transportation sector are consistent with the state of the country's transportation development and national energy policy. The USA is the first country to implement carbon emission standards. It has implemented the optimization of energy structure and transportation structure very early, and it has the world’s most complete integrated transportation system. Figure 7 shows that carbon emissions from the US transportation sector are in a relatively stable state. The growth of automobiles and population has brought fluctuations within a reasonable range, which shows that the prediction results of this article are in line with the actual situation.

(3) The overall carbon emission of Japan’s transportation sector shows a steady downward trend, with an average decline rate of 1.2231%. The slope of the curve shows a slow decline from 2019 to 2020 and a faster decline after 2020, which can be partly explained by the impact of the COVID-19. The COVID-19 in Japan was gradually revealed by 2020, and the policy of staying at home and reducing travel made the decline of transportation carbon emissions faster in 2020–2021. The characteristics of Japan’s carbon emissions are related to its national conditions and transportation energy policies. On the one hand, Japan is the most aging country in the world, with a decreasing trend in population changes. The number of transportation vehicles also grows slowly, so the level of carbon emissions from the transportation sector is low and on a decreasing trend. On the other hand, according to the statistics and reports of the National Energy Agency, Japan is short of energy resources, has a low self-sufficiency rate, and mostly relies on imports. Therefore, the Japanese gov-
ernment began to implement new energy policies and low-carbon reforms in the transportation industry in the twentieth century, and achieved better results as early as around 2000. Japan’s national conditions, energy, and transportation conditions provide support for the predicted results of this paper.

Analysis and expansion

In the “Further prediction” section, the ENGM(1,4) model is used to forecast and analyze the carbon emissions of the transportation sector in three countries, China, the USA, and Japan, respectively, for the 7 years from 2019 to 2025. Based on the projection results, the following conclusions and insights are drawn.

(1) The different economic developments, energy policies, and transportation reforms of these three countries have resulted in three carbon emission data profiles for the transport sector. As developed countries, the USA and Japan have implemented energy structure optimization and low-carbon travel policies in the twentieth century, and the related technology development has been at the world’s leading level, so the overall change trend of carbon emissions is not significant. Among them, the impact of Japan’s serious population aging on transportation travel has led to a decreasing trend of carbon emissions. China’s rapid economic development in recent years has led to major changes in the transportation industry. China’s implementation of related environmental protection policies is lagging behind the USA and Japan. Therefore, the carbon emissions of China’s transportation sector will continue to grow for a period of time in the future, but the growth rate will slow down.

(2) Drawing on the energy policies of the USA and Japan, we focus on providing some suggestions for the future development of carbon emissions from the transportation sector in China. First of all, the government and relevant units can learn from the US transportation structure reform measures, and accelerate the progress of transportation energy-saving and emission reduction-related technology research, so as to improve the level of transportation intelligence thus enhance the intelligent transportation system, and establish a more scientific and environmentally friendly travel system from a macro perspective. Second, complete fuel economy standards and low-carbon fuel standards must be formulated. On the one hand, advanced technologies can be used to improve engine equipment and in turn fuel economy efficiency. On the other hand, the government should strengthen the management of transportation tools and strictly control the emissions carbon in the transportation sector through policies and methods. Finally, the research and development (R&D) of new automotive technologies and popularization of new energy vehicles should be strengthened by relying on China’s well-known auto companies. Energy-saving and environmentally friendly vehicles should be vigorously developed, and automotive innovation and industry alliances must be promoted. A public R&D led by the government should be established, with the joint participation of enterprises and research institution platforms.

Conclusion

In this study, a new nonlinear multivariate grey model based on the EKC is proposed for use in the analysis of carbon emissions from the transportation sector in China, the USA, and Japan. According to the verification cases and predictions of the three countries, the following conclusions are drawn:

(1) Considering the change of energy structure and the data characteristics of carbon emissions with incomplete information in short-term forecasting and basing on the theory related to EKC, this study combines the characteristics of IPAT equation and extended STIRPAT model, and identifies the factors affecting carbon emissions in transportation sector as GDP, energy consumption in transportation sector, and population. The index of GDP in the EKC assumption is extended from 2 to \( r \) to further enhance the flexibility of the model. A new type of ENGM(1,N) model is established on the basis of the differential information principle. The least squares estimation is used to obtain the model parameters, and the quantum particle swarm optimization algorithm is designed to optimize the unknown index and improve the model accuracy. The new model is a coupling of multiple models and methods, with multiple characteristics, and it provides a new idea for carbon emission prediction.

(2) The economic development, energy policy, and transportation conditions of China, the USA, and Japan differ greatly, with carbon emissions in the transportation sector showing increasing, fluctuating, and decreasing characteristics in the past 20 years, respectively. The ENGM(1,4) model is compared with five models, namely, GM(1,4), GMC(1,4), SVR, ANN, and ARIMA, through three validation case sets. The new model is used in the prediction of carbon emission data from the transportation sector for these three countries for the 7 years (2019–2025). The results show that the ENGM(1,4) model has the best fitting and prediction.
effect than other models, and can better adapt to the data characteristics of carbon emissions.

Although the new model has strong theoretical and practical significance, the model needs to be further improved and perfected:

(1) The degree of influence of different factors on carbon emissions should be further explored. A data-based structure selection algorithm can be considered to determine the coefficient values in a more intuitive and scientific way, so as to clarify the laws of parameters’ values and analyze the correlation degree of influencing factors and the development trend of prediction results, and fully explore the value of the model.

(2) Through the research in this paper, it can be seen that the prediction and early warning of carbon emissions are of great significance for planning the low-carbon development of specific industries. In the future, from a sectoral perspective, the model can be further extended to carbon emissions research in other industries, such as the industrial sector and the power sector, which account for a large proportion of carbon emissions. The new model can also be used to study carbon emissions from different fossil energy sources (e.g., coal, oil, natural gas) from an energy perspective. In addition, from the model perspective, the new model of this paper is proposed based on the EKC hypothesis to explore the relationship between the environment and the economy, so that it can also be considered for application to study the association between energy consumption and economic and other influencing factors in different types and sectors.

Acknowledgements We would like to thank the editor and the anonymous reviewers for their helpful comments.

Author contribution Siyuan Huang: methodology, data curation, writing—original draft, visualization, formal analysis. Xinqing Xiao: conceptualization, methodology, supervision, writing—review and editing. Huan Guo: writing—review and editing, supervision.

Funding This work is supported by the National Natural Science Foundation of China (Nos. 71871174) and the Fundamental Research Funds for the Central Universities (WUT: 2019IA005).

Data availability The authors confirm that the data supporting the findings of this study are available within the article.

Declarations

Ethics approval and consent to participate There is no ethical approval and patient consent to participate is required for this study.

Consent for publication The authors confirm that the final version of the manuscript has been reviewed, approved, and consented for publication by all authors.

Competing interests The authors declare no competing interests.

References

Agbulut U (2022) Forecasting of transportation-related energy demand and CO2 emissions in Turkey with different machine learning algorithms. Sustainable Production and Consumption 29:141–157. https://doi.org/10.1016/j.spc.2021.10.001

Ali G, Ashraf A, Bashir MK et al (2017) Exploring environmental Kuznets curve (EKC) in relation to green revolution: a case study of Pakistan. Environ Sci Policy 77:166–171. https://doi.org/10.1016/j.envsci.2017.08.019

Altıntaş H, Kassouri Y (2020) Is the environmental Kuznets curve in Europe related to the per-capita ecological footprint or CO2 emissions. Ecol Ind 113:106187. https://doi.org/10.1016/j.ecolind.2020.106187

Aslam B, Hu JS, Hafeez M et al (2021) Applying environmental Kuznets curve framework to assess the nexus of industry, globalization, and CO2 emission. Environ Technol Innov 21:101377. https://doi.org/10.1016/j.eti.2021.101377

Bibi F, Jamil M (2021) Testing environment Kuznets curve (EKC) hypothesis in different regions. Environ Sci Pollut Res 28(11):13581–13594. https://doi.org/10.1007/s11356-020-11516-2

Bilgili F, Koçak E, Bulut Ü et al (2016) The dynamic impact of renewable energy consumption on CO2 emissions: a revisited environmental Kuznets curve approach. Renew Sustain Energy Rev 54:838–845. https://doi.org/10.1016/j.rser.2015.10.080

Bilgili F, Kuşkaya S, Khan M et al (2021a) The roles of economic growth and health expenditure on CO2 emissions in selected Asian countries: a quantile regression model approach. Environ Sci Pollut Res 28:44949–44972. https://doi.org/10.1007/s11356-021-13639-6

Bilgili F, Nathaniel SP, Kuşkaya S et al (2021b) Environmental pollution and energy research and development: an environmental Kuznets curve model through quantile simulation approach. Environ Sci Pollut Res 28:53712–53727. https://doi.org/10.1007/s11356-021-14506-0

Bilgili F, Ozturk I, Koçak E et al (2022a) The nexus between access to electricity and CO2 damage in Asian Countries: the evidence from quantile regression models. Energy and Buildings 256(1):115161. https://doi.org/10.1016/j.enbuild.2021.11761

Bilgili F, Khan M, Awan A (2022b) Is there a gender dimension of the environmental Kuznets curve? Evidence from Asian countries. Environ Sci Pollut Res. https://doi.org/10.1007/s11356-022-02139-3

Brown L, McFarlane A, Campbell K et al (2020) Remittances and CO2 emissions in Jamaica: an asymmetric modified environmental Kuznets curve. The Journal of Economic Asymmetries 20:e00166. https://doi.org/10.1016/j.jea.2020.e00166

Cao Y, Yin KD, Li XM et al (2021) Forecasting CO2 emissions from Chinese marine fleets using multivariable trend interaction grey model. Appl Soft Comput 104:107220. https://doi.org/10.1016/j.asoc.2021.107220

Chontanawat J (2018) Decomposition analysis of CO2 emission in ASEAN: an extended IPAT model. Energy Procedia 153:186–190. https://doi.org/10.1016/j.egypro.2018.10.057

Deng JL (1982) Control problems of grey systems. System Control Letter 1:288–294. https://doi.org/10.1016/j.tsl.2017.09.014
Ding S, Xu N, Ye J et al (2020) Estimating Chinese energy-related CO₂ emissions by employing a novel discrete grey prediction model. J Clean Prod 259(20):120793. https://doi.org/10.1016/j.jclepro.2020.120793

Dogan E, Inglesi-Lotz R (2020) The impact of economic structure to the environmental Kuznets curve (EKC) hypothesis: evidence from European countries. Environ Sci Pollut Res 27(11):12717–12724. https://doi.org/10.1007/s11356-020-07878-2

Dong KY, Sun RJ, Jiang HD et al (2018) CO₂ emissions, economic growth, and the environmental Kuznets curve in China: what roles can nuclear energy and renewable energy play. J Clean Prod 196(20):51–63. https://doi.org/10.1016/j.jclepro.2018.05.271

Duan HM, Wang D, Pang XY et al (2020) A novel forecasting approach based on multi-kernel nonlinear multivariable grey model: a case report. J Clean Prod 260(1):120929. https://doi.org/10.1016/j.jclepro.2020.120929

Ehrlich P, Holdren J (1970) The people problem. Saturday Review 4:42–43

Fang DB, Zhang XL, Yu Q et al (2018) A novel method for carbon dioxide emission forecasting based on improved Gaussian processes regression. J Clean Prod 173(1):143–150. https://doi.org/10.1016/j.jclepro.2017.05.102

Feng Z, Niu W, Cheng W (2017) Multi-objective quantum-behaved particle swarm optimization for economic environmental hydrothermal energy system scheduling. Energy 131:165–178. https://doi.org/10.1016/j.energy.2017.05.013

Gao MY, Yang HL, Xiao QZ et al (2021) A novel method for carbon emission forecasting based on Gompertz’s law and fractional grey model: evidence from American industrial sector. Renewable Energy 181:803–819. https://doi.org/10.1016/j.renene.2021.09.072

Gao MY, Yang HL, Xiao QZ et al (2022) COVID-19 lockdowns and air quality: evidence from grey spatiotemporal forecasts. Socio-Economic Planning Sciences 101228. https://doi.org/10.1016/j.seps.2022.101228

Grossman GM, Krueger AB (1991) Environmental impacts of a North American free trade agreement

Grossman GM, Krueger AB (1995) Economic growth and the environment. Quart J Econ 112:353–378

Guo JL, Liu W, Tu LP et al (2021) Forecasting carbon dioxide emissions in BRICS countries by exponential cumulative grey model. Energy Rep 7:7238–7250. https://doi.org/10.1016/j.egyr.2021.10.075

Hosseini SM, Safioddin A, Shirmohammadi R (2019) Forecasting of CO₂ emissions in Iran based on time series and regression analysis. Energy Rep 5:619–631. https://doi.org/10.1016/j.egyr.2019.05.004

IBRD (2021) The Word Bank. https://www.worldbank.org/en/home

IEA (2021) International Energy Agency. https://www.iea.org/

Kennedy J, Eberhart R (1995) Particle swarm optimization. Proceedings of ICNN’95-International Conference on Neural Networks IEEE 4: 1942–1948

Kharbach M, Chfadi T (2017) CO₂ emissions in Moroccan road transport sector: Divisia. Cointegration, and EKC Analyses 35:396–401. https://doi.org/10.1016/j.ecs.2017.08.016

Kuznets S (1995) Economic growth and income inequality. American Economic Review 45:1–28

Le HP, Van DTB (2020) The energy consumption structure and African EMDEs’ sustainable development. Helioyn 6(4):e03822. https://doi.org/10.1016/j.helioyn.2020.e03822

Leerbeck K, Bacher P, Junker RG (2020) Short-term forecasting of CO₂ emission intensity in power grids by machine learning. Appl Energy 277(1):15527. https://doi.org/10.1016/j.apenergy.2020.115527

Li Y, Xiao J, Chen Y, Jiao L (2019) Evolving deep convolutional neural networks by quantum behaved particle swarm optimization with binary encoding for image classification. Neurocomputing 362:156–165. https://doi.org/10.1016/j.neucom.2019.07.026

Li YM, Dong HK, Lu SS (2021) Research on application of a hybrid heuristic algorithm in transportation carbon emission. Environ Sci Pollut Res 28:48610–48627. https://doi.org/10.1007/s11356-021-14079-y

Lin SF, Wang SY, Marinova D et al (2017) Impacts of urbanization and real economic development on CO₂ emissions in non-high income countries: empirical research based on the extended STIRPAT model. J Clean Prod 166(10):952–966. https://doi.org/10.1016/j.jclepro.2017.08.107

Liu DN, Xiao BW (2018) Can China achieve its carbon emission peaking? A scenario analysis based on STIRPAT and system dynamics model. Ecol Ind 93:647–657. https://doi.org/10.1016/j.ecolind.2018.05.049

Ofoolu-Adarkwa J, Xie NM, Javed SA (2020) Forecasting CO₂ emissions of China’s cement industry using a hybrid Verhulst-GM(1, N) model and emissions’ technical conversion. Renew Sustain Energy Rev 130:109945. https://doi.org/10.1016/j.rser.2020.109945

Pao HT, Fu HC, Tseng CL (2012) Forecasting of CO₂ emissions, energy consumption and economic growth in China using an improved grey model. Energy 40(1):400–409. https://doi.org/10.1016/j.energy.2012.01.037

Qiao WB, Lu HF, Zhou GF et al (2020) A hybrid algorithm for carbon dioxide emissions forecasting based on improved lion swarm optimizer. J Clean Prod 244(20):18612. https://doi.org/10.1016/j.jclepro.2020.125324

Rao CJ, Yan BJ (2020) Study on the interactive influence between economic growth and environmental pollution. Environ Sci Pollut Res 27(31):39442–39465. https://doi.org/10.1007/s11356-020-10017-6

Ren F, Long DH (2021) Carbon emission forecasting and scenario analysis in Guangdong Province based on optimized Fast Learning Network. J Clean Prod 317(1):128408. https://doi.org/10.1016/j.jclepro.2021.128408

Saidi K, Omri A (2020) Reducing CO₂ emissions in OECD countries: do renewable and nuclear energy matter. 126: 103425. https://doi.org/10.1016/j.jpucene.2020.103425

Sevgüll BM, Ağbuhut U (2021) Electricity production based forecasting of greenhouse gas emissions in Turkey with deep learning, support vector machine and artificial neural network algorithms. J Clean Prod 285(20):125324. https://doi.org/10.1016/j.jclepro.2020.125324

Song W, Cattani C, Chi C (2020) Multifractional Brownian motion and quantum-behaved particle swarm optimization for short term power load forecasting: An integrated approach. Energy 194:116847. https://doi.org/10.1016/j.energy.2019.116847

Sun W, Wang CF, Zhang CC (2017) Factor analysis and forecasting of CO₂ emissions in Hebei, using extreme learning machine based on particle swarm optimization. 162(20):1095–1101. https://doi.org/10.1016/j.jclepro.2017.06.016

Sutthichaimeethee P, Ariyasajakorn D (2018) Forecast of carbon dioxide emissions from energy consumption in industry sectors in Thailand. Environmental and Climate Technologies 22(1):107–117. https://doi.org/10.1007/s116847-018-0007

Tian X, Geng Y, Zhong SZ et al (2018) A bibliometric analysis on trends and characters of carbon emissions from transport sector. Transp Res Part d: Transp Environ 59:1–10. https://doi.org/10.1016/j.trd.2017.12.009

Turgut OE (2016) Hybrid Chaotic Quantum behaved Particle Swarm Optimization algorithm for thermal design of plate fin heat exchangers. Appl Math Model 40(1):50–69. https://doi.org/10.1016/j.apm.2015.05.003
Wang D, Nie R, Shi HY (2011) Scenario analysis of China’s primary energy demand and CO₂ emissions based on IPAT model. Energy Procedia 5:365–369. https://doi.org/10.1016/j.egypro.2011.03.062

Wang CJ, Wang F, Zhang XL (2017) Examining the driving factors of energy related carbon emissions using the extended STIRPAT model based on IPAT identity in Xinjiang. Renew Sustain Energy Rev 67:51–61. https://doi.org/10.1016/j.rser.2016.09.006

Wang Q, Li SY, Pisarenko Z (2020) Modeling carbon emission trajectory of China, US and India. J Clean Prod 258(10):120723. https://doi.org/10.1016/j.jclepro.2020.120723

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.