THE INTERNATIONAL SOCIETY FOR ENERGYTRANSITION STUDIES

ISETS Working Paper No. 2024001

A Study on the Efficiency of Sustainable Development of Green Economy in Chinese Cities in the Context of Carbon Neutrality Based on Deep Learning

Ying Zhang, Keyi Ju

Abstract: While the global economy is developing, the problem of environmental pollution is also becoming more and more serious. "Carbon peaking and carbon neutrality" are important goals for China's green and low-carbon economic development, which China has successively proposed. In order to reflect China's major decision and social responsibility to cope with global warming, it is necessary to incorporate the concept of "carbon peak and carbon neutral" into China's second centur goal, so as to truly demonstrate China's power and improve its international status. Therefore, in order to solve the problems in the process of achieving "peak carbon and c arbon neutrality", this paper analyzes China's green low - carbon economic developmen t model, mainly discusses the scientific planning of low-carbon environmental protection in China, considers how to reduce energy emissions and control the total amount of carbon emissions. In the context of deep learning, this paper proposes a long and short-term memory model (LSTM) for predicting carbon emissions. The model uses a genetic algorithm to search for optimality of the parameters of the LSTM network, and the LSTM network is supplemented with empirical modal decomposition and structural optimality search. Experiments on the carbon emission data set show that this method can accurately and effectively predict the carbon emission trend because the genetic algorithm ensures the goodness of the network and the empirical modal decomposition removes some noise.

www.isets.org

Citiation: Zhang, Y. and Ju K.Y., A Study on the Efficiency of Sustainable Development of Green Economy in Chinese Cities in the Context of Carbon Neutrality Based on Deep Learning. ISETS Working Paper, No.2024001, 2024.

A Study on the Efficiency of Sustainable Development of Green Economy in Chinese Cities in the Context of Carbon Neutrality Based on Deep Learning

Ying Zhang[1,2] Keyi Ju[1]

[1] School of Economics and Management, Jiangsu University of Science and Technology, Zhenjiang, Jiangsu, 212100, China
[2] Business School, Suzhou Institute of Technology, Jiangsu University of Science and Technology, Zhangjiagang, Jiangsu, 215600, China
Email：michellezhang8888@163.com

[1] School of Economics and Management, Jiangsu University of Science and Technology, Zhenjiang, Jiangsu, 212100, China
Email: jky1983@163.com

Abstract: While the global economy is developing, the problem of environmental pollution is also becoming more and more serious. "Carbon peaking and carbon neutrality" are important goals for China's green and low-carbon economic development, which China has successively proposed. In order to reflect China's major decision and social responsibility to cope with global warming, it is necessary to incorporate the concept of "carbon peak and carbon neutral" into China's second century goal, so as to truly demonstrate China's power and improve its international status. Therefore, in order to solve the problems in the process of achieving "peak carbon and carbon neutrality", this paper analyzes China's green low-carbon economic development model, mainly discusses the scientific planning of low-carbon environmental protection in China, considers how to reduce energy emissions and control the total amount of carbon emissions. In the context of deep learning, this paper proposes a long and short-term memory model (LSTM) for predicting carbon emissions. The model uses a genetic algorithm to search for optimality of the parameters of the LSTM network, and the LSTM network is supplemented with empirical modal decomposition and structural optimality search. Experiments on the carbon emission data set show that this method can accurately and effectively predict the carbon emission trend because the genetic algorithm ensures the goodness of the network and the empirical modal decomposition removes some noise.
Keywords: carbon neutral, deep learning, LSTM, genetic algorithm, empirical model decomposition

1 Introduction

In the process of building ecological civilization in China, considering the sustainable development of ecological civilization, the future development of future generations, from the perspective of human destiny, it is necessary to build a green home and promote the rapid development of China's green economy. China has carried out the development concept of green environment protection and put forward the goal of "carbon peak and carbon neutral". We have been steadily and steadily implementing the goal of "carbon peaking and carbon neutral" in our current work, contributing to the global environmental protection. Achieving the goal of "carbon peaking and carbon neutral" is China's civilizational commitment as a responsible country, which reflects a strong nationalism and harmonious society construction idea [1]. At present, China has embarked on the new journey of the second century goal, and needs to implement the development concept of "carbon peak and carbon neutral", strive to achieve the core development goal of low-carbon economy, and actively explore the green low-carbon road belonging to China's economic development.

In the process of economic development, China has always advocated an economic development strategy based on green and low-carbon, and our emission reduction targets have played a big role in this [2]. The "carbon peak" refers to the peak of carbon dioxide emissions at a certain point in China's economic development, after which carbon emissions will not grow, but will gradually decrease. This goal is determined by the scale of carbon dioxide emissions in relation to economic development, which means that after carbon emissions peak, the economy will continue to grow, but the overall carbon dioxide emissions will be controlled. The growth of the economy will be accompanied by a downward trend in carbon emissions. "Carbon neutral" refers to the balance between carbon dioxide emissions and the amount of carbon dioxide offset by afforestation, energy conservation and emission reduction within a certain period of time, to achieve positive and negative offset, to achieve relative "zero emissions" [3]. In other words, all industries are required to control carbon dioxide emissions, adhere to the concept of energy conservation and environmental protection, reduce energy consumption, achieve a balance of carbon emissions, and achieve "net zero emissions", thus reflecting China's green, low-carbon development ideas.

The problem of carbon emissions caused by economic development must be solved through better development. In the new journey of building a modernized socialist country, China must better coordinate economic development and carbon peak and carbon neutral goals [4]. To the greatest extent possible, we should avoid repeating the developed countries' detours of developing first and then reducing carbon, first high carbon and then low carbon, and come out with a new development path of building a modern country in a low carbon way [5].

To achieve the goal of carbon neutrality, we must fundamentally rely on the comprehensive green transformation of economic and social development, and promote the economy to embark on the road of green, low-carbon and cyclic development, which is the basic solution to the ecological problems of China's resources and environment, and the primary way to achieve the goal of carbon neutrality. To promote the realization of carbon peak and carbon neutral is essentially to promote the gradual "decoupling" of economic and social development and carbon emissions [6].

Achieving the new goal of carbon peaking and carbon neutral, which is unprecedented in human history, cannot be achieved without greater scientific and technological innovation and policy innovation. We need to organize national major special projects to carry out research and development for the basic theories and key technologies of deep carbon reduction, such as large-scale energy storage, hydrogen energy for steel making, fuel cells, carbon dioxide capture and storage, and carbon dioxide chemical industry [7]. Accelerate the promotion of new energy-saving and low-carbon technologies such as near-zero energy buildings, electric vehicles, heat pumps for heating, and industrial waste heat for heating. Implement new energy-saving and low-carbon models such as renewable energy building integration, shared transportation, and industry-city integration. Promote the accelerated development of new online economies such as video conferencing, telemedicine, and network education. Accelerate the use of new technologies, new models and new business models to promote energy conservation and carbon reduction. To adhere to the market-oriented, green credit, green bonds, green insurance as the main body, play a good role in the National Green Development Fund. Explore new ways to promote comprehensive demonstration and promotion of energy-saving and low-carbon technologies in a concentrated manner in parks, business districts and communities. Promote the deep integration and mutual benefit of green industry and green finance. Explore the application of information technology such as big data, cloud computing, data crawler, block chain, digital twin, etc. in carbon emission source locking, carbon emission data analysis, carbon emission supervision, carbon emission prediction and early warning and other scenarios to improve digital carbon reduction capability. Accelerate the improvement of price, taxation, finance and other economic policies conducive to green low-carbon development, and promote the innovative development of service models such as contract energy management, third-party management of environmental pollution and environmental trusteeship. Accelerate the construction of a national energy-use rights and carbon emissions trading market, and promote the formation of market expectations conducive to carbon peaking and carbon neutrality [8].

In summary, the conflict between the growth of carbon emissions and environmental protection is becoming more and more serious. The use of scientific and effective forecasting methods to detect carbon emissions can help protect the environment as well as prevent the blind development of the economy from causing irreversible damage to the ecology and affecting the survival and development of future generations. Every penny spent on development to protect the environment can bring greater economic benefits in the future. Predicting carbon emissions can provide technical support for a sustainable development path and provide a scientific basis for the development of emission reduction plans [9]. In the current context of the continuous development of big data technology, various technologies have emerged, and also bring many new ideas and methods for data prediction work. Deep learning is a branch of machine learning, and what machine learning can accomplish whether complex or simple learning work, deep learning can accomplish or even do better, and the learning ability of deep learning is proportional to the depth of the model.

In this paper, a deep learning model is used to study the daily carbon emission data in China and to predict them from the perspective of time series. The main body of the model is a long short-term memory (LSTM) model, which is able to fully utilize and learn from the data, obtain the characteristics of the data and make predictions. However, this model has many parameters to be adjusted, so each time the model is constructed, it is necessary to find the optimal parameters through a large number of experiments, and the obtained parameters are not always optimal. The combination of genetic algorithm (GA) and LSTM model enables the optimization of the parameters in the LSTM model, thus eliminating the need for manual tuning. The above two algorithms can learn the data to a certain extent, but the noise of the model itself will also interfere with the learning effect of the model, so this paper introduces empirical modal decomposition (EMD), which is combined with the GA-LSTM model. of data on the prediction effect of the model.

2 Related Works

2.1 Current status of green economy development in Chinese cities

Climate change is related to the survival and development of all human beings. Scientific research and observation data show that the global climate is undergoing a change with warming as the main feature in the last hundred years. Since the industrial revolution, human activities, especially the industrialization of developed countries, have emitted large amounts of greenhouse gases, which are the main factors of the current global climate change [10]. At present, global climate change has led to accelerated melting of glacial snow, rising sea level, imbalance in water resources distribution, frequent extreme weather, direct threats to agricultural production and biodiversity, and frequent occurrence of catastrophic climate events, which have increasingly profound impacts on human beings. The adverse effects of global climate change will further increase in the future, and climate change has become the greatest challenge to the sustainable development of the international community.

In 2020, natural disasters occurred in many places around the world, all related to climate change. The spread of the New Crown pneumonia epidemic has triggered a profound rethinking of the relationship between humans and nature, forcing us to refocus on global policies to address climate change [11]. The Newcastle pneumonia epidemic has reaffirmed the fragility of the balance of nature. Zoonotic diseases have become the latest concern worldwide. Ecological construction and environmental protection have become a topic of importance to countries around the world today, and China is actively responding to the call to develop a low-carbon economy. In the background of the country's strong advocacy of environmental protection, China has carried out a series of energy-saving and emission reduction work to promote sustainable social development. Meanwhile, as people's living standards continue to improve and their awareness of the importance of ecological environment deepens, the concept of "green water and green mountains are the silver mountain of gold, and green cities are the highest standard on earth" is accepted and recognized by more and more people, and the rapid economic development also brings a series of environmental problems [12]. For example, haze, water pollution and so on. These phenomena have seriously affected people's quality of life and health, and even threaten the sustainable and stable development of human society. Therefore, green low-carbon economy has become one of the essential topics for China to achieve the goal of building a well-off society in all aspects.

In China, the energy structure is mainly disposable energy sources such as coal and oil. With the development of the economy and the improvement of people's living standards and the enhancement of the awareness of sustainable development, people's demand for clean energy has gradually increased [13]. At the same time, the traditional fossil fuel production and use methods have brought about serious pollution and environmental problems, and it has become an inevitable trend to develop and use new technologies in large quantities to save energy and reduce emissions. This will promote the green industrial revolution, provide strong support for low carbonization, and to a certain extent ease the pressure of protecting the ecological environment and sustainable socio-economic development under a series of severe situations such as resource scarcity and environmental degradation in China. At present, China's industrial development is rapid and the demand for energy is increasing. As China's economy and society continue to grow rapidly, China's renewable energy resources consumption ranks first in the world [14]. In 2013, the national available coal was about 500 million tons and 600 million cubic meters of natural gas; electricity accounted for about 76% of the amount used for primary electrical equipment; and coal required for wind power generation accounted for nearly 30%-40% of the total installed capacity. Therefore, there is an inextricable link between energy conservation and environmental protection, and the development of green industry is particularly important in the context of low carbonization [15]. China is a large population and small resource country with scarce resources per capita. Therefore, the development of low-carbon economy is an inevitable choice for China to achieve sustainable development, and will also promote socio-economic and ecological civilization to a higher level. Green energy and industry have the same attributes: first, no pollutants are produced in the production process; second, less energy consumption, less pollution and other characteristics in line with the world trend of environmental protection trends [16]. This is conducive to improving the quality of the environment as well as reducing the harm or damage caused to human survival.

The development of green agriculture can not only promote the further optimization of China's agricultural production structure, but also help to improve the utilization of resources and save energy. First of all, we start from the "three rural" problem [17]. Focus on the green development of fossil energy, low-carbon utilization, pollution reduction and carbon innovation, focusing on strengthening the key technology development of source emission reduction such as multi-energy complementary coupling, low-carbon construction materials, low-carbon industrial raw materials, low-fluorine raw materials, etc.; strengthening the key technology development of process emission reduction such as integrated coupling of low-carbon technologies across the whole industrial chain and industries, low-carbon industrial process reengineering, and efficiency improvement in key areas; strengthening the synergy of pollution reduction and carbon reduction, synergistic the key technology development for end-of-pipe emission reduction such as treatment and ecological recycling, transportation storage and non-CO2 greenhouse gas emission reduction [18]. During the period of rapid economic growth in China, it is necessary to solve the outstanding conflicts such as low income of farmers, insufficient food production and serious environmental pollution in order to achieve sustainable development. Secondly, we should make more efforts to protect the ecological environment from damage and build ecological agriculture. Finally, to establish a sound market system of green agricultural products and related information service platform, to provide consumers with high-quality, convenient supply of green products. Green agriculture in the process of development in China, in order to achieve the "three rural" problem, we must first solve the problem of unreasonable production structure of agricultural products.

2.2 The state of the art of deep learning-based air quality monitoring

Since the 20th century, due to the accelerated industrialization of mankind, steam engines and coal were widely used in industrial production and life, releasing large amounts of soot and toxic gases. However, people neglected environmental protection issues, resulting in a series of major air pollution incidents, such as the London smog incident in the United Kingdom, the leakage of highly toxic raw materials in Bhopal, India, and air pollution in the Maas Valley industrial zone in Belgium, which caused many deaths and induced serious diseases. Since then, the United States, the United Kingdom and other developed countries began to actively seek ways to prevent and control air pollution [19]. By the 1970s, a large number of prediction tools were used in air quality modeling, while prediction research in China started in the late 1980s [20].

Artificial neural networks are self-learning, self-organizing, and self-adaptive, and possess powerful nonlinear function fitting and fault tolerance capabilities, making them a powerful way to deal with nonlinear data. With the dramatic increase in computer storage capacity and computing power in recent years, plus the emergence of deep learning, more and more research has begun to use neural networks to solve practical application problems, and air quality prediction with deep learning methods has attracted the attention of many researchers [21]. The fully connected neural network-based model uses independent distributed subnetworks composed of multiple layers of neural networks so as to fuse multiple sources of data and extract the potential factors affecting air quality, while taking into account the air quality of neighboring regions by means of spatial transformation.

Although computer technology and the technology of storing data are constantly evolving, with the advent of the information age and the increasing expansion of data, the problem of research and processing of massive data not only has an impact on human production and life, but also accelerates a new round of innovation and development of data science [22]. The multi-layer structure of deep learning is used to abstract complex data more deeply, so that certain intrinsic laws can be found in the huge data. Therefore, when using a large amount of data to learn certain features, the study of deep learning models is more capable of revealing the essential information inside the data [23]. Deep learning has evolved from traditional machine learning. The research of deep learning has made new advances in both models and algorithms, which has not only led to innovations in machine learning in research but also promoted the rapid development of artificial intelligence. The shortcomings of shallow networks are addressed. Innovations have been made on top of the traditional neural networks, and the concept of deep networks has been proposed, thus pushing deep learning into the research boom. Deep learning is to process complex information from the outside world by simulating the human brain and building it artificially to extract deeper features so that machines can solve complex problems similar to the human brain and dig out information from the inside of the data [24].

Deep learning is also widely used in air quality prediction due to its powerful nonlinear fitting ability. Better prediction results can be obtained through deep learning. Chen et al [25] proposed a deep learning-based prediction model, which uses correlation analysis to find the most appropriate set of features and is used for air quality prediction. Xu et al [26] embedded feature selection and semi-supervised learning in different layers of deep learning networks to integrate interpolation, prediction and feature analysis of air quality into one model. Yang et al [27] proposed a hybrid model combining complementary integrated empirical modal decomposition, improved cuckoo search and differential evolution algorithms and Elman neural network for air quality prediction. Saini et al [28] used a recurrent neural network to extract information from the input sequence and measure the influence of other stations to obtain prediction results using a fully connected layer. Du et al [29] considered the spatial correlation of air pollutants and used a spatial transformation component to process sparse air quality data with a neural distribution structure to fuse data affecting air quality from multiple cities. Zhang et al [30-31] proposed a spatio-temporal deep learning-based air quality prediction method that extracts inherent air quality features with a stacked autoencoder model and trains them in a greedy layer-by-layer manner, which can predict air quality at multiple stations simultaneously and shows the temporal stability of seasons.

For air pollutant emissions, the prediction of carbon emissions based on the traditional time series analysis method has certain prediction effect. However, compared with machine learning methods, it is slightly inadequate, and the use of data is not sufficient. Machine learning and deep learning can solve this problem well when making predictions, and are the mainstream development direction in data analysis at present. So based on this, this paper explores the use of machine learning and deep learning based methods for prediction research of carbon emission problems.

3 Algorithm Design

3.1 Long and short-term memory networks and genetic algorithms

Long and short-term memory neural networks are derived from a variant of recurrent neural networks. Since recurrent neural networks can update slowly during gradient descent gradient disappearance occurs, resulting in not getting the global optimum. Gradient disappearance is a situation where the network fails to learn or learns slowly due to slow updating or stagnation of weights during gradient descent. The cause of gradient disappearance is that when the absolute value of the gradient at each point in time is less than 1, the product of the two is approaching 0, which causes the update of the weights to stagnate. When the absolute value of the gradient at each time point is greater than 1, the product of the two will keep increasing, and the resulting oscillation in the updating of the weights will occur, which is the gradient explosion. Moreover, in deep learning, the number of network layers is often very large, which increases the rate of gradient convergence to 0 or gradient explosion, which largely limits the training effect of recurrent neural networks. Figure 1 shows the basic structure of LSTM.

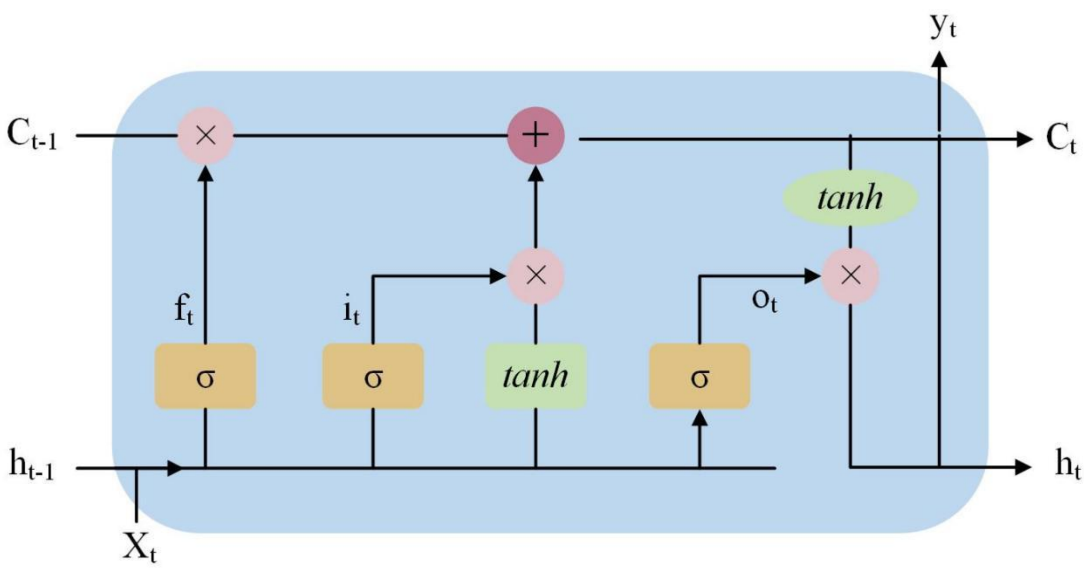

Figure 1 LSTM model structure

The genetic algorithm searches the hypothesis space mainly by encoding, fitness selection, genetic variation, and control parameters, and the comparison between the magnitudes of the fitness function values determines the best hypothesis, i.e., the optimal solution. Encoding is the transformation of the solution in the solution space of the parameters to be estimated into binary codes 0 and 1, and the feasible solution thus becomes a chromosome of the individuals in the population. Genetic variation is to mimic the population reproduction and variation in nature, and after the selection of the fitness function in the initialized population, the more superior parent individuals are retained, and each superior parent individual is copied, and the crossover of their codes and random inversions are made with a certain probability to complete the genetic and variation operations of the individuals to perform the parameter search. The genetic inheritance is to select the best existing individuals in the population. The variation is to ensure that the search gets better individuals that do not exist in the population. Without variation, it is difficult for the algorithm to continue the search for superiority. Variation breaks the stability of the population and allows the search to continue downward, thus filtering out the more adaptive individuals.

3.2 Empirical modal decomposition

The empirical mode decomposition method, as the name implies, is to decompose the data according to certain criteria, and the components obtained by different decomposition criteria are different. The following is a brief explanation of the method: The Intrinsic Mode Function (IMF). Almost all signals that we encounter in daily life can be represented by their combinations. By setting the decomposition criteria, the intrinsic mode function can be extracted from the signal, and the differences in the characteristic scales and frequencies of the constituent intrinsic mode functions cause the differences of each intrinsic mode function. So the eigenmode functions obtained from the decomposition are unique and symmetric based on the upper and lower X-axis, so some of the eigenmode functions show linearity and some show nonlinearity, and the complex signals are generated based on the combination of different eigenmode functions.

Due to the complexity of the data in reality, it is often difficult for the data itself to meet the restrictions of the eigenmodal function. The empirical modal decomposition method is needed to decompose the signal so that it can meet the requirements of the eigenmodal function and facilitate its analysis on a case-by-case basis. The steps of empirical modal decomposition depend mainly on the transformation of the data itself, and the main operation procedure is as follows.

Step 1. Use the cubic function curve through the extreme value points of the original signal $r_0(t) = S(t)$ to get the upper and lower envelopes of this signal. The upper envelope is generated by the function of the local maxima. The lower envelope is generated by the function of the local minima.

Step 2. After the first step is completed, the upper and lower envelopes are obtained, and the mean values of the upper and lower envelopes at each moment are calculated, which are defined as the mean envelope $m_i(t)$ by the function of the mean value.

Step 3. Subtract the mean envelope $m_i(t)$ from the original signal S_t to get the intermediate signal curve $l_i(t)$.

Step 4. Each eigenmodal function has to meet the above two constraints, so the two conditions need to be verified for the intermediate signal. If the conditions are satisfied, an eigenmode function $imf_i(t)$ is successfully separated, and the original signal is used to remove the $imf_i(t)$. That is, subtract this eigenmode function and get the combination of the remaining other eigenmode components:$r_i(t)$. Steps 1 to 4 are used for $r_i(t)$, and if the intermediate signal obtained does not satisfy the two constraints of the eigenmodal function, step 5 is performed.

Step 5. If the signal does not satisfy the two constraints, it means that the eigenmode function component of the signal is not fully refined. Next, redefine $r_i(t)$ as the original signal. Continue to fit the upper and lower envelopes using the cubic function, calculate the mean value for each moment of the upper and lower envelopes, and fit the mean envelope $m_{i+1}(t)$ according to the mean value.

Step 6. Subtract the mean envelope $m_{i+1}(t)$ from $r_i(t)$ to obtain the intermediate signal $l_{i+1}(t)$. Check whether the signal satisfies the two constraints of the eigenmodular function. If they are satisfied, the signal is an eigenmodular function $imf_i(t)$. If the two constraints are not satisfied, proceed to steps 5 and 6 until all the eigenmodes are decomposed or the stopping condition is reached. For the stopping condition, the standard deviation s_d can be used to control:

$$S_d = \sum_{t=1}^{n} [\frac{|l_{j-1}(t)-l_j(t)|}{l_j^2(t)}] \tag{1}$$

Where $l_j(t)$ is the modal component to be determined and s_d generally takes a value between 0.2 and 0.3. A complete original signal can be obtained about the representation of the original signal $l(t)$ by the above decomposition:

$$l(t) = \sum_{i=1}^{n} imf_i(t) + r_n(t) \tag{2}$$

The i th eigenmode function obtained after decomposition can be represented by $imf_i(t)$, and after decomposing the original signal to obtain all IMF components, the remaining components are represented by $r_n(t)$. Figure 2 illustrates the steps of the EMD method.

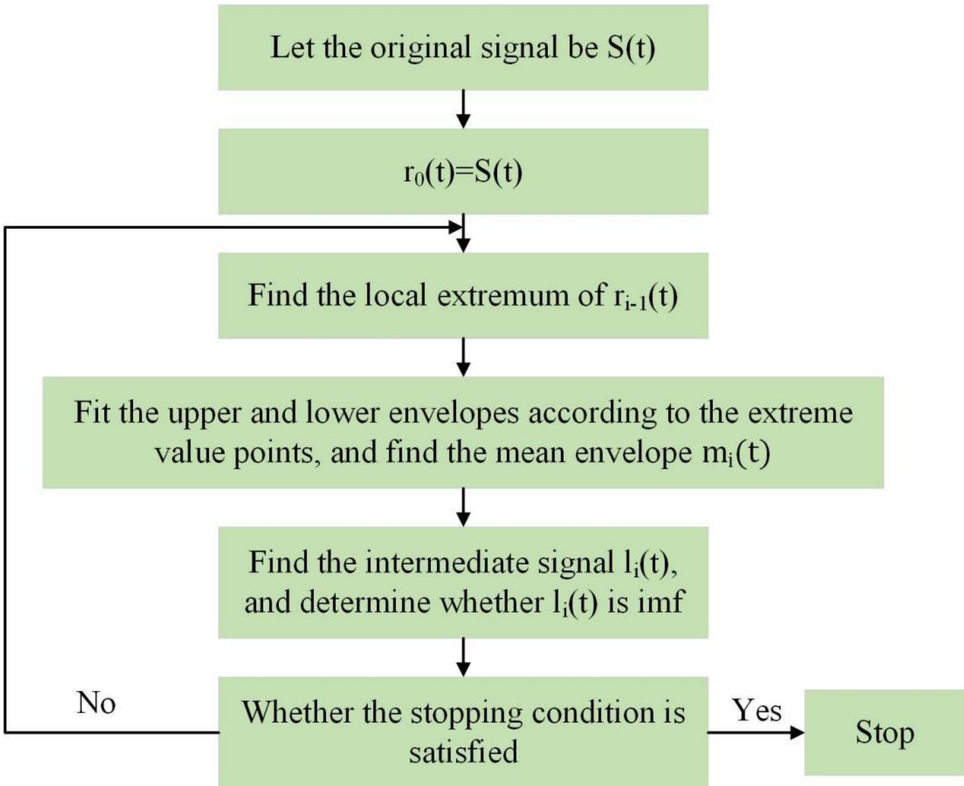

Figure 2 Flow chart of EMD method

The empirical modal decomposition method is adaptive in that the signal screening process is directly dependent on its own transformation, and there is no need to select the basis function in advance. The transformation is simple and easy to understand. The original signal can be decomposed to obtain a finite number of eigenmode function components, and the eigenmode functions obtained after the decomposition are different due to different eigenscale parameters. Empirical modal decomposition can also be understood as a data filtering process. By decomposing the data, the frequencies of each eigenmode component are arranged from high to low, i.e., the original signal keeps subtracting the low frequency components to obtain the high frequency components, so the empirical modal decomposition method itself is also an adaptive high-pass filter.

3.3 Fusion Algorithm

The empirical modal decomposition theory is able to decompose nonlinear and non-stationary time series data into smoother small signals, and the decomposed small signals can have the same properties as the original data. Therefore, the same model as the original data can be used in the prediction of the components. Carbon emission data are very volatile and non-linear, and can be decomposed using empirical modal decomposition. A genetic algorithm is used to find the optimal parameters of the LSTN model structure for each eigenmodal component. Finally, the test set data of each eigenmodal function component is predicted using the LSTM model after the optimization search. The above steps are repeated for each component, and the predicted data for each eigenmodal function component are summed to obtain the final predicted data. The modeling steps of the EMD-GA-LSTM model are shown in Figure 3.

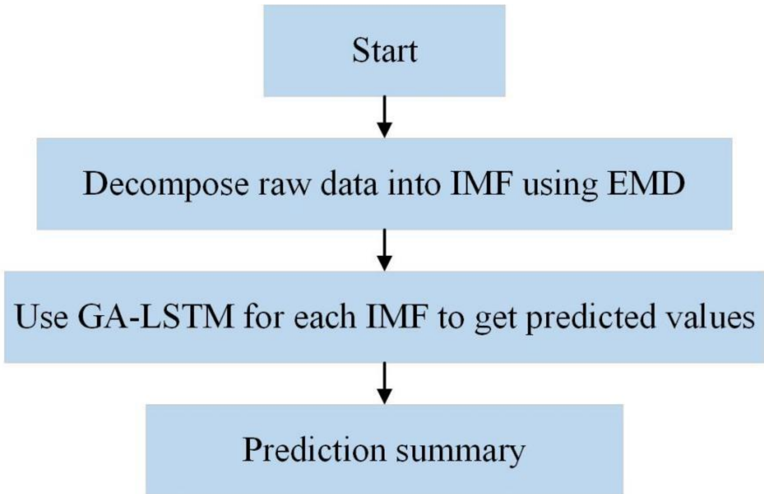

Figure 3 Flow chart of EMD-GA-LSTM model

4 Experiments

4.1 LSTM training process analysis

The experiments in this paper use the collected daily carbon emissions of a region in 2019. The total amount of data is 700, and the first 525 data are selected as the training set and the last 175 data are the test set. The training set is used to train the parameters of the model, and the test set is used to test the performance of the model after training. The data structure is divided into two columns, one for time and one for daily emission data in tens of thousands of tons, and the first ten rows of data are shown in Table 1.

Table 1 Daily carbon emissions

Date	Emission
2019/1/1	31.111
2019/1/2	33.084

2019/1/3	32.596
2019/1/4	32.118
2019/1/5	32.349
2019/1/6	31.401
2019/1/7	32.143
2019/1/8	31.699
2019/1/9	32.310
2019/1/10	32.129

The effectiveness of the models can be measured using evaluation metrics for whether there is an improvement in the models, comparisons between models, etc. The three evaluation metrics used are as follows: mean square error (MSE), root mean square error (RMSE), and mean absolute error (MAE). The mean square error is the square of the difference between the predicted value and the true value, which is then averaged by the following equation:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2 \tag{3}$$

The root mean square error is the square of the difference between the predicted value and the true value, which is averaged and then squared, with the following equation:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2} \tag{4}$$

The mean absolute error is the absolute value of the difference between each predicted value and the true value, which is then summed and averaged, with the following equation:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|\hat{y}_i - y_i| \tag{5}$$

The structure of the established LSTM model is an input layer, an output layer and two LSTM layers, and a Dense layer, which is a fully connected layer. Each LSTM layer is followed by a Dropout layer with a parameter of 0.2, which means that 20% of the links are randomly broken in the learning network of each layer. the Dropout layer is used to prevent overfitting of the model. The model uses stochastic gradient descent as the optimizer, the momentum is set to 0.9, the learning rate is 0.001, and a total of 100 epochs are iterated. Figure 4 shows the mean square error variation during the training of the model.

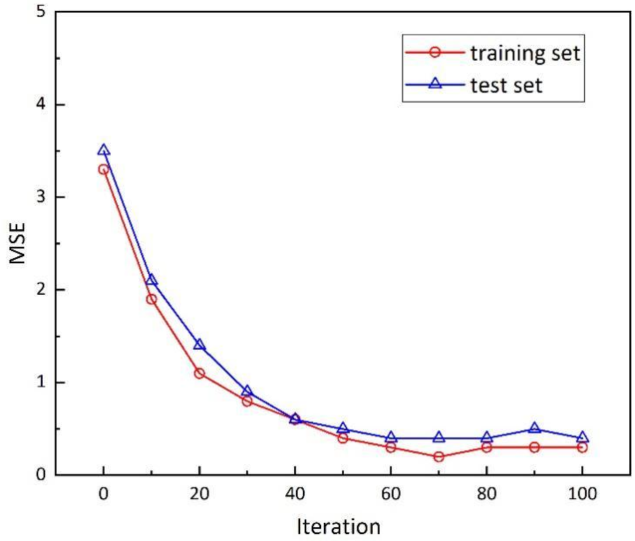

Figure 4 Error variation during model training

4.2 Analysis of GA-LSTM training process

Since a single LSTM network tuning parameter will fall into a local optimum situation, and the genetic algorithm can perform targeted optimization for this problem. The genetic algorithm uses the processed training set data with some parameters selected from the above LSTM model. The GA-LSTM model is trained and the trained LSTM structure is obtained as one input layer, two LSTM layers, two Dense layers and one output layer, respectively. Each LSTM layer and Dense layer is followed by a Dropout layer with the parameter of 0.20.

The test set is brought into the trained LSTM model for prediction, and the prediction results are collected. the root mean square error, the mean square error and the mean absolute error of the test set predictions are 1.33, 1.769, 0.973. the test set curve is fitted with the model predictions, and the fitting results are shown in Figure 5. The blue color represents the true value and the green color represents the predicted value. From the model curve fitting results, the model fit is very good and can show the complete trend and fluctuation of the data.

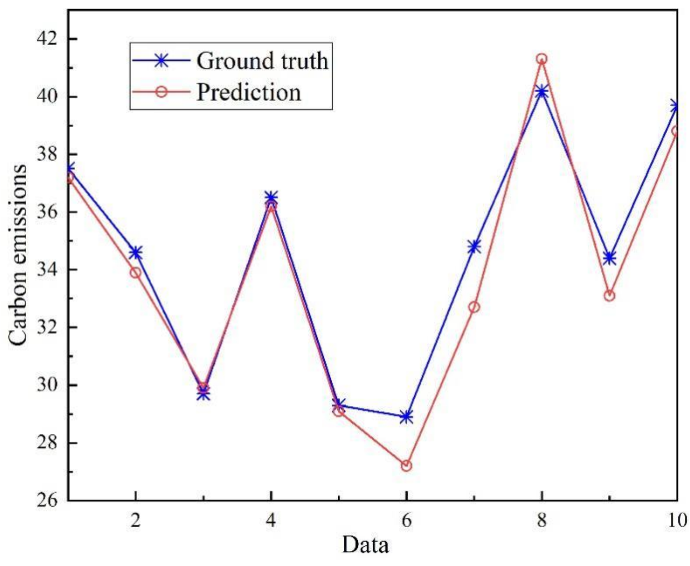

Figure 5 Fitting effect of GA-LSTM model

4.3 Analysis of EMD-GA-LSTM training process

The data were first decomposed into six eigenmodal components and residuals using empirical modal decomposition, named imfl to imf6, and res. The decomposed data were plotted as line graphs to see if the two constraints of EMD were satisfied. The structure of the decomposed data is shown in Table 2, and all the decomposed components satisfy the two constraints of EMD. In Table 2, Imfl to Imf6 represent the components of the eigenmodes, and res represents the residuals after decomposition. Since only the data with 5 decimal places are retained and the residuals are of order $1 \times e^{-15}$, the error of all the eigenmodes is $1 \times e^{-15}$, which is not much different from the true value, so the residuals are taken as 0. Therefore, in the next prediction, only the data from Imfl to Imf6 are used as input for the prediction summary.

Table 2 Table of eigenmodal components

Date	imf_1	imf_2	imf_3	imf_4	imf_5	imf_6	res
2019/1/1	-0.692	-0.462	4.262	2.285	-1.193	26.912	0
2019/1/2	-0.637	0.151	4.288	2.286	-1.196	26.916	0
2019/1/3	-0.147	0.442	4.308	2.271	-1.198	26.919	0
2019/1/4	-0.433	0.268	4.323	2.236	-1.198	26.923	0
2019/1/5	-0.360	-0.255	4.332	2.108	-1.197	26.927	0
2019/1/6	-0.325	-0.456	4.335	2.182	-1.194	26.932	0
2019/1/7	-0.293	-0.248	4.333	2.018	-1.189	26.936	0

The model predictions obtained from the training of each component and the LSTM models of the six components are taken for this experiment, and the errors of each model are shown in Table 3. The root-mean-square error values of all the components are relatively small except for the first component which has a high root-mean-square error value. This indicates that the LSTM network has a better learning effect for the low-frequency components and a poorer learning performance for the high-frequency components, because the fluctuation frequency of the high-frequency components is faster and there is some difficulty in learning using the model, while the low-frequency components fluctuate less frequently and are easier to learn. Figure 6 shows the prediction error of EDM-GA-LSTM and other methods. The root-mean-square error (RMSE) of the prediction results for each component is 0.926, a decrease of 0.443, or 32.4%, compared with the RMSE of 1.369 for the LSTM model. Compared with the root mean square error of 1.33 of the GA-LSTM, it decreased by 0.404, or 30.4%.

Table 3 Component error table

Serial number	RMSE	MSE	MAE
imf_1	0.904	0.818	0.702
imf_2	0.224	0.050	0.152
imf_3	0.048	0.002	0.024
imf_4	0.005	0.000	0.004
imf_5	0.004	0.000	0.003
imf_6	0.000	0.000	0.002

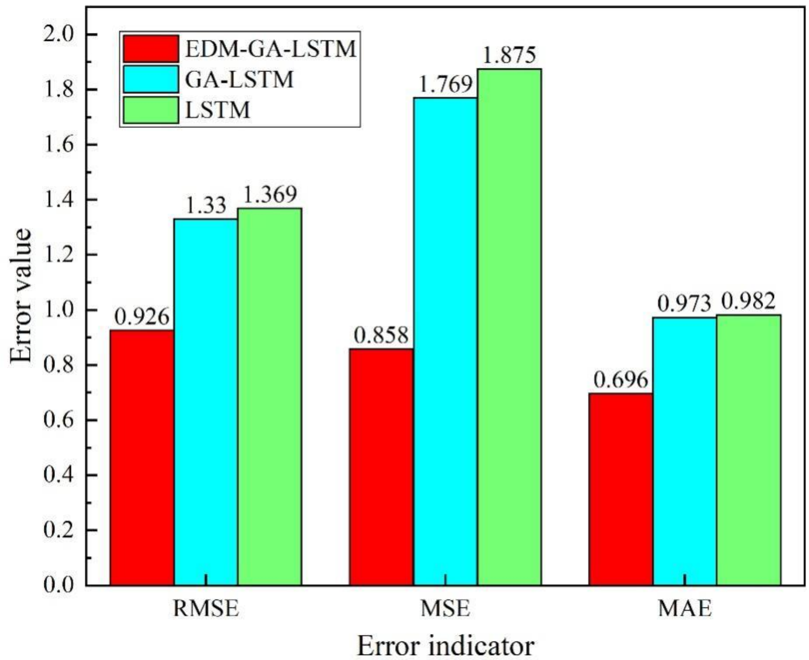

Figure 6 EDM-GA-LSTM prediction error

5 Conclusions

Analyzed from the perspective of China's economic development, the importance of industrial development is reflected in the economic development of the past few centuries. Because of the rapid development of industry, the overall economic development is promoted, but the industrial development is accompanied by a large amount of carbon dioxide emissions, which also brings serious pollution problems to the whole natural environment. Mechanized production is the starting point of industrial revolution, economic development cannot be separated from industrial production, and the mechanized operation will cause serious carbon dioxide pollution. In the development of global technology, carbon emissions show a gradual rise, affecting the global climate and causing environmental pollution problems. According to the current development of the environment, controlling carbon emissions can help reduce environmental problems such as the greenhouse effect caused by industrial development. In this paper, we have explored some analysis of carbon emission time series by combining signal decomposition and genetic algorithm starting from the perspective of deep learning. In this paper, an LSTM model incorporating genetic algorithm and empirical model decomposition algorithm is proposed. For the eigenmodal components obtained from EMD decomposition, the LSTM network can learn the low-frequency components well, but the learning effect of the high-frequency components is poor, indicating that the noise in the high-frequency components is large. Experimental results on real carbon emission datasets show that the method in this paper can accurately and effectively predict carbon emission trends.

Availability of data and material

The datasets used during the current study are available from the corresponding author on reasonable request.

Competing interests

Declares that he has no conflict of interest.

Reference

[1] Bouscayrol A, Chevallier L, Cimetiere X, et al. EPE'13 ECCE Europe, a carbon-neutral conference[J]. EPE Journal, 2018, 28(1): 43-48.
[2] Wang C, Dong G. Research on Green Financial Ecology Construction Based on Low Carbon Economy[J]. Ekoloji, 2019, 28(107): 3635-3641.
[3] Li Z, Wang J. The Dynamic Impact of Digital Economy on Carbon Emission Reduction: Evidence City-level Empirical Data in China[J]. Journal of Cleaner Production, 2022, 351: 131570.
[4] Paramati S R, Mo D, Huang R. The role of financial deepening and green technology on carbon emissions: evidence from major OECD economies[J]. Finance Research Letters, 2021, 41: 101794.
[5] Pan W, Pan W, Hu C, et al. Assessing the green economy in China: an improved framework[J]. Journal of cleaner production, 2019, 209: 680-691.
[6] Liu W, Liu M, Liu T, et al. Does a Recycling Carbon Tax with Technological Progress in Clean Electricity Drive the Green Economy?[J]. International Journal of Environmental Research and Public Health, 2022, 19(3): 1708.
[7] Despins C, Labeau F, Le Ngoc T, et al. Leveraging green communications for carbon emission reductions: Techniques, testbeds, and emerging carbon footprint standards[J]. IEEE Communications Magazine, 2011, 49(8): 101-109.
[8] Wang Y. Research on the relationship between green energy use, carbon emissions and economic growth in Henan province[J]. Frontiers in Energy Research, 2021: 390.
[9] Du K, Li P, Yan Z. Do green technology innovations contribute to carbon dioxide emission reduction? Empirical evidence from patent data[J]. Technological Forecasting and Social Change, 2019, 146: 297-303.
[10] Anser M K, Usman M, Godil D I, et al. Does globalization affect the green economy and environment? The relationship between energy consumption, carbon dioxide emissions, and economic growth[J]. Environmental Science and Pollution Research, 2021, 28(37): 51105-51118.
[11] Mao D, Liu Y P, Wang F Z, et al. Growth Trend Analysis Of Carbon Dioxide (Co2) Emissions And Urban Green Open Spaces: A Case Study Of Henan, China[J]. Oxidation Communications, 2016, 39(4): 3305-3312.
[12] Jin Y, Tang Y M, Chau K Y, et al. How government expenditure Mitigates emissions: A step towards sustainable green economy in belt and road initiatives project[J]. Journal of environmental management, 2022, 303: 113967.
[13] Jiang B, Sun Z, Liu M. China's energy development strategy under the low-carbon economy[J]. Energy, 2010, 35(11): 4257-4264.
[14] Yu X, Li M, Kang W. Heterogeneity of decoupling between economic development and carbon emissions in China's Green Industrial Parks[J]. Earth's Future, e2022EF002753.

[15] Hu Y, Zheng J. How does green credit affect carbon emissions in China? A theoretical analysis framework and empirical study[J]. Environmental Science and Pollution Research, 2022: 1-15.
[16] Zhang F, Deng X, Phillips F, et al. Impacts of industrial structure and technical progress on carbon emission intensity: Evidence from 281 cities in China[J]. Technological Forecasting and Social Change, 2020, 154: 119949.
[17] Hanif I, Aziz B, Chaudhry I S. Carbon emissions across the spectrum of renewable and nonrenewable energy use in developing economies of Asia[J]. Renewable Energy, 2019, 143: 586-595.
[18] Wang M, Feng C. How will the greening policy contribute to China's greenhouse gas emission mitigation? A non-parametric forecast[J]. Environmental Research, 2021, 195: 110779.
[19] Du K, Li J. Towards a green world: How do green technology innovations affect total-factor carbon productivity[J]. Energy Policy, 2019, 131: 240-250.
[20] Wei S, Yuwei W, Chongchong Z. Forecasting CO2 emissions in Hebei, China, through moth-flame optimization based on the random forest and extreme learning machine[J]. Environmental Science and Pollution Research, 2018, 25(29): 28985-28997.
[21] Patterson D, Gonzalez J, Hölzle U, et al. The carbon footprint of machine learning training will plateau, then shrink[J]. Computer, 2022, 55(7): 18-28.
[22] Xu H, Wang M, Jiang S, et al. Carbon price forecasting with complex network and extreme learning machine[J]. Physica A: Statistical Mechanics and its Applications, 2020, 545: 122830.
[23] Sun C. The correlation between green finance and carbon emissions based on improved neural network[J]. Neural Computing and Applications, 2021: 1-15.
[24] Sun W, Huang C. Predictions of carbon emission intensity based on factor analysis and an improved extreme learning machine from the perspective of carbon emission efficiency[J]. Journal of Cleaner Production, 2022, 338: 130414.
[25] Chen B. Air quality index forecasting via deep dictionary learning[J]. IEICE TRANSACTIONS on Information and Systems, 2020, 103(5): 1118-1125.
[26] Xu Z, Cao Y, Kang Y. Deep spatiotemporal residual early-late fusion network for city region vehicle emission pollution prediction[J]. Neurocomputing, 2019, 355: 183-199.
[27] Yang Z, Wang J. A new air quality monitoring and early warning system: Air quality assessment and air pollutant concentration prediction[J]. Environmental research, 2017, 158: 105-117.
[28] Saini T, Chaturvedi P, Dutt V. Modelling particulate matter using multivariate and multistep recurrent neural networks[J]. Frontiers in Environmental Science, 2021: 614.
[29] Du S, Li T, Yang Y, et al. Deep air quality forecasting using hybrid deep learning framework[J]. IEEE Transactions on Knowledge and Data Engineering, 2019, 33(6): 2412-2424.
[30] Zhang B, Zhang H, Zhao G, et al. Constructing a PM2. 5 concentration prediction model by combining auto-encoder with Bi-LSTM neural networks[J].

Environmental Modelling & Software, 2020, 124: 104600.
[31] Bo-Lun Chen, Guo-Chang Zhu, Yi-Yun Sheng et al. An Empirical Mode Decomposition Fuzzy Forecast Model for Air Quality, 09 July 2021, PREPRINT (Version 1) available at Research Square [https://doi.org/10.21203/rs.3.rs-676389/v1]