Heliyon 9 (2023) e21241

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Modeling and forecasting CO_2 emissions in China and its regions using a novel ARIMA-LSTM model☆

Tingxin Wen, Yazhou Liu *, Yun he Bai, Haoyuan Liu

College of Business Administration, Liao Ning Technical University, XingCheng, 125100, China

ARTICLE INFO

Keywords:
Carbon dioxide emissions
Factors analysis
ARIMA+LSTM
Forecast

ABSTRACT

Since China joined the WTO, its economy has experienced rapidly growth, resulting in significantly increase in fossil fuel consumption and corresponding rise in CO_2 emissions. Currently, China is the world's largest emitter of CO_2, the regional distribution is also extremely uneven. so, it is important to identify the factors influence CO_2 emissions in the three regions and predict future trends based on these factors. This paper proposes 14 carbon emission factors and uses the random forest feature ranking algorithm to rank the importance of these factors in three regions. The main factors affecting CO_2 emissions in each region are identified. Additionally, an ARIMA + LSTM carbon emission predict model based on the inverse error combination method is developed to address the linear and nonlinear relationships of carbon emission data. The findings suggest that the ARIMA + LSTM is more accurate in predicting the trend of CO_2 emissions in China. Moreover, the ARIMA + LSTM is employed to forecast the future CO_2 emission trends in China's east, central, and west regions, which can serve as a foundation for China's CO_2 emission reduction initiatives.

1. Introduction

In the era of economic globalization, issues related to CO_2 (carbon dioxide) emissions and climate change have emerged and attracted the attention of scholars. According to the Intergovernmental Panel on Climate Change, the quantity of CO_2 in the atmosphere is rising at a rate of 2 % per year. In 2014, global CO_2 emissions amounted to 36.14 billion tons, which is triple the amount emitted in 1960. The massive CO_2 emissions are aggravating the Global Warming process and powerful measures are necessary to maintain a sustainable global development [1].

China is now the world's biggest emitter of CO_2 (IEA, 2018). To deal with this issue, the Chinese government has implemented various pledges and initiatives. China ratified the Kyoto Protocol in 1998 and the Paris Climate Agreement in 2016. During the general debate of the 75th United Nations General Assembly on September 22, 2020, China announced its goals are to achieve peak CO_2 emissions by 2030 and carbon neutrality by 2060. To effectively decrease CO_2 emissions, the primary reasons for the increase in CO_2 emissions must be identified, forecasting China's future CO_2 emission tendencies based on these causes, and carry out suitable emission reduction measures.

☆ As the submitted manuscript has already been uploaded, no modification options are provided in the function bar. Among them, highlighting revisions made is the version before the modification, and manuscript is the final version.Due to the inability to modify the final manuscript, any questions raised by reference errors can be corrected during the proofreading process.

* Corresponding author.
E-mail address: lyz661155@163.com (Y. Liu).

https://doi.org/10.1016/j.heliyon.2023.e21241
Received 22 August 2023; Received in revised form 8 September 2023; Accepted 18 October 2023
Available online 24 October 2023

The paper introduces two main innovations: (1) the use of the random forest feature screening method for CO_2 emission impact factor analysis. This method effectively identifies the factors with the most influential factor on CO_2 emissions, improving the accuracy and interpret ability of the model. (2) The construction of an ARIMA + LSTM CO_2 emissions prediction model based on a weighted combination of ARIMA and LSTM. This model combines the advantages of both methods and provides more accurate predictions of carbon dioxide emissions. Additionally, the model predicts CO_2 emissions for China as a whole and its three regions (East, West, and Central), provide valuable insights for developing carbon reduction policies.

This article is organized into six parts. Section 2 reviews the factors and prediction methods that affect CO_2 emissions. Section 3 introduces the LSTM and ARIMA, as well as the integrated prediction weighting methods and model evaluation indicators. Section 4 describes the data collection, analysis and processing process in detail. In section 5, the prediction results of LR, BPNN, ARIMA, LSTM and ARIMA + LSTM are compared and analyzed. The conclusion of the study is that the ARIMA + LSTM is more precise in predicting CO_2 emissions. The model is used to predict CO_2 emissions from 2018 to 2025 in China as a whole and in three regions, east, west and central, to assess the efficacy of China's carbon reduction strategies. In section 6, credible proposals are presented based on the analysis of the carbon emission forecast results for China as a whole and for the three regions of east, west, and central. These recommendations are intended to support China in attaining its carbon emission reduction goals.

2. Literature review

2.1. Impacting factors of CO_2 emissions

In the last few years, the scholars have extensively investigated the factors affecting CO_2 emissions, focusing on the economy, energy consumption, population, urbanization, transportation, and technological progress [2]. analyzed 14 energy-related emission drivers and their provincial characteristics in the eastern and southern coastal regions of China using the log-averaged dies index approach. They discovered that there was a remarkable correlation between economical growth and CO_2 emissions even in the socio-economically developed regions of China [3]. studied the factors influencing CO2 emissions in China from 2006 to 2012 with production theory and data envelopment analysis. They demonstrated that economic activities is the primary reason for the increase CO_2 emissions in China [4]. have analyzed the intensity of CO_2 emissions based on alternative, fossil and renewable energy sources and investigated the economic growth on the basis of CO_2 emission intensity. They found that the intensity of CO_2 emissions from solid fuels had the greatest effect on the economic growth.

In terms of energy consumption [5], researched the correlation on energy consumption and economic growth for 30 mainland Chinese provinces from 1985 to 2007. They used the panel unit root method, the homogeneous panel co-integration method, and the panel-based dynamical OLS method to revisit the synergistic movement of energy consumption and economic growth. The research found that increased energy consumption, which includes oil, gas and electricity, contributed to increased CO_2 emissions [6]. [7] found that under the current energy consumption scenario, China is not expected to reach the peak of CO_2 emissions. Therefore, energy consumption must be restructured to achieve an earlier and less severe peak in CO_2 emissions.

In terms of population [8], used the panel auto-regressive distributed lag (ARDL) model framework and combined mean group (PMG) estimator to explore the association between demographic drivers, low-carbon technologies, and CO_2 emissions in 285 cities. The study found that population size and density can increase CO_2 emissions, while the quality of population and low-carbon innovations are significant elements in mitigating the CO_2 emission pressure in long term. Acheampong et al. (2019) conducted a research of China's carbon intensity based on economic growth, population size, technology research development, energy consumption, financial growth, trade liberalization, foreign investment, industrialization, and urbanization [9]. analyzes the determinants of CO_2 emissions in China by using extended STIRPAT model and finds that population size, energy intensity, GDP per person, the degree of urbanization, the coal consumption share, the secondary sector share and investment each have a positive effect on CO_2 emissions in a significant way. In particular, population size has the largest influence, while energy intensity has the smallest influence.

In terms of urbanization [10], studied the influencing factors of CO_2 emissions in five major regions of China by using the STIRPA model. The study shows that population, industrial structure, GDP per capita, level of technological development and urbanization have various effects on CO_2 emissions in different regions, but they are primarily the major affecting factors. The study finds that urbanization and GDP per capita are more influential on CO_2 emissions than other factors [11]. analyzed the urbanization process in Shandong province and found that rapid urbanization accompanied by high energy consumption has increased the province's CO_2 emissions significantly over the past 20 years [12]. used an autoregressive distributed lag (ARDL) constraint test to examine the cointegration of carbon emission data from 1972 to 2014. Following a causal analysis, the research found that urbanization increases CO_2 emissions in the feature [13]. assessed CO_2 emission efficiency based on panel data of 30 Chinese provinces from 2000 to 2016, which found a positive correlation between urbanization rate, economic development level, energy consumption structure and CO_2 emission efficiency. In addition, there is also a noticeable regional heterogeneity in the mechanism of urbanization level on regional CO_2 emissions [14].

In transportation [15], investigated the decoupling association of economic growth and transportation CO_2 emissions of Jiangsu province by using Tapio model. The study shows that transportation is an important affecting factor of CO_2 emissions in Jiangsu province [16]. has identified six primary factors affecting China's CO_2 emissions from the transportation industries and developed a CO_2 emission prediction model using support vector regression (SVR). The experimental results have shown that the CO_2 emissions in the transportation industry vary depending on the intensity of CO_2 emissions. To accelerate the achievement of peak CO_2 emissions in the transportation industries. it is essential to make changes to economic growth pattern and appropriately decrease economic development rate.

According to Refs. [17,18], technological progress has played a vital role in China's CO_2 emissions. The decrease of CO_2 emissions in different regions is influenced by the degree of technological progress and the types of technologies. These researches have studied the factors affecting CO_2 emissions in different perspectives, which are crucial to reduce CO_2 emissions. In this article, we have analyzed the factors that are responsible for CO_2 emissions, for example, economy, energy consumption, population, urbanization rate, transportation and technology. We summarize the factors affecting CO_2 emissions from four perspectives: economy, population, transportation, and technology, taking into account the features of CO_2 emissions in China's different regions.

2.2. The models of CO_2 emissions prediction

The determination of CO_2 emission factors is crucial to reduce emissions volume, and the investigation on carbon emission projection methods is indispensable. Recently, numerous scholars have studied the methods of CO_2 emission forecasting [19]. established CO_2 emission trends and total world CO_2 emissions for the top 25 countries from 1971 to 2007 b y using linear tendency analysis, and concluded that this method can forecast the CO_2 emissions of these 25 countries effectively [20]. delves into the association of urban population, energy consumption, economic growth and CO_2 emissions in BRICS countries from 2004 to 2010 through a multi-variate gray model, and proposes positive carbon reduction recommendations [21]. used a non-linear gray multivariate model to predict CO_2 emissions caused by fossil energy consumption in China and attained a better prediction accuracy than the traditional gray model [22]. modeled the correlation between CO_2 emissions and economic growth in China by using gray vehulst model based on the PSO algorithm and concluded that the correlation of CO_2 emissions and economic growth is inverted U-shaped, emissions are at a rapidly increase rate, and the government should take efficient initiatives to decrease CO_2 emissions. Although the above prediction model is relatively simple, has fewer parameters, easily trained, and has high forecasting accuracy, it is weak in extracting the nonlinear series of carbon emission data. To further improve the forecasting accuracy and use the data nonlinear series information fully, machine learning-based carbon emission forecasting methods have become a new research direction.

[23] predicted CO_2 emissions resulting from oil consumption in OPEC countries by utilizing hybrid cuckoo search algorithm and neural networks. They concluded that their prediction accuracy was higher than that of artificial neural networks [24]. predicted the carbon intensity of Brazil, Australia, India, China, and the United States by using artificial neural networks (ANN), the results showed that the accuracy of the artificial neural network predictions was higher [25]. introduced a hybrid algorithm to predict the CO2 emissions of individual countries in 2018–2025 based on an improved lion swarm optimizer [26]. has used 12 machine learning algorithms to forecast costs and CO_2 emissions in an energy-water integrated optimization model of buildings, which has achieved excellent forecasting results [27].have developed a fast learning network (FLN) prediction algorithm for forecasting CO_2 emissions in Guangdong Province for the years 2020–2060 based on the chicken flock algorithm. Their results show a significant improvement compared to the basic fast learning network. These machine learning algorithms show good fitting and generalization ability for non-linear and unfit data. But, these models have many training parameters, which can lead gradient disappearance and get stuck in local optima and over fitting.

Based on the literature reviewed, it is evident that the accuracy and stability of single linear or nonlinear models in predicting carbon dioxide emissions require improvement. To address this issue, combined prediction methods that consider both linear and nonlinear information within the data have emerged as a new research direction. Hybrid methods that have been successfully applied in various fields on the basis of linear econometric models and machine learning models [28,29]. However, in CO_2 emission prediction filed, there has been limited research on the combination of linear and no-linear forecasting models. the ARIMA model is a classical linear statistical model that is excellent at identifying linear relationships in time series. In contrast, the LSTM is great for handling nonlinear relationships in time series due to its unique gate structure [30]. has forecast China's main energy consumption by using an ARIMA model and found that the energy consumption growth rate at 2014–2020 showing a substantial increase, but is smaller than the start of the first decade of the 21st century [31]. used LSTM to predict China's CO_2 emissions with significantly improved accuracy compared to models such as GP (Gaussian Process Regression) and BPNN (Back Propagation Neural Network) [32]. predicted the air quality index effectively by using LSTM [33]. used an LSTM-ARIMA hybrid model to effectively forecast the future export volume of Indonesia, which supported the Indonesian government in formulating economic policies [34]. accurately and effectively predicts well production by combining ARIMA and LSTM models, which is critical for prolonging the life cycle of wells and enhancing the recovery of reservoirs.

To use of the time series data fully, in this article we used the error inverse method to combine the ARIMA and LSTM in a weighted manner. After comparing and analyzing the LR, BPNN, ARIMA as well as LSTM, we find that the ARIMA + LSTM has a better predict results.

3. Construction and evaluation of ARIMA + LSTM

3.1. ARIMA model

ARIMA model is a method of time series analysis, based on stochastic theory and was proposed by Box and Jenkins in 1970. Time series is a collection of time-varying stochastic variables that exhibit certain regularities that make it possible to predict future trends. the structure of the ARIMA model is shown in Eq. (1).

$$\Delta^d y_t = \theta_0 + \sum_{i=1}^{p} \varphi_i \Delta^d y_{t-1} + \varepsilon_t + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} \tag{1}$$

In Eq. (1). ε_t presents the series of y_t after d-differential transformation; ε_t presents the residual term at t time, it is a mutually independent white noise series and obeys a normal distribution with mean 0 and variance constant σ^2; $\Phi_i(i=1,2,...,p)$ and $\theta_j(j=1,2,...,q)$ are the parameters to be estimated, and p and q are the orders. Thus, the above model can be written as $ARIMA(p,d,q)$, ε_t is an process of $ARIMA(p,q)$, y_t is an $ARIMA(p,d,q)$ process. ARIMA is a linear model, Therefore, its ability is restricted to describe the non-linear features of the time series.

ARIMA modeling and forecasting includes four steps: (1) series smoothing process. If the series is non-stationary, it needs to satisfy the smoothness test by differential variation; (2) determining the p and q orders through the auto-correlation coefficients and partial autocorrelation coefficients; (3) evaluating the parameters of the model as well as testing them to determine whether the model is desirable; and (4) using the model with the appropriate parameters selected to make predictions.

3.2. Long short term memory networks(LSTM)

LSTM is a recurrent neural network with gate structure proposed by Ref. [35]. (Hochreiter et al., 1997). Compared with the traditional RNN, LSTM has three additional gates unit structure: input gate, forget gate and output gate. These gates effectively control the updating and discarding of data, overcoming the disadvantages of RNN, such as excessive influence of weights, gradient vanishing and exploding. Therefore, the LSTM can converge faster and increase the carbon emission prediction accuracy. Fig. 1 below is a diagrammatic representation of the LSTM (see Fig. 2).

$$\text{Input gate}: i_t = sigmoid(W_i x_t + W_i h_{t-1} + b_i) \tag{2}$$

$$\text{Forget gate}: f_t = sigmoid\left(W_f x_t + W_f h_{t-1} + b_f\right) \tag{3}$$

$$\text{Unit}: c_t = f_t \odot c_{t-1} + i_t \odot \tanh(W_c x_t + W_c h_{t-1} + b_c) \tag{4}$$

$$\text{Output gate}: o_t = sigmoid(W_o x_t + W_o h_{t-1} + b_o) \tag{5}$$

$$\text{Final output}: h_t = o_t \odot \tanh(c_t) \tag{6}$$

in Eqs. (2)–(6), W_i is input gate's weight matrix, b_i is input gate's bias term, W_f is input gate's weight matrix; b_f is forget gate's bias term; the cell state c_t is obtained by multiplying the cell c_{t-1} at the previous moment with the oblivion gate f_t , the input gate i_t controls how much information flows into the internal memory unit c_t at the current moment, the forgetting gate controls how much information from the previous moment internal memory unit c_{t-1} can be accumulated to the current moment internal memory unit c_t; $\odot$ denotes the element-wise product of the corresponding elements of two vectors.

3.3. ARIMA + LSTM model

Based on the review above, it is evident that ARIMA exhibits excellent linear prediction ability, while LSTM demonstrates more

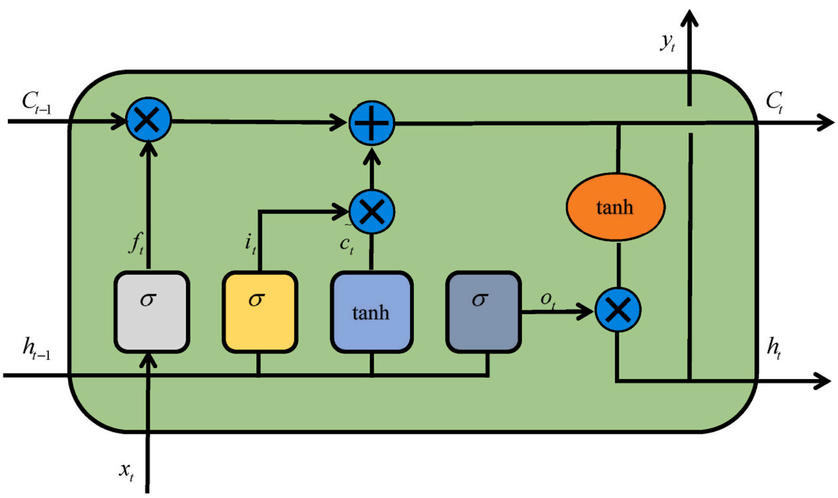


Fig. 1. Structure of LSTM

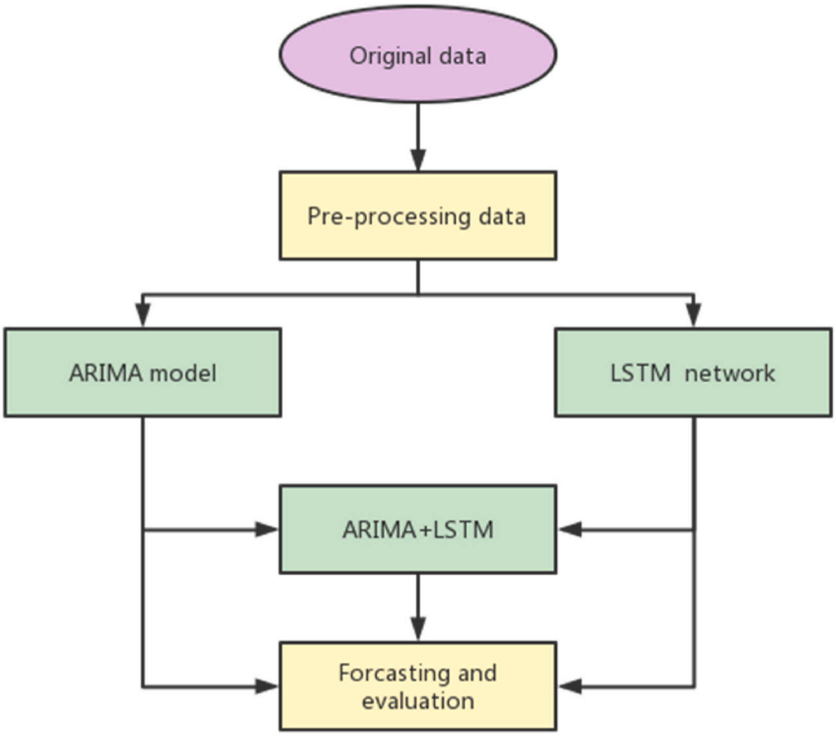


Fig. 2. Flowchart of ARIMA + LSTM model.

accurate prediction effects for time series data. Both ARIMA and LSTM can achieve more accurate prediction tasks. This paper analyzes the collected carbon emission data and identifies its obvious time series characteristics. To better utilize the time series information of carbon emission data, we assume that it contains both linear and nonlinear information. To verify the existence of linear and nonlinear information in carbon emission time series data and improve prediction accuracy, this paper employs the inverse error combination method to weigh the combination of ARIMA and LSTM. The weighted combination is presented in Eqs. (7)–(9) in this paper, which aims to enhance the prediction accuracy.

$$f_t = \omega_1 f_{1t} + \omega_2 f_{2t}, t = 1, 2, ..., n \tag{7}$$

$$\omega_1 = \frac{\lambda_2}{\lambda_1 + \lambda_2} \tag{8}$$

$$\omega_2 = \frac{\lambda_1}{\lambda_1 + \lambda_2} \tag{9}$$

in Eq. (7), f_{it} is the prediction value obtained from LSTM or ARIMA, ω_i is the weight coefficient, and λ_1, λ_2 are the prediction errors of LSTM and ARIMA, respectively. According to formula, it shown that the model with smaller errors will be assigned larger weight coefficients, so that the whole combined model error tends to be reduced and the prediction value error is smaller, which achieves the effect of improving the whole prediction accuracy [17].

The application process model described in the figure above can be divided into three steps: (1) Collect carbon emission data from each region and account for the CO_2 emissions of various regions. (2) Use ARIMA model to fit and forecast carbon emission data linearly and LSTM to fit and predict CO_2 emission data non-linearly. (3) Combine ARIMA + LSTM to fit and forecast CO_2 emissions data and evaluate the indicators Finally, forecasting the total CO_2 emissions of China and three regions by using the model.

3.4. Evaluation indicators

To evaluate the accuracy of different models, in this paper, we select root mean square error (RMSE), mean absolute error (MAE) and mean absolute percentage error (MAPE) as prediction evaluation indicators, The evaluation formula are Eqs. (10)–(12).

$$MAE = \frac{1}{m}\sum_{i=1}^{m}|(y_i - \widehat{y_i})| \tag{10}$$

$$RMSE = \sqrt{\frac{1}{m}\sum_{i=1}^{m}(y_i - \widehat{y_i})^2} \tag{11}$$

$$MPAE = \frac{100\%}{m}\sum_{i=1}^{m}\left|\frac{\widehat{y_i} - y_i}{y_i}\right| \tag{12}$$

in the above evaluation indicators, $\widehat{y_i}$ indicates the predicted value of CO_2 emissions in different years, y_i indicates the real value of CO_2 emissions in different years, and m indicates the number of time series; generally speaking, the smaller the value of MAE, RMSE and MAPE, the smaller the mistake between the real value and the forecast value, the more accurate the forecast result.

4. Regional CO_2 emissions accounting and impact factors analysis

4.1. CO_2 emission regional division

China's vast territory and differences in geographic location, resource endowment, and national regional development strategies result in significant regional differences in carbon dioxide emissions. CO_2 emissions studies use various regional classification methods, including the three-division method, the four-division method, and the eight-division method. The three-division method is the most widely used and better reflects the characteristics of China's regional CO_2 emissions. This paper selects the trichotomous method as the CO_2 emissions region classification method, combining the data structure published by the China Bureau of Statistics and the geographical location and economic development of each province in China. The country is classified into three carbon emission regions: East, West, and Central. Table 1 shows the specific regional division.

4.2. Accounting for CO_2 emissions

The data adopted in this article are all from public data published by the National Bureau of Statistics (NBS), and considering the possible impact of the novel coronavirus pneumonia epidemic on the data analysis thus leading to errors, for this reason, this paper selected data before 2018 for analysis, considering the completeness of the data published by the NBS. Data sources include the China Statistical Yearbook, Energy Statistical Yearbook, and Science and Technology Statistical Yearbook from 1997 to 2017.

Currently, there are several methods for calculating CO_2 emissions, including material balance algorithms, c-measurement methods, model estimation methods, and IPCC inventory methods. Of these methods, the IPCC inventory method is regarded as the most authoritative method for accounting for CO_2 emissions in the world (Liu et al., 2014). This method calculates CO_2 emissions based on energy consumption, using simple and mature accounting formulas, emission factors, and a large number of application examples. Therefore, this article adopts the IPCC inventory method to account for CO_2 emissions caused by energy consumption across China and three regions in the east, west, and central regions, without considering cross-regional or international transfer issues. The National Greenhouse Gas Inventory Guidelines are referred to for this purpose [1]. To calculate CO_2 emissions by fossil fuel combustion, this paper focuses on the combustion of eight major fossil fuels: raw coal, coke, diesel, kerosene, fuel oil, gasoline, crude oil, and natural gas. CO2 emissions are calculated as shown in Eq. (13).

$$EC = \sum_{i=1}^{8} C_i = \sum_{i=1}^{8} E_i \times F_i \times \frac{44}{12} = \sum_{i=1}^{8} E_i \times CF_i \times CC_i \times COF_i \times \frac{44}{12} \tag{13}$$

in Eq. (13), i expressed the energy type; EC expressed the total carbon dioxide emission of energy consumption; C_i expressed the carbon emission; E_i expressed the ratio of consumption to standard quantity; F_i expressed the carbon emission factor; CF_i expressed the low-level heat generation; CC_i expressed the carbon content; COF_i expressed the oxidation rate; $\frac{44}{12}$ expressed the carbon dioxide to carbon element molecular weight ratio; $CF_i \times CC_i \times COF_i$ expressed carbon emission factor; $CF_i \times CC_i \times COF_i \times \frac{44}{12}$ expressed carbon dioxide emission factor, the specific emission factors are listed in Table 2 below.

4.3. Analysis of CO_2 emission impact factors

This paper provides a comprehensive summary of the four factors that influence CO_2 emissions, namely population, economy, transportation, and technology. The relevant literature was analyzed, and the actual situation in the three major regions of East, Central, and West China was taken into account. The demographic factors considered were population size, urbanization rate, and

Table 1
Provinces of each region.

Regions	Provinces
Eastern	Beijing, Tianjin, Hebei, Liaoning, Shanghai, Jiangsu, Zhejiang, Fujian, Shandong, Guangdong, Hainan
Central	Jilin; Heilongjiang; Shanxi, Anhui, Jiangxi, Henan, Hunan; Hubei
Western	Xinjiang, Qinghai, Ningxia, Gansu, Inner Mongolia, Shaanxi, Chongqing, Sichuan, Guizhou, Yunnan, Guangxi,

Table 2
Carbon emission factors of fossil fuels.

Types	Low level heat generation (KJ/Kg)	Carbon content (KgC/GJ)	Carbon oxidation rate	Carbon emission factors ($t-C/t$) or ($t-C/m^3$)	CO_2 emission factors ($t-CO_2/t$) or ($t-CO_2/m^3$)
Raw Coal	20908	26.4	0.98	0.54	2.0
Coke	28435	29.5	0.93	0.78	2.9
Diesel	42653	20.3	0.98	0.85	3.1
Kerosene	43070	19.6	0.98	0.83	3.0
Fuel oil	41816	21.1	0.98	0.87	3.2
Gasoline	43070	18.9	0.98	0.80	2.9
Crude Oil	41816	20.1	0.98	0.82	3.0
Natural Gas	389310	15.3	0.99	5.90	21.6

consumption level. The economic factors included the GDP of primary, secondary, and tertiary industries, total import and export, and energy consumption structure. The transportation factors included road and railroad mileage, passenger and freight volume. The technological factors considered were energy intensity and technological progress. Table 3 provides a detailed explanation of each influencing factor.

To assess the impact of different factors on CO_2 emissions in three regions, this study employs the random forest feature filtering method to rank the factors based on their importance. Initially, the study identified 14 carbon emission factors as variables to be ranked, which were then input into the random forest algorithm. To minimize errors and avoid randomness, a five-fold cross-validation method was utilized, and the experiment was repeated 20 times. The importance of each CO_2 emissions factor was then averaged to determine its influence on the three regions. The results are presented in Table 4, which displays the importance of each carbon emission effect factor to CO_2 emissions in the eastern region, Table 5, which shows the contribution to the central region, and Table 6, which presents the contribution to the western region. According to the data in Tables 4 and it is evident that various factors influence CO_2 emissions in the whole China. Population accounts for 21.2 %, the economy accounts for 25 %, transportation has the largest share with 45.5 %, and technology has the smallest share with only 9.3 %. Specifically, freight volume, secondary industry, population size, railroad mileage, and road mileage are the top five factors that contribute the most, with 13.9 %, 10.2 %, 10.2 %, 9 %, and 8.6 %, respectively. These factors significantly impact China's CO_2 emissions. On the other hand, passenger volume, technological progress, and energy consumption structure have the smallest contribution, at 2.8 %, 1.9 %, and 1.3 %, respectively.

According to Tables 5 and it is evident that different factors in the eastern region. Specifically, demographic factors account for 22.4 %, economic factors for 36.1 %, transportation factors for 28.9 %, and technological factors for 12.6 %. Notably, the top five factors that significantly influence CO_2 emissions in the region are urbanization rate (13.6 %), secondary industry (12.4 %), railroad mileage (10.7 %), tertiary industry (10.4 %), and energy intensity (9.3 %). Conversely, technological progress, passenger traffic, and energy consumption structure have the least impact, accounting for only 3.3 %, 1.7 %, and 1.3 %, respectively.

According to Tables 6 and it is evident that various factors in the central region. Specifically, demographic factors account for 17.7 %, economic factors account for 32.3 %, transportation factors account for 44.9 %, and technological factors account for 5.1 %. The five top factors that significantly affect carbon dioxide emissions in the central region are railroad mileage (15.3 %), passenger traffic (11.5 %), road mileage (10.4 %), total import and export (8.7 %), and secondary industry (7.9 %). Additionally, population size, technological progress, and energy consumption structure have the least impact, accounting for only 2.6 %, 1.5 %, and 1.3 %, respectively.

According to Table 7, it is evident that different factors in the western region. Specifically, demographic factors account for 22.8 %, economic factors for 35.5 %, transportation factors for 33.4 %, and technological factors for 8.4 %. Among these factors, primary industry, road mileage, urbanization rate, freight volume, and railroad mileage are the five top influences to CO_2 emissions in the

Table 3
Explanation of impact factors.

Factors		Explanation	Unit
Population	Population size	Total population at year-end	10000 people
	Urbanization rate	Ratio of non-rural population to total population	Percentage
	Consumption level	Ratio of total personal consumption to average total population per year	Percentage
Economy	Primary industry	The Agriculture	Trillion yuan
	Secondary industry	The Manufacturing industry	Trillion yuan
	Tertiary industry	All industries other than primary and secondary industries	Trillion yuan
	Total import and export value	Total amount of goods actually entering and leaving China's borders	Million dollars
	Energy consumption structure	Ratio of coal consumption to total energy consumption	Percentage
Transportation	Road mileage	Total highway mileage at year-end	Kilometre
	Railway mileage	Total railway mileage at year-end	Kilometre
	Passenger	traffic Number of passengers transported throughout the year	10000 people
	Freight	traffic Weight of goods transported throughout the year	10000 tons
Technology	Energy intensity	Ratio of total energy consumption to GDP	Percentage
	Tech progress	Ratio of science and technology expenditure to GDP	Percentage

Table 4
Contribution of national CO_2 emissions impact factors.

Influencing factors	Contribution	Influencing factors	Contribution
Freight	13.90 %	Energy intensity	7.40 %
Secondary industry	10.20 %	Total import and export value	6.90 %
Population size	10.20 %	Urbanization rate	5.70 %
Railway mileage	9.00 %	Consumption level	5.30 %
Highway mileage	8.60 %	Passenger	2.80 %
Tertiary industry	8.40 %	Technological advancement	1.90 %
Primary industry	8.20 %	Energy consumption structure	1.30 %

Table 5
Contribution of CO_2 emissions impact factors in the eastern region.

Influencing factors	Contribution	Influencing factors	Contribution
Urbanization rate	13.60 %	Primary industry	6.10 %
Secondary industry	12.40 %	Total import and export value	5.90 %
Railway mileage	10.70 %	Population size	5.30 %
Tertiary industry	10.40 %	Consumption level	3.50 %
Energy intensity	9.30 %	Technological advancement	3.30 %
Freight	9.00 %	Passenger	1.70 %
Highway mileage	7.50 %	Energy consumption structure	1.30 %

Table 6
Contribution of CO_2 emissions impact factors in the central region.

Influencing factors	Contribution	Influencing factors	Contribution
Railway mileage	15.30 %	Freight	7.70 %
Passenger	11.50 %	Urbanization rate	7.30 %
Highway mileage	10.40 %	Primary industry	6.70 %
Total import and export value	8.70 %	Energy intensity	3.60 %
Secondary industry	7.90 %	Population size	2.60 %
Consumption level	7.80 %	Technological advancement	1.50 %
Tertiary industry	7.70 %	Energy consumption structure	1.30 %

Table 7
Contribution of CO_2 emissions impact factors in the western region.

Influencing Factors	Contribution	Influencing Factors	Contribution
Primary industry	11.50 %	Consumption level	6.70 %
Highway mileage	11.50 %	Energy intensity	6.40 %
Urbanization rate	10.80 %	Population size	5.30 %
Freight	10.50 %	Total import and export value	4.90 %
Railway mileage	9.70 %	Technological advancement	2.00 %
Secondary industry	9.30 %	Passenger	1.70 %
Tertiary industry	8.40 %	Energy consumption structure	1.40 %

western region, accounting for 11.5 %, 11.5 %, 10.8 %, 10.5 %, and 9.7 %, respectively. Conversely, technological progress, passenger volume, and energy consumption structure have the smallest impact, accounting for only 2 %, 1.7 %, and 1.4 %, respectively.

Fig. 3 illustrated the degree of effect of various factors on the three regions, in the eastern region, urbanization is the greatest impact factor, indicating that the rapid rise in urban population along with economic growth has led to a significant raise in CO_2 emissions. To address this issue, the government should take significant measures to control population size and establish a reasonable urban and rural spatial layout. Additionally, the eastern region has the most complex transportation network in China, with the highest mileage of roads and railroads, a complex intercity transportation network, and the largest number of private transportation. Therefore, this region should implement policies to reduce the proportion of fuel vehicles, develop energy-efficient public transportation systems, reduce private transportation, and promote the use of new energy transportation. It is essential to note that the eastern region has the highest energy intensity in the country, indicating a high dependence on traditional energy consumption. Therefore, the region should optimize its energy layout and increase the percentage of renewable energy sources, for example, hydrogen and wind energy. Technological advances have a relatively small impact on CO_2 emissions reduction; therefore, eastern region should improve the targeting of emission reduction technologies and increase efforts to decrease CO_2 emissions with precision.

For central region, transportation is the primary factor contributing to CO_2 emissions. To decrease these emissions, the region should optimize its transportation strategy, adjust the market structure for fuel and new energy vehicles, and provide financial and

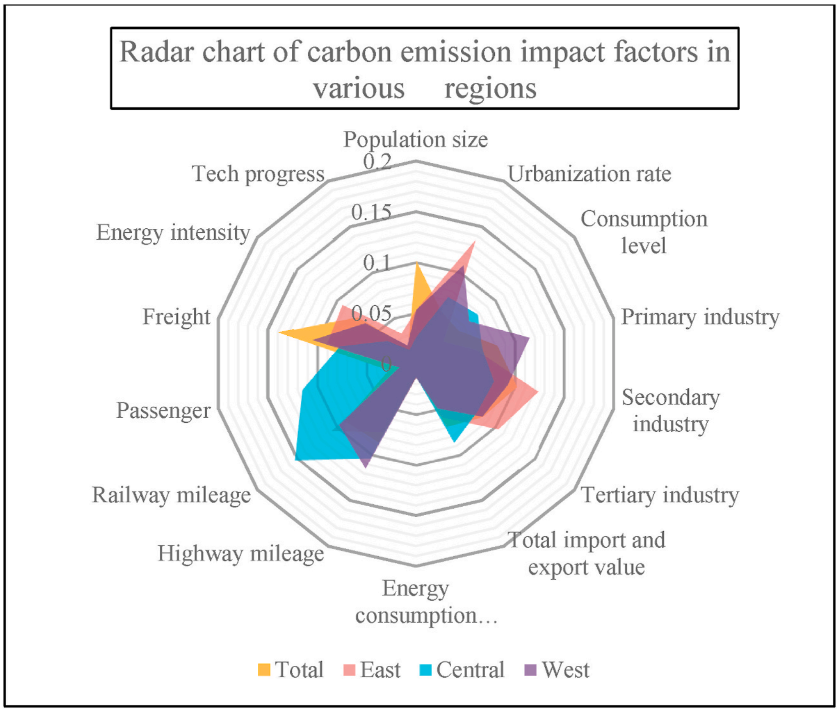


Fig. 3. Radar map of CO2 emissions impact factors.

policy support to new energy vehicle companies. The central region has a relatively large proportion of secondary and tertiary industries, and while plans to upgrade these industries have promoted economic development, they have also increased CO_2 emissions. To solve this problem, the region should build a reasonable industrial layout while maintaining economic growth, incentivizing developing new low-carbon and green enterprises, and increase the investment in the tertiary industry. Technological advances have a restricted influence on reducing CO_2 emissions in the central region; therefore, the region should put more effort into R&D, especially the developing and applying of carbon decrease technologies.

For western region, the primary industry has the largest influence on CO_2 emissions, next to the secondary and tertiary industries. Due to the region's reliance on traditional industries and the concentration of industries, relevant authorities should plan the industrial structure and increase the market share of tertiary industries. The government also should accelerate the development of renewable Energy industries, for example, photovoltaic and wind energy, according to the geographical characteristics of the region. Although technological advances have a restricted affect on reducing CO_2 emissions, investment in technological innovation and application should remain a priority.

In China, there are evident regional disparities in CO_2 emissions, and each region should develop and implement tailored emission reduction policies based on their specific circumstances. The primary and secondary industries have a close association with the three major regions, and a rational industrial structure and regional layout are critical to achieving carbon emission reduction goals. Transportation is a primary contributor to CO_2 emissions in China, mainly caused by the use of traditional energy sources. To reduce the use of fuel vehicles, relevant authorities should develop effective policies to promote the adoption of cleaner energy, increase the market share of new energy mobility Scooter and encourage the use of public transportation. Furthermore, technological advancements can also facilitate carbon emission reduction by encouraging enterprises and research institutions to create carbon-control and carbon-reduction products.

Table 8
Order *p* and *q* of ARIMA models.

Regions	p	d	q
Nationwide	1	1	0
Eastern	1	1	0
Central	1	1	0
Western	0	1	1

5. Prediction of CO_2 emissions

5.1. Prediction of CO_2 emissions by using ARIMA model

The previous section examined the factors that affect CO_2 emissions in different regions and identified the primary factors for each region. In this part, an ARIMA model was used to forecast the CO_2 emissions of three regions. At first, the CO_2 emission data of three regions were smoothed, and after the smoothing test, we determine that the ARIMA model is a smoothed time series with d-values of 1, 1, and 1 for the nationwide, eastern, and central regions, respectively, when the difference is of order 2. For the western region, the ARIMA model is a smoothed time series with a d-value of 1 when the difference is of order 1. Subsequently, by calculating the autocorrelation coefficient ACF and partially autocorrelation coefficient PACF values, we determine the orders of p and q of the ARIMA model for each region based on the principle of minimizing the AIC error measure, as presented in Table 8.

After determining the parameters of the $ARIMA(p,d,q)$ model for both the entire country and each region, CO_2 emissions were predicted for China as a whole and its three regions: East, Central, and West. In this process, 80 % of the data set was used for training and 20 % for testing. The results of each evaluation indicators were calculated and presented in Table 9. The MAPE values of the ARIMA model for predicting CO_2 emissions for China and its three regions were 0.0261, 0.0260, 0.0336, and 0.0411, respectively. These values shown that the ARIMA model has efficient linear prediction ability, and Changes in CO_2 emissions are reflected in the selected impact factors. However, the ability of ARIMA to express the nonlinear relationships within the data is unknown. Therefore, to test the nonlinear relationships in the carbon emission data, an LSTM model was used to forecast the CO_2 emissions for three regions (see Table 10).

5.2. Prediction of CO_2 emissions by using LSTM

In this paper, we input carbon emission data from 1997 to 2017 into an LSTM and utilize a sliding window forecasting method. Specifically, we use data points y_t, y_{t+1}, …, y_{t+n} to predict CO_2 emissions in year $y_{t+(n+1)}$. To prevent over fitting and maintain the model's generalization ability, we standardize the data. We employ minimum-maximum normalization to linearly transform the original data x, mapping the data values to the range $[0,1]$. The specific formula is shown in the following Eq. (14).

$$\widehat{x} = \frac{x - \min(x)}{\max(x) - \min(x)} \tag{14}$$

in this paper, various parameters such as the Dropout layer, optimizer, and regularization were configured to enhance the performance of the LSTM. Training times were set to four gradients of 100, 200, 300, 400, and 500. The data was split into 80 % for training and 20 % for testing, with 10 iterations performed each time. The training number corresponding to the minimum RMSE was selected. After several experiments, it was found that the error was minimized at a training number of 200, with the optimizer set to SGD and the regularization parameter L1 set to 0.001. Based on these results, carbon emission forecasts were carried out for China as a whole and for three regions (east, west, and east), with evaluation indicators MAPE of 0.0192, 0.0258, 0.0200, and 0.0148, respectively. These findings suggest that the LSTM model can effectively reflect the no-linear prediction correlation within the data, and that the selected influence factors can represent the change of CO_2 emissions. To further investigate the linear and nonlinear nature of carbon emission data, this paper combines ARIMA and LSTM to predict CO_2 emissions.

5.3. Prediction of CO_2 emissions by using ARIMA + LSTM

Although a single ARIMA or LSTM can provide more accurate forecasts, it can only utilize linear or nonlinear information from the data and cannot fully verify the correlation relationship within the CO_2 emissions data. By combining ARIMA and LSTM, we can better predict future carbon emission trends in China. Table 11 shows that the ARIMA + LSTM is more accurate than either model used alone.

5.4. Comparison of prediction results

Various models, including ARIMA, LSTM, ARIMA + LSTM, LR, and BPNN, are considered in this paper, to predict CO_2 emissions of total China and three regions. Fig. 4 and Table 12 present the predicted index values of CO_2 emissions for total China and for the three regions by different models. Results show that the ARIMA linear prediction model and the LSTM nonlinear prediction model were more accurate than the conventional LR and the BPNN. For example, for the eastern region, the ARIMA has an RMSE of 1.0094, MAE of 0.7821, and MAPE of 0.0260, while the LSTM model has an RMSE of 1.1582, MAE of 1.1427, and MAPE of 0.0258. after using the

Table 9
Prediction evaluation indicators values of ARIMA model.

Regions	RMSE	MAE	MAPE
Nationwide	2.1751	1.7083	0.0261
Eastern	1.0094	0.7821	0.0260
Central	0.7598	0.6202	0.0336
Western	0.7448	0.6066	0.0411

Table 10
Prediction evaluation indicators values of LSTM.

Regions	RMSE	MAE	MAPE
Nationwide	2.1082	1.8948	0.0192
Eastern	1.1582	1.1427	0.0258
Central	0.6328	0.5541	0.0200
Western	0.5205	0.4087	0.0148

Table 11
Prediction evaluation indicators values of ARIMA + LSTM model.

Regions	RMSE	MAE	MAPE
Nationwide	1.9842	1.8213	0.0186
Eastern	1.1019	1.0808	0.0243
Central	0.5774	0.5107	0.0185
Western	0.4966	0.3917	0.0146

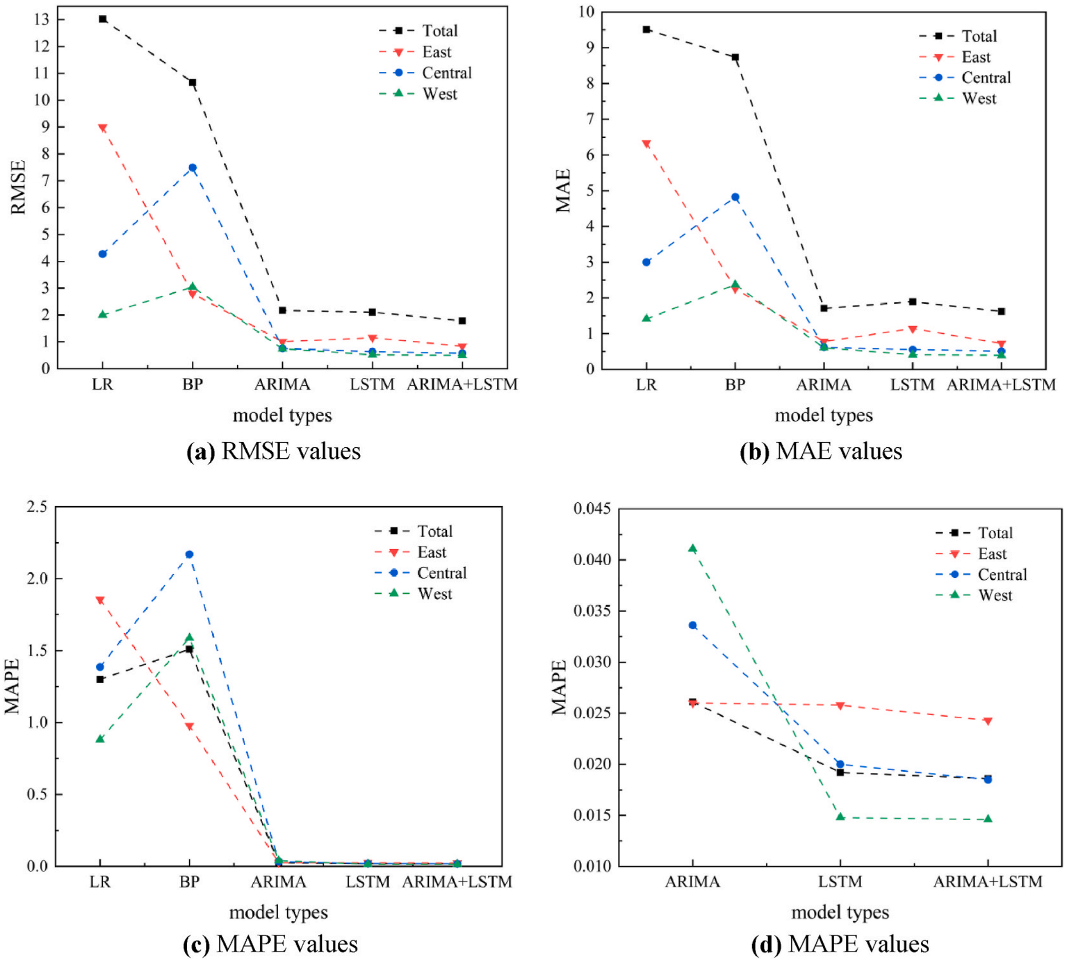


Fig. 4. Changes in evaluation indicators of different models.

combined ARIMA and LSTM, the model is further improved. Table 12 and Fig. 4 (d) show that the combined model is similar to the single model with the lowest value of the evaluation index, indicating that the inverse error combination method can effectively decrease the error caused by model singularity and the combined model is reasonable. This also validates the hypothesis that there are linear and non-linear relationships within the data.

Table 12
Comparison of different model prediction results.

Regions	Models	Evaluation indicators		
		RMSE	MAE	MAPE
Nationwide	LR	13.0206	9.5111	1.3010
	BPNN	10.6617	8.7357	1.5093
	ARIMA	2.1751	1.7083	0.0261
	LSTM	2.1082	1.8948	0.0192
	ARIMA + LSTM	1.7842	1.6213	0.0186
Eastern	LR	9.0006	6.3393	1.8566
	BPNN	2.8001	2.2576	0.9793
	ARIMA	1.0094	0.7821	0.0260
	LSTM	1.1582	1.1427	0.0258
	ARIMA + LSTM	0.8417	0.7241	0.0243
Central	LR	4.2760	2.9988	1.3875
	BPNN	7.4927	4.8271	2.1703
	ARIMA	0.7598	0.6202	0.0336
	LSTM	0.6328	0.5541	0.0200
	ARIMA + LSTM	0.5774	0.5107	0.0185
Western	LR	2.0035	1.4122	0.8814
	BPNN	3.0440	2.3719	1.5890
	ARIMA	0.7448	0.6066	0.0411
	LSTM	0.5205	0.4087	0.0148
	ARIMA + LSTM	0.4966	0.3917	0.0146

5.5. *Application of ARIMA + LSTM to predict future CO_2 emissions*

After comparison, it shown that the ARIMA + LSTM shows better forecast accuracy and is more suitable for predicting China's CO_2 emissions Table 13 presents the results of CO_2 emission projections for each region, while Fig. 5 illustrates the prospective tendency of CO_2 emissions in each region. Table 13 and Fig. 5 shown that China's CO_2 emissions are expected to continue to increase from 2018 to 2025, albeit at a reduced rate compared to the period from 2000 to 2015. This suggests that China's efforts to implement a range of carbon reduction measures are beginning to yield positive results.

By comparing (b), (c) and (d) in Fig. 5, it can be seen that CO_2 emissions in all each regions of China will increase in the coming years. Compared to the central and western regions, CO_2 emissions in the eastern region will grow more slowly, suggesting that the eastern region may earlier reaching the peak of CO_2 emissions. However, due to the industrial transfer from the eastern region and the economic development needs of the central and western regions, CO_2 emissions will continue to climb rapidly year by year thus making the mission to reduce CO_2 emissions in these regions more urgent.

General speaking, China's efforts to reduce CO_2 emissions have yielded positive results. However, there is still a need to introduce relevant policies, strengthen the monitoring of carbon emission reduction plans, and restructure primary, secondary and tertiary industries. Targeted carbon emission reduction measures should be implemented in different regions to promote economic development while reducing CO_2 emissions. Additionally, the promotion of a CO_2 emissions trading system and the upgrading of new environmental industries are necessary to promote a low-carbon economy.

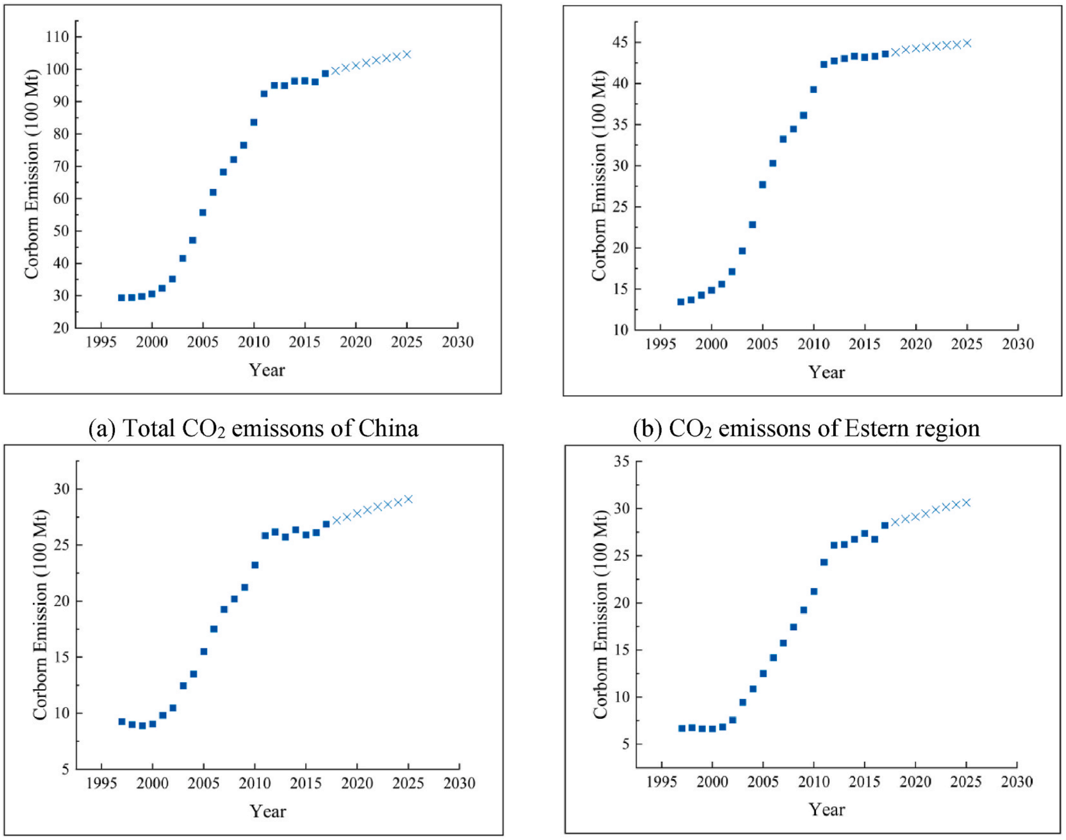
6. Conclusions

In the article, a new hybrid model, ARIMA + LSTM, is proposed to predict CO_2 emissions for total China and for the three regions We selected 14 factors that influence CO_2 emissions and used random forest method to calculate and rank their contributions. These factors are then fed into the ARIMA + LSTM hybrid model, which better than the LR, BPNN, ARIMA and LSTM. As a result, the new model can predict CO_2 emission trends more accurately. Our results show that this new model can be effectively applied in the field of CO_2 prediction.

After analyzing the influences of demographic, economy, transportation, and technological advances on CO_2 emissions in China as a whole and in the three regions, To summarize the following conclusions. Firstly, the production activities of the primary and secondary industries in different regions cause the largest amount of CO_2 emissions. Therefore, reducing their shares can effectively

Table 13
Forecasting carbon dioxide emissions by region, 2018–2025.

Regions	Years							
	2018	2019	2020	2021	2022	2023	2024	2025
Nationwide	99.56	100.5	101.2	102.01	102.83	103.43	103.94	104.66
Eastern	43.8	44.1	44.25	44.41	44.52	44.63	44.71	44.92
Central	27.2	27.51	27.83	28.14	28.42	28.63	28.81	29.1
Western	28.56	28.89	29.12	29.46	29.89	30.17	30.42	30.64



(a) Total CO_2 emissons of China

(b) CO_2 emissons of Estern region

(c) CO_2 emissons of central region

(d) CO_2 emissons of Western region

Fig. 5. China's national and three major regional CO_2 projections, 2018–2025.

reduce CO_2 emissions. Each region should continue to adjust its industrial structure according to its own situation and encourage the development of tertiary industries. However, the transfer of industries from the eastern region to the central and western regions will inevitably lead to an increase in CO_2 emissions. Therefore, the central government should integrate the actual situation of each region to optimize the industrial layout. Secondly, transportation factors have a great impact on all three regions, particularly the eastern and central regions. These two regions have a large share of the country's population, economy and transportation, resulting in a much larger number of fuel cars than in the western region. Therefore, China should promote the use of cleaner energy in all regions, encouraging the adoption of clean energy and new energy cars, and providing policies and funding support and assistance to new energy companies. Finally, the low influence of science and technology on the three regions, especially the western region, indicates that the importanceof science and technology in improving energy efficiency and emission control has not been realized. China should optimize the structure of scientific research and guide enterprises and scientific research institutions to move closer to CO_2 emission reduction.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:Tingxin Wen reports article publishing charges, statistical analysis, and writing assistance were provided by Liaoning Social Science Planning Foundation. Tingxin Wen reports article publishing charges, statistical analysis, and writing assistance were provided by National Natural Science Foundation of China.

Acknowledgements

This work is supported by the National Natural Science Foundation of China (Grant Number 71771111), Project of Liaoning Social Science Planning Foundation (Grant number L14BTJ004).

References

[1] 2006 IPCC Guidelines for National Greenhouse Gas Inventories - IPCC, n.d.
[2] C. Gao, Y. Liu, J. Jin, T. Wei, J. Zhang, L. Zhu, Driving forces in energy-related carbon dioxide emissions in east and south coastal China: commonality and variations, J. Clean. Prod. 135 (2016) 240–250, https://doi.org/10.1016/j.jclepro.2016.05.131.
[3] K. Du, C. Xie, X. Ouyang, A comparison of carbon dioxide (CO2) emission trends among provinces in China, Renew. Sustain. Energy Rev. 73 (2017) 19–25, https://doi.org/10.1016/j.rser.2017.01.102.
[4] T. Vujović, Z. Petković, M. Pavlović, S. Jović, Economic growth based in carbon dioxide emission intensity, Phys. Stat. Mech. Its Appl. 506 (2018) 179–185, https://doi.org/10.1016/j.physa.2018.04.074.
[5] L. Fei, S. Dong, L. Xue, Q. Liang, W. Yang, Energy consumption-economic growth relationship and carbon dioxide emissions in China, Energy Pol. 39 (2011) 568–574, https://doi.org/10.1016/j.enpol.2010.10.025.
[6] K. Munir, N. Riaz, Asymmetric impact of energy consumption on environmental degradation: evidence from Australia, China, and USA, Environ. Sci. Pollut. Res. 27 (2020) 11749–11759, https://doi.org/10.1007/s11356-020-07777-6.
[7] G. Xu, P. Schwarz, H. Yang, Adjusting energy consumption structure to achieve China's CO2 emissions peak, Renew. Sustain. Energy Rev. 122 (2020), 109737, https://doi.org/10.1016/j.rser.2020.109737.
[8] Q. Zhu, X. Peng, The impacts of population change on carbon emissions in China during 1978–2008, Environ. Impact Assess. Rev. 36 (2012) 1–8, https://doi.org/10.1016/j.eiar.2012.03.003.
[9] H. Zhao, J. Hu, F. Hao, H. Zhang, Determinants of carbon dioxide emissions and their peaking prospect: evidence from China, Front. Environ. Sci. 10 (2022), 913835, https://doi.org/10.3389/fenvs.2022.913835.
[10] H. Li, H. Mu, M. Zhang, S. Gui, Analysis of regional difference on impact factors of China's energy – related CO2 emissions, Energy 39 (2012) 319–326, https://doi.org/10.1016/j.energy.2012.01.008.
[11] L. Ren, W. Wang, J. Wang, R. Liu, Analysis of energy consumption and carbon emission during the urbanization of Shandong Province, China, J. Clean. Prod., Carbon Emissions Reduction: Policies, Technologies, Monitoring, Assessment and Modeling 103 (2015) 534–541, https://doi.org/10.1016/j.jclepro.2014.08.098.
[12] R. Ali, K. Bakhsh, M.A. Yasin, Impact of urbanization on CO2 emissions in emerging economy: evidence from Pakistan, Sustain. Cities Soc. 48 (2019), 101553, https://doi.org/10.1016/j.scs.2019.101553.
[13] V.J. Schweizer, K.L. Ebi, D.P. van Vuuren, H.D. Jacoby, K. Riahi, J. Strefler, K. Takahashi, B.J. van Ruijven, J.P. Weyant, Integrated climate-change assessment scenarios and carbon dioxide removal, One Earth 3 (2020) 166–172, https://doi.org/10.1016/j.oneear.2020.08.001.
[14] Y. Ding, F. Li, Examining the effects of urbanization and industrialization on carbon dioxide emission: evidence from China's provincial regions, Energy 125 (2017) 533–542, https://doi.org/10.1016/j.energy.2017.02.156.
[15] Y. Wang, T. Xie, S. Yang, Carbon emission and its decoupling research of transportation in Jiangsu Province, J. Clean. Prod. 142 (2017) 907–914, https://doi.org/10.1016/j.jclepro.2016.09.052.
[16] C. Zhu, M. Wang, W. Du, Prediction on peak values of carbon dioxide emissions from the Chinese transportation industry based on the SVR model and scenario analysis, J. Adv. Transport. (2020) 1–14, https://doi.org/10.1155/2020/8848149, 2020.
[17] J. Chen, M. Gao, S.K. Mangla, M. Song, J. Wen, Effects of technological changes on China's carbon emissions, Technol. Forecast. Soc. Change 153 (2020), 119938, https://doi.org/10.1016/j.techfore.2020.119938.
[18] W. Gu, X. Zhao, X. Yan, C. Wang, Q. Li, Energy technological progress, energy consumption, and CO2 emissions: empirical evidence from China, J. Clean. Prod. 236 (2019), 117666, https://doi.org/10.1016/j.jclepro.2019.117666.
[19] A.Ç. Köne, T. Büke, Forecasting of CO2 emissions from fuel combustion using trend analysis, Renew. Sustain. Energy Rev. 14 (2010) 2906–2915, https://doi.org/10.1016/j.rser.2010.06.006.
[20] L. Wu, S. Liu, D. Liu, Z. Fang, H. Xu, Modelling and forecasting CO 2 emissions in the BRICS (Brazil, Russia, India, China, and South Africa) countries using a novel multi-variable grey model, Energy 79 (2015) 489–495, https://doi.org/10.1016/j.energy.2014.11.052.
[21] Z.-X. Wang, D.-J. Ye, Forecasting Chinese carbon emissions from fossil energy consumption using non-linear grey multivariable models, J. Clean. Prod. 142 (2017) 600–612, https://doi.org/10.1016/j.jclepro.2016.08.067.
[22] Z.-X. Wang, Q. Li, Modelling the nonlinear relationship between CO2 emissions and economic growth using a PSO algorithm-based grey Verhulst model, J. Clean. Prod. 207 (2019) 214–224, https://doi.org/10.1016/j.jclepro.2018.10.010.
[23] H. Chiroma, S. Abdul-kareem, A. Khan, N.Mohd Nawi, A.Y. Gital, L. Shuib, A.I. Abubakar, M.Z. Rahman, T. Herawan, Global warming: predicting OPEC carbon dioxide emissions from petroleum consumption using neural network and hybrid cuckoo search algorithm, PLoS One 10 (2015), e0136140, https://doi.org/10.1371/journal.pone.0136140.
[24] A.O. Acheampong, E.B. Boateng, Modelling carbon emission intensity: application of artificial neural network, J. Clean. Prod. 225 (2019) 833–856, https://doi.org/10.1016/j.jclepro.2019.03.352.
[25] W. Qiao, H. Lu, G. Zhou, M. Azimi, Q. Yang, W. Tian, A hybrid algorithm for carbon dioxide emissions forecasting based on improved lion swarm optimizer, J. Clean. Prod. 244 (2020), 118612, https://doi.org/10.1016/j.jclepro.2019.118612.
[26] M. Emami Javanmard, S.F. Ghaderi, M. Hoseinzadeh, Data mining with 12 machine learning algorithms for predict costs and carbon dioxide emission in integrated energy-water optimization model in buildings, Energy Convers. Manag. 238 (2021), 114153, https://doi.org/10.1016/j.enconman.2021.114153.
[27] F. Ren, D. Long, Carbon emission forecasting and scenario analysis in Guangdong Province based on optimized Fast Learning Network, J. Clean. Prod. 317 (2021), 128408, https://doi.org/10.1016/j.jclepro.2021.128408.
[28] T.-T.-H. Phan, X.H. Nguyen, Combining statistical machine learning models with ARIMA for water level forecasting: the case of the Red river, Adv. Water Resour. 142 (2020), 103656, https://doi.org/10.1016/j.advwatres.2020.103656.
[29] A. Soy Temur, M. Akgun, G. Temur, Predicting housing sales in Turkey using arima, lstm and hybrid models, J. Bus. Econ. Manag. 20 (2019) 920–938, https://doi.org/10.3846/jbem.2019.10190.
[30] C. Yuan, S. Liu, Z. Fang, Comparison of China's primary energy consumption forecasting by using ARIMA (the autoregressive integrated moving average) model and GM(1,1) model, Energy 100 (2016) 384–390, https://doi.org/10.1016/j.energy.2016.02.001.
[31] Y. Huang, L. Shen, H. Liu, Grey relational analysis, principal component analysis and forecasting of carbon emissions based on long short-term memory in China, J. Clean. Prod. 209 (2019) 415–423, https://doi.org/10.1016/j.jclepro.2018.10.128.
[32] Q. Wu, H. Lin, Daily urban air quality index forecasting based on variational mode decomposition, sample entropy and LSTM neural network, Sustain. Cities Soc. 50 (2019), 101657, https://doi.org/10.1016/j.scs.2019.101657.
[33] E. Dave, A. Leonardo, M. Jeanice, N. Hanafiah, Forecasting Indonesia exports using a hybrid model ARIMA-LSTM, Procedia Comput. Sci. 179 (2021) 480–487, https://doi.org/10.1016/j.procs.2021.01.031.
[34] D. Fan, H. Sun, J. Yao, K. Zhang, X. Yan, Z. Sun, Well production forecasting based on ARIMA-LSTM model considering manual operations, Energy 220 (2021), 119708, https://doi.org/10.1016/j.energy.2020.119708.
[35] S. Hochreiter, J. Schmidhuber, Long short-term memory, Neural Comput. 9 (1997) 1735–1780, https://doi.org/10.1162/neco.1997.9.8.1735.