



A novel short-term carbon emission prediction model based on secondary decomposition method and long short-term memory network

Feng Kong¹ · Jianbo Song¹ · Zhongzhi Yang¹

Received: 5 January 2022 / Accepted: 18 April 2022 / Published online: 28 April 2022

© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

Abstract

Grasping the dynamics of carbon emission in time plays a key role in formulating carbon emission reduction policies. In order to provide more accurate carbon emission prediction results for planners, a novel short-term carbon emission prediction model is proposed. In this paper, the secondary decomposition technology combining ensemble empirical mode decomposition (EEMD) and variational mode decomposition (VMD) is used to process the original data, and the partial autocorrelation function (PACF) is applied to select the optimal model input. Then, the long short-term memory network (LSTM) is chosen for prediction. The secondary decomposition algorithm is innovatively introduced into the field of carbon emission prediction, and the empirical results illustrate that the secondary decomposition technology can further improve the prediction accuracy. Combined with the secondary decomposition, the R^2 , MAPE, and RMSE of the model are improved by 2.20%, 43.08%, and 36.92% on average. And the proposed model shows excellent prediction accuracy ($R^2=0.9983$, MAPE=0.0031, RMSE=118.1610) compared with other 12 comparison models. Therefore, this model not only has potential value in the formulation of carbon emission reduction plans, but also provides a valuable reference for future carbon emission forecasting research.

Keywords Short-term carbon emission prediction · The secondary decomposition algorithm · The long short-term memory network

Introduction

The rise of global temperature has aroused more and more discussion and attention. With the continuous deterioration of the global climate, glaciers continue to melt, biodiversity is lost, and the hydrological cycle model is destroyed, which leads to frequent droughts and serious damage to the ecosystem (Park et al. 2020). The increase of greenhouse

gas emissions is considered to be a major factor resulting in this phenomenon (Hashim et al. 2015). To cope with this situation, the Chinese government has formulated a series of carbon emission reduction plans, for instance, the newly proposed “double carbon” target (carbon peak in 2030 and carbon neutrality in 2060). By 2019, China’s carbon intensity has decreased by 18.2% compared to 2015, completing the binding targets of the 13th Five-Year Plan ahead of schedule. During the 14th Five-Year Plan period, China designs to reduce energy consumption and carbon dioxide emissions per unit of GDP by 13.5% and 18% respectively. Confronting with severe challenges, it is necessary for the Chinese government to formulate a series of practical carbon emission reduction strategies.

Rational prediction of the future CO₂ is essential to the carbon emission reduction targets (Zhao and Yang 2020). The long-term trend of carbon emission can be observed through the annual carbon emission prediction. However, in order to adjust carbon emission reduction policies and measures in a timely and flexible manner, it is necessary

Responsible Editor: Ilhan Ozturk

✉ Jianbo Song
songjianbo@ncepu.edu.cn

Feng Kong
kongfeng@ncepu.edu.cn

Zhongzhi Yang
yzz@ncepu.edu.cn

¹ Department of Economics and Management, North China Electric Power University, Hebei 071003, People’s Republic of China

to have a clearer grasp of the dynamics and fluctuations of short-term carbon emissions, which can provide sufficient time for planners to respond. The prediction of daily carbon emissions not only can help us gain a clearer picture of carbon emissions, but also plays a key role in formulating short-term carbon emission reduction plans and controlling carbon emissions. Similar to wind speed prediction and carbon price prediction, short-term carbon emission prediction can also apply time series prediction method. Moreover, in contrast with traditional machine learning algorithms, LSTM does not need feature engineering in the construction process and can better mine the time relationship between data series (Sun and Huang 2020). And there are nonlinearity and fluctuation in the original sequence of daily carbon emissions, and the volatility of data will have a certain impact on the prediction (Sun and Ren 2021). When the amount of data is large and highly irregular, data decomposition technology will improve the precision of prediction to a certain extent (E et al. 2021). Mi et al. (2017) proved in practice that the secondary decomposition method can well deal with nonlinear time series with high complexity and volatility. It is a good attempt to apply the secondary decomposition method to the field of short-term carbon emission. Under this background, a novel hybrid short-term carbon emission time series prediction model based on ensemble empirical mode decomposition (EEMD), variational mode decomposition (VMD), and long short-term memory network (LSTM) is proposed.

The possible innovations and contribution of this paper are mainly described as three aspects:

- (1) The secondary decomposition method is applied to carbon emission prediction for the first time. Taking into account the nonlinear characteristics of sample data, the idea of secondary decomposition is innovatively introduced into the field of carbon emission prediction, which enormously increases the forecasting precision. This enriches the research of secondary decomposition and provides a new thought for the exploration of carbon emission prediction.
- (2) This paper puts forward a carbon emission prediction model based on LSTM. LSTM is innovatively applied in the field of short-term carbon emission, which expands the application of deep learning algorithm.
- (3) The proposed EEMD-VMD-LSTM model accurately predicts short-term carbon emissions. In this paper, 912 daily carbon emission monitoring data from January 1, 2019 to June 30, 2021 are selected for prediction. The increase of sample size can greatly improve the learning ability and prediction performance of LSTM. The model can not only provide guidance for high carbon emission industries and regions to rationally control emissions but also contribute to improving the decision-making level of relevant departments.

The remaining sections of this paper are organized as follows: “[Literature review](#)” section shows the literature review. “[Methodology](#)” section describes the decomposition algorithm and prediction model, meanwhile, the construction process of the proposed model is explained. “[Empirical analysis](#)” section contains the data source and processing method, evaluating indicators, and model parameters. “[Comparison experiments](#)” section introduces a comparative experimental analysis of five groups and supplementary case. Finally, the discussion, the conclusion of the research, and prospects are presented.

Literature review

There are a large number of domestic and international research on carbon emission prediction; many researchers predict carbon emissions by analyzing the influential factors. They first screen out the influential factors highly correlated to carbon emission and then employ them as input data for carbon emission prediction research. Table 1 lists the relevant literature.

In the literature, they almost all noted that the limitation of the research was ignoring the potential influencing factors and the lack of data. Time series forecasting is a means of using historical data to predict the data in a certain period of time in the future. It emphasizes the importance of the time factor in prediction and temporarily ignores other external influential factors (Sun and Ren 2021). Applying the time series forecasting method effectively overcomes the above problems. But due to the small sample characteristics of the annual carbon dioxide emission series (Gao et al. 2021), only a few scholars chose time series method to predict. For instance, Yang and O’Connell (2020) combined the idea of time series prediction and applied the autoregressive integrated moving average (ARIMA) to forecast the carbon emissions during air transportation in Shanghai in the next 5 years. Malik et al. (2020) selected the same method to predict the annual carbon dioxide emissions of Pakistan. Similarly, Heydari et al. (2019) did equal research and assessed annual carbon dioxide emission trends in Italy, Iran, and Canada. However, they all mentioned that the small sample size will affect the accuracy of the model to a certain extent. Therefore, we break through the limitation and expand our horizons to the field of short-term carbon emission time series prediction.

As an effective algorithm in the field of deep learning (Memarzadeh and Keynia 2020), it has the merits of nonlinear prediction ability, fast convergence speed, and capturing the long-term correlation of time series (Ma et al. 2015). Moreover, LSTM was applied to many hot prediction fields such as electricity price prediction (Chang et al. 2019), flood prediction (Ding et al. 2020), financial market prediction

Table 1 Research on carbon emission prediction

Reference	Means	Period	Major affecting factors
(Wen and Cao 2020)	The support vector machine based on improved chicken swarm optimization	2000–2016	Economic level, living standard, social condition, education level
(Huang et al. 2019)	Long short-term memory network	1990–2016	The proportion of thermal power generation, industrial structure, energy consumption per unit of GDP
(ŞEntürk and Zehra 2021)	Artificial neural network	1990–2015	Energetic and economic indicators
(Aslam et al. 2021)	ARDL bound test	1971–2016	Economic growth, energy
(Wang and Wang 2021)	Extreme learning machine optimized by manta rays foraging optimization	2000–2017	Urbanization rate, energy intensity, freight turnover
(Boamah et al. 2021)	Brainstorm optimization algorithm	1990–2014	Economy, urbanization rate, imports, and exports
(Hosseini et al. 2019)	Multiple linear regression and multiple polynomial regression analysis	1975–2015	Population, CO ₂ intensity, GDP, electricity production from fossil fuels, energy use
(Li et al. 2021b)	Genetic algorithm optimized extreme learning machine	2000–2017	Economic scale, population size, transportation structure, energy consumption
(Wang et al. 2020)	Support vector machine optimized by particle swarm optimization	1980–2020	Total population, urbanization rate, GDP per capita, transport energy intensity, etc

(Fischer and Krauss 2018), traffic flow prediction (Luo et al. 2019), and wind speed prediction (Gu et al. 2021). The practice in these fields proved that LSTM is an advanced technology for time series prediction. To forecast short-term carbon emissions more accurately, this paper chooses LSTM as the prediction method.

Data decomposition algorithm can effectively reduce the difficulty of prediction. Wu and Huang (2009) proposed ensemble empirical mode decomposition (EEMD), which can significantly reduce the fluctuation of original data. Sun and Ren (2021) constructed a hybrid prediction model combining EEMD and backpropagation neural network optimized by particle swarm optimization (PSO-BP) to predict short-term carbon emissions in China. This research conclusion highlighted that EEMD is more suitable to deal with short-term carbon emission data than other primary decomposition methods. Thus EEMD is selected as the first decomposition method in this paper.

Although EEMD improves the prediction accuracy of the model to some extent, it also has limitations. Liu et al. (2013) noted that in all the modal functions decomposed, the first intrinsic mode function (IMF1) has the highest irregularity and complexity, which could affect the accuracy of prediction. Sun and Huang (2020) found that further decomposition of IMF1 can effectively solve this problem. In addition, the excellent data recognition and time series information extraction ability of the secondary decomposition method have been proved in some fields (Li et al. 2021a; Liu and Zhang 2021) since VMD has the characteristics of high adaptability and good denoising effect (Zhao et al. 2020). And in practice, compared with fast ensemble empirical mode decomposition (FEEMD) and wavelet decomposition (WD), reprocessing IMF1 with VMD can

more significantly reduce the complexity and forecasting difficulty of time series (Sun and Huang 2020). Therefore, this paper firstly chooses EEMD to preprocess the original short-term carbon emission sequence, and IMF1 obtained by EEMD is re-decomposed by VMD to improve the performance of the prediction model.

Through the literature review, it can be found that there are some defects in influencing factor prediction method. For example, some potential influencing factors have not been taken into account, the data collection cannot be completed due to the defects of the existing database, and the small sample size will lead to poor model prediction accuracy. The previous literature mainly focuses on the long-term carbon emission prediction, which can only reflect the trend of carbon emission. Moreover, the secondary decomposition is an effective data processing method, but the carbon emission prediction model combined with the method to further improve the prediction accuracy needs to be explored. In consideration of the shortcomings of the existing research, this study develops the EEMD-VMD-LSTM time series prediction model to predict the short-term carbon emissions.

Methodology

Ensemble empirical mode decomposition (EEMD)

The EMD algorithm is suitable for processing and analyzing the non-stationary and nonlinear signals, but this method has the following shortcomings: (1) the IMF component obtained by EMD decomposition has the phenomenon of mode aliasing; (2) the end-points-extending problem affects the decomposition effect. To restrain the phenomenon,

Huang and Wu (2008) added Gaussian white noise to the original sequence to effectively extract periodic and trend information. They successfully overcame the weaknesses of EMD and proposed EEMD.

Implementation steps of the EEMD algorithm are as follows:

- (1) A Gaussian white noise $n_i(t)$ with standard normal distribution is added to the original sequence $x(t)$ to generate a new signal:

$$\hat{x}_i(t) = x(t) + n_i(t) \quad (1)$$

In which $x(t)$ is the initial sequence and $n_i(t)$ is the Gaussian white noise.

- (2) EMD decomposition is performed on the obtained signal $\hat{x}_i(t)$ containing noise.

$$\hat{x}_i(t) = \sum_{j=1}^J c_{i,j}(t) + r_{i,j}(t) \quad (2)$$

where $c_{i,j}(t)$ is the k th IMF decomposed after adding Gaussian white noise for the j th time. In addition, $r_{i,j}(t)$ stands for the residual function, which represents the average trend of the signal.

- (3) Repeat step 1 and step 2, and the IMF set is obtained by adding white noise signals with different amplitudes in each decomposition.
- (4) Perform the averaging operation on the corresponding IMF set above to get the final result:

$$C_j(t) = \frac{1}{M} \sum_{i=1}^M c_{i,j}(t) \quad (3)$$

Where $C_j(t)$ is the j th component, and M is the number of Gaussian white noise sequences.

Variational mode decomposition (VMD)

The VMD algorithm is a non-recursive signal decomposition method proposed by Dragomiretskiy and Zosso (2014), which includes constructing variational problem and solving variational problem. The bandwidth and center frequency of each IMF component can be updated alternately and iteratively during the decomposition process. At last, a series of effective components with limited bandwidth can be obtained from the original signal through adaptively decomposing.

Construction of variational problem

For each IMF, the single-sided spectrum is obtained by applying Hilbert transform to calculate its analytic signal. And

through adding an exponential term $e^{-j\omega_k t}$, each single-sided spectrum is modulated to the corresponding baseband.

$$\left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-j\omega_k t} \quad (4)$$

Estimate the bandwidth of each IMF with Gaussian smoothing. Then, the constrained variational problem that needs to be settled by VMD is represented as follows:

$$\min_{u_k, \omega_k} \left\{ \sum_{k=1}^k \left\| \partial_t \times \left[\delta(t) + \frac{j}{\pi t} \right] * u_k(t) \times e^{-j\omega_k t} \right\|^2 \right\} \quad (5)$$

$$s.t. \sum_{k=1}^k u_k = x(t)$$

In which $\{u_k\} = \{u_1, \dots, u_k\}$ represents the IMF component set obtained by VMD; $\{\omega_k\} = \{\omega_1, \dots, \omega_k\}$ represents the frequency center of each IMF component.

Solution to variational problems

In this process, the quadratic penalty factor ξ and Lagrange multiplier $\lambda(t)$ are introduced to transform the constrained variational problem into an unconstrained variational problem. The new expression is as follows:

$$L(u_k, \omega_k, \lambda) = \xi \sum_{k=1}^k \left\| \partial_t * \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-j\omega_k t} \right\|^2$$

$$+ \left\| f(t) - \sum_{k=1}^k u_k(t) \right\|^2 + \langle \lambda(t), f(t) - \sum_{k=1}^k u_k(t) \rangle \quad (6)$$

Solve the above equation using the alternate direction method of multipliers. To find the minimum point of the augmented Lagrange expression, we update u_t^{n+1} , ω_t^{n+1} and λ_k^{n+1} alternately.

The following formula represents the value of u_t^{n+1} :

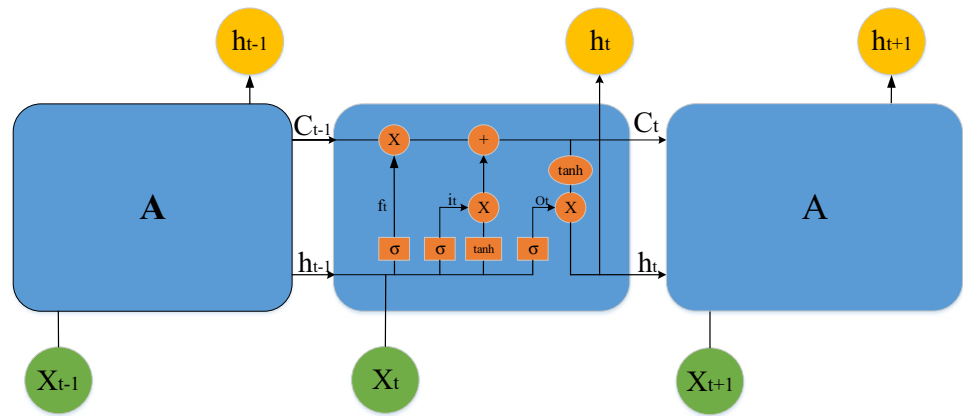
$$u_t^{n+1} = \underset{u_k \in X}{\operatorname{argmin}} \left\{ \xi \left\| \partial_t * \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-j\omega_k t} \right\|^2 + \left\| f(t) - \sum_{k=1}^k u_k(t) \right\|^2 \right\} \quad (7)$$

Among them, ω_k^n is equal to ω_t^{n+1} and $\sum_{i=1}^i u_i(t)$ is the same to $\sum_{i \neq k}^i u_i(t)^{n+1}$.

Update formula (7) by the inverse Fourier transform method.

$$\hat{u}_t^{n+1} = \underset{\hat{u}_k u_k \in X}{\operatorname{argmin}} \left\{ \xi \left\| j\omega \left[(1 + \operatorname{sgn}(\omega + \omega_k)) \hat{u}_k(\omega + \omega_k) \right] \right\|_2^2 \right.$$

$$\left. + \left\| \hat{f}(\omega) - \sum_{i=1}^i \hat{u}_i(\omega) + \frac{\hat{\lambda}(\omega)}{2} \right\|_2^2 \right\} \quad (8)$$

Fig. 1 Framework of LSTM

Then use $\omega - \omega_k$ to replace ω in the first item to obtain formula (9).

$$\hat{u}_t^{n+1} = \underset{\hat{u}_k u_k \in X}{\operatorname{argmin}} \left\{ \xi \|j(\omega - \omega_k) [(1 + \operatorname{sgn}(\omega)) \hat{u}_k(\omega)]\|_2^2 + \|\hat{f}(\omega) - \sum_{i=1}^i \hat{u}_i(\omega) + \frac{\hat{\lambda}(\omega)}{2}\|_2^2 \right\} \quad (9)$$

Next, the solution of the quadratic optimization problem is obtained through calculating, as shown in formula (10):

$$\hat{u}_t^{n+1}(\omega) = \frac{\hat{f}(\omega) - \sum_{i \neq k}^i \hat{u}_i(\omega) + \frac{\hat{\lambda}(\omega)}{2}}{1 + 2\xi(\omega - \omega_k)^2} \quad (10)$$

We use the similar method to update the center frequency:

$$\omega_k^{n+1} = \frac{\int_0^\infty \omega |\hat{u}_k(\omega)|^2 d\omega}{\int_0^\infty |\hat{u}_k(\omega)|^2 d\omega} \quad (11)$$

Long short-term memory model (LSTM)

LSTM is a special recurrent neural network (RNN) (Chang et al. 2011) model initially developed by Hochreiter and Schmidhuber (1997). Previous studies have proved that because of the vanishing gradient problem, the traditional RNN cannot capture the long-term dependence of time series in the training process. However, LSTM can enhance the learning process of time series to prevent the gradient of information transmission from disappearing. The cells of the hidden layer in the LSTM allow long sequences to flow through the gradient, and the memory blocks containing self-connected memory units can remember the temporal state. By designing forget gate, input gate, and output gate, LSTM can control the information flow in and out of the memory block. Comparing with the traditional RNN,

LSTM is not only easier to train, but also can learn time series with long time spans. The construction of LSTM is visualized in Fig. 1.

All the three gates are sigmoid units, and σ represents the standard logistics sigmoid function defined in formula (12):

$$\sigma(x) = \frac{1}{1 + e^{-x}} \quad (12)$$

The forget gate layer controls the forgetting information of the cell:

$$f_t = \sigma(w_f \cdot [h_{t-1}, x_t] + b_f) \quad (13)$$

Among them, h_{t-1} represents the previous output, x_t is the current input, w stands for the input weight matrices, and b represents the offset.

The input information flowing into the memory unit is controlled by the input gate layer.

$$i_t = \sigma(w_i \cdot [h_{t-1}, x_t] + b_i) \quad (14)$$

$$\hat{C}_t = \tanh(w_c \cdot [h_{t-1}, x_t] + b_c) \quad (15)$$

where the updated value is calculated by the sigmoid function, and the tanh function generates a new candidate value \hat{C}_t .

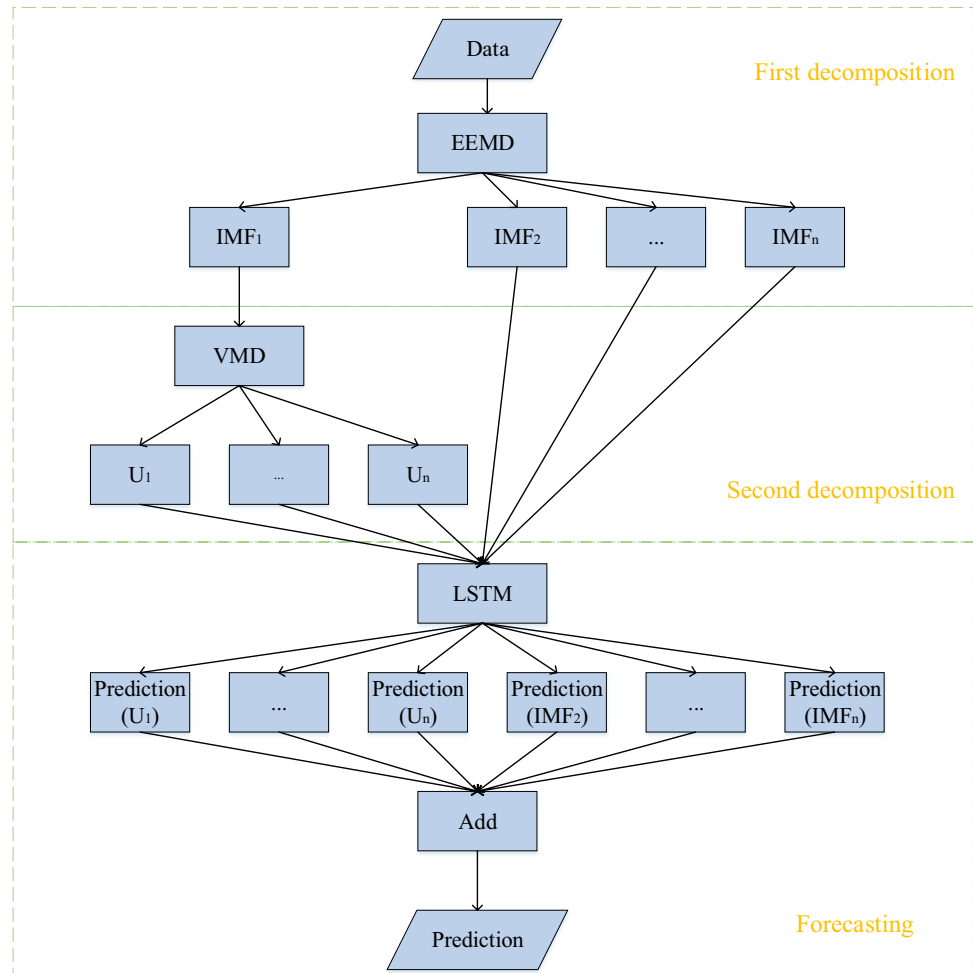
By moderating input features and forgetting partial information, we can update the state of memory cells.

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \hat{C}_t \quad (16)$$

The output gate layer controls the output information of the cell, which yields:

$$O_t = \sigma(w_o \cdot [h_{t-1}, x_t] + b_o) \quad (17)$$

Fig. 2 The process of proposed model



Use the tanh function to scale the C_t value. Then multiply $\tanh C_t$ and O_t to obtain the expected output.

$$h_t = O_t \cdot \tanh(C_t) \quad (18)$$

Construction of the proposed model

The framework of EEMD-VMD-LSTM model is shown in Fig. 2. The specific construction process is as step 1 to step 3 and step 4 introduce the comparative experiment:

- (1) Step 1: Data processing. EEMD is used to process the irregular and complex daily carbon emission monitoring data, which obtains a series of IMF components with low nonlinearity. After that, IMF1 with the most random components is further decomposed into 11 subsequences by VMD. Meanwhile, the remaining IMF components are not processed. This process is known as “the secondary decomposition.”
- (2) Step 2: Selection of input data. The PACF is utilized to analyze the relationship between each sequence and the relative data. Then the appropriate data set is selected as the input of LSTM model.
- (3) Step 3: Prediction. The LSTM model, which belongs to the field of deep learning, is used to predict each internal mode function. Through accumulating the prediction results of each IMF component, the final result is obtained.
- (4) Step 4: Comparative experiment. To prove the accuracy and stability of the proposed model, five groups of comparative experiments are constructed in this paper. Through the combination of decomposition algorithm and prediction model, we can get EEMD-VMD-LSTM and the other 12 comparison models. Figure 3 clearly shows the work process.

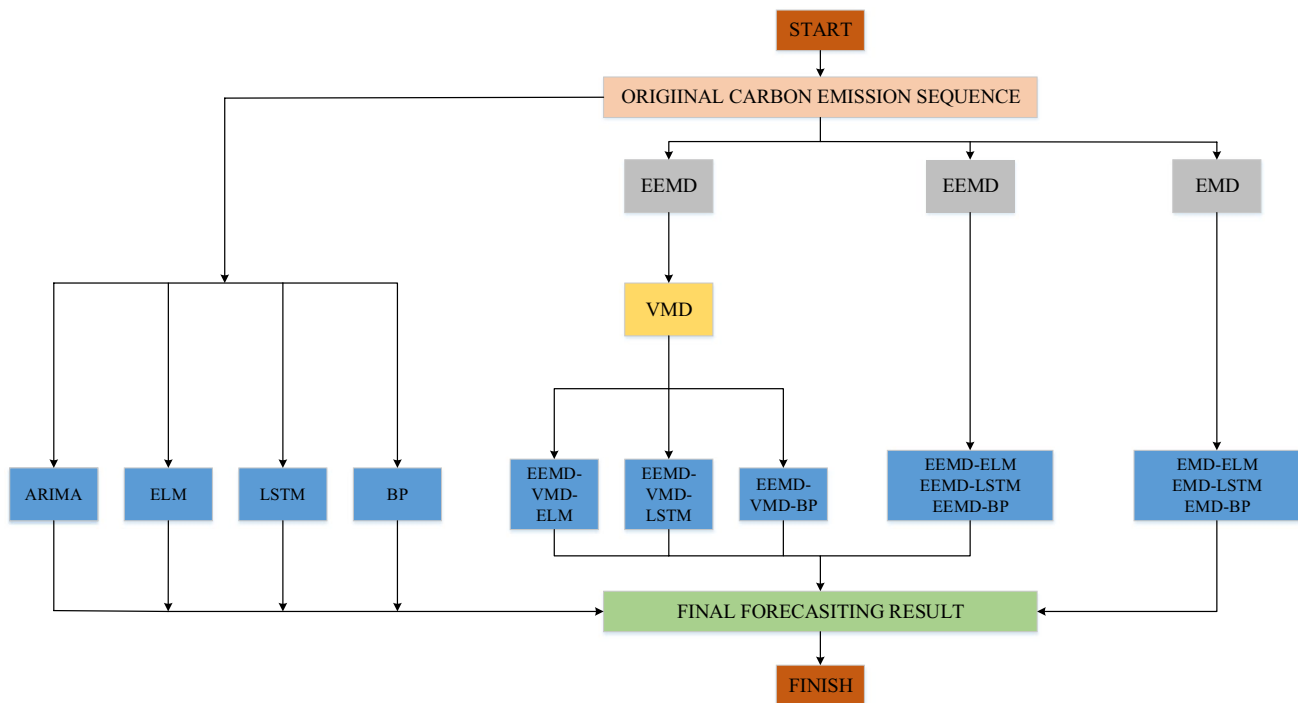


Fig. 3 Comparative experiment process

Empirical analysis

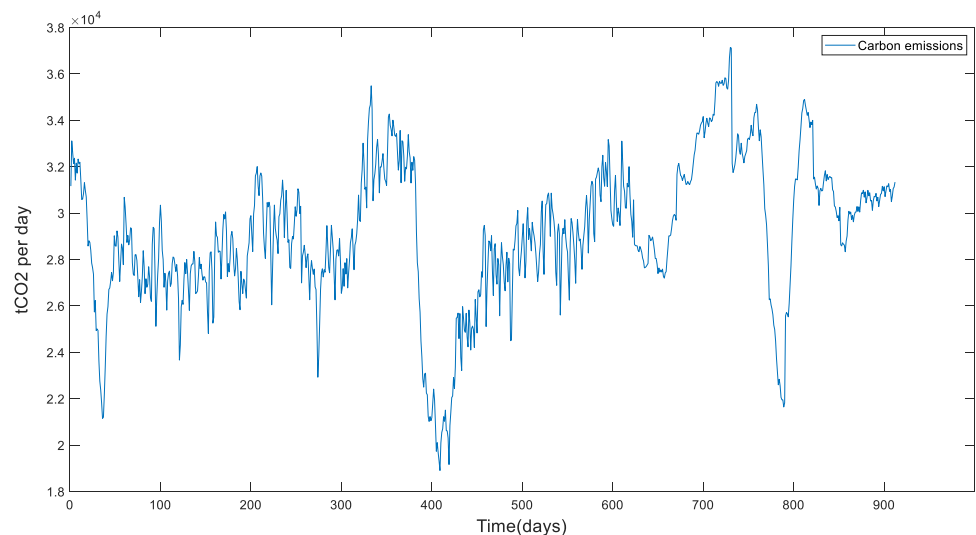
Data

This paper collects the daily carbon emission data of China from January 1, 2019 to June 30, 2021 (data from Carbon Emission Accounts & Datasets, <https://www.ceads.net/>), which are taken as a sample set for simulation and

empirical analysis. For the sample data set, the last 180 data are retained as the testing set of the model, and the remaining data are used for model training. The proportion of the two sets of data is about 80% and 20% respectively.

As visualized in Fig. 4, apparently, the original data of Chinese daily carbon emissions shows high complexity and strong volatility.

Fig. 4 Original data of carbon emission in China



Data processing

The primary decomposition

The original data has high-level nonlinearity and irregularity, which greatly increase the difficulty of forecasting. EEMD algorithm is an adaptive non-stationary signal processing method, which can decompose the signal adaptively depending on the characteristics of the signal itself. After EEMD processing, 11 modal functions are obtained. As is delineated in Fig. 5, in contrast with the original signal, the subsequence after decomposing is smoother and more regular. Besides, as a comparison means, the EMD decomposition results are illustrated in Fig. 6. As shown in Fig. 6, the modal functions obtained by EMD have the problem of mode aliasing, which means that the modal function is not a component of a single frequency, so it is impossible to accurately obtain the instantaneous frequency. Compared with the modal functions from EMD, the fluctuation of the modal functions decomposed by EEMD is more regular.

The secondary decomposition

In this paper, there are two decomposition processes. At First, EEMD is chosen to deal with the original sequence for the first time, and then VMD is applied to further process the sub-sequence IMF1 with the highest complexity. As shown in Fig. 7, 11 subsequences are obtained from IMF1.

In contrast with IMF1, the subsequences are smoother and easier to forecast.

Model accuracy evaluation index

To quantify the prediction accuracy of the models, the goodness of fit value (R^2), the mean absolute percentage error (MAPE), and the root mean square error (RMSE) are introduced as indicators to evaluate the prediction effect of the models. The closer R^2 is to 0, the worse the model fit. The range of MAPE and RMSE is $[0, +\infty]$. The larger the indicators, the greater the error, indicating that the prediction performance of the model is worse. The calculation formulas for the above three evaluation indexes are as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y}_i)^2} \quad (19)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100\% \quad (20)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (21)$$

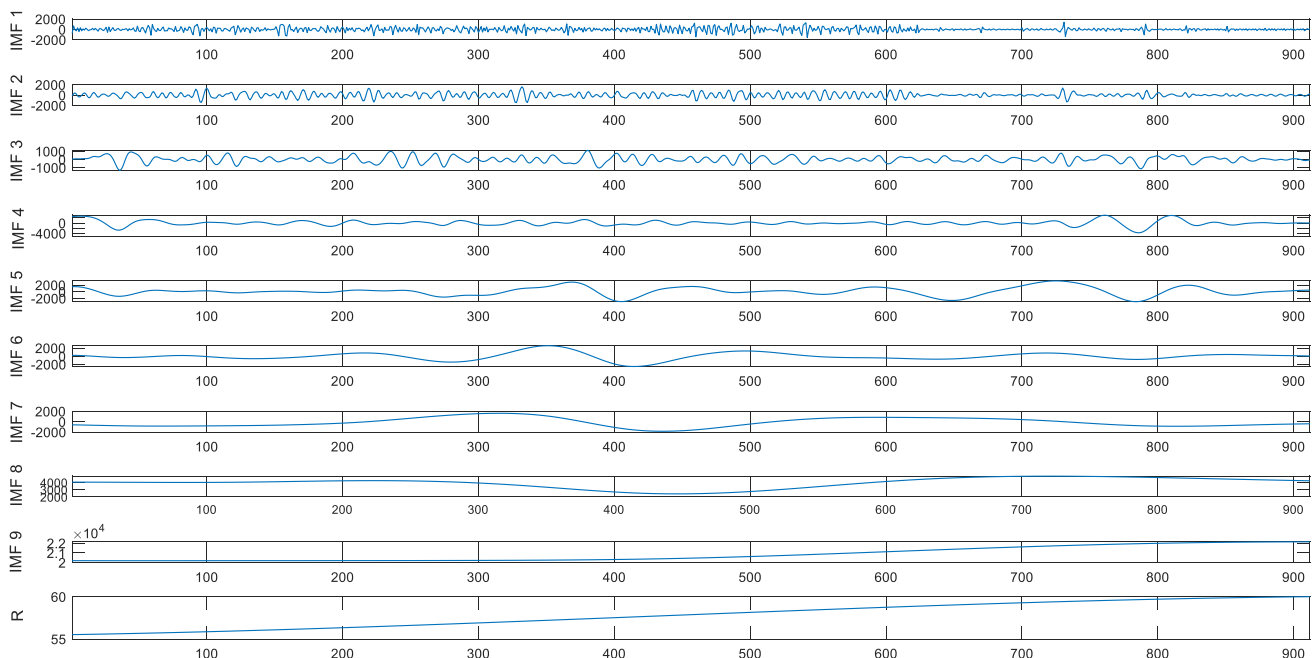


Fig. 5 EEMD results of China dataset

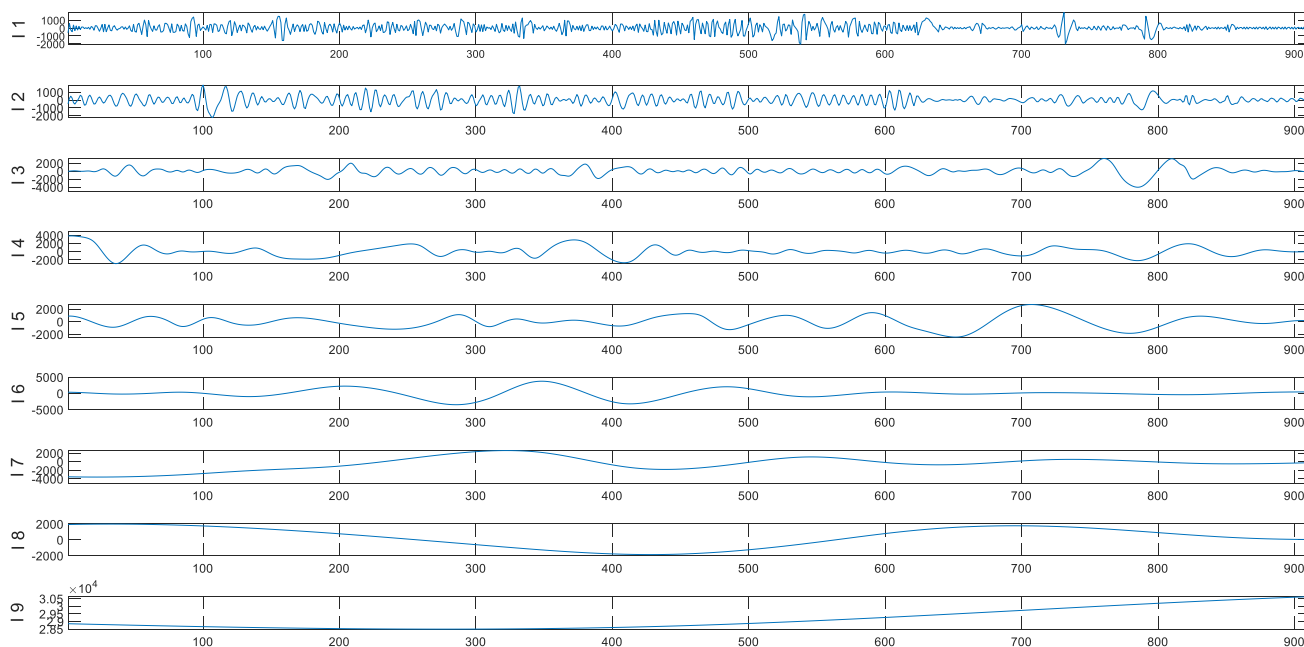


Fig. 6 EMD results of China dataset

Input selection and model parameters

Determine the time-lagged factors according to the partial autocorrelation analysis results of each modal function, and select the data corresponding to the time-lagged factors as the input of the follow-up prediction

procedure. The input variables of the model are demonstrated in Table 2.

Model parameter setting is a very significant step, for the inappropriate parameters will affect the results of prediction. The number of hidden units directly determines the fitting ability of the model. If the number of hidden units is too small, LSTM cannot establish complex

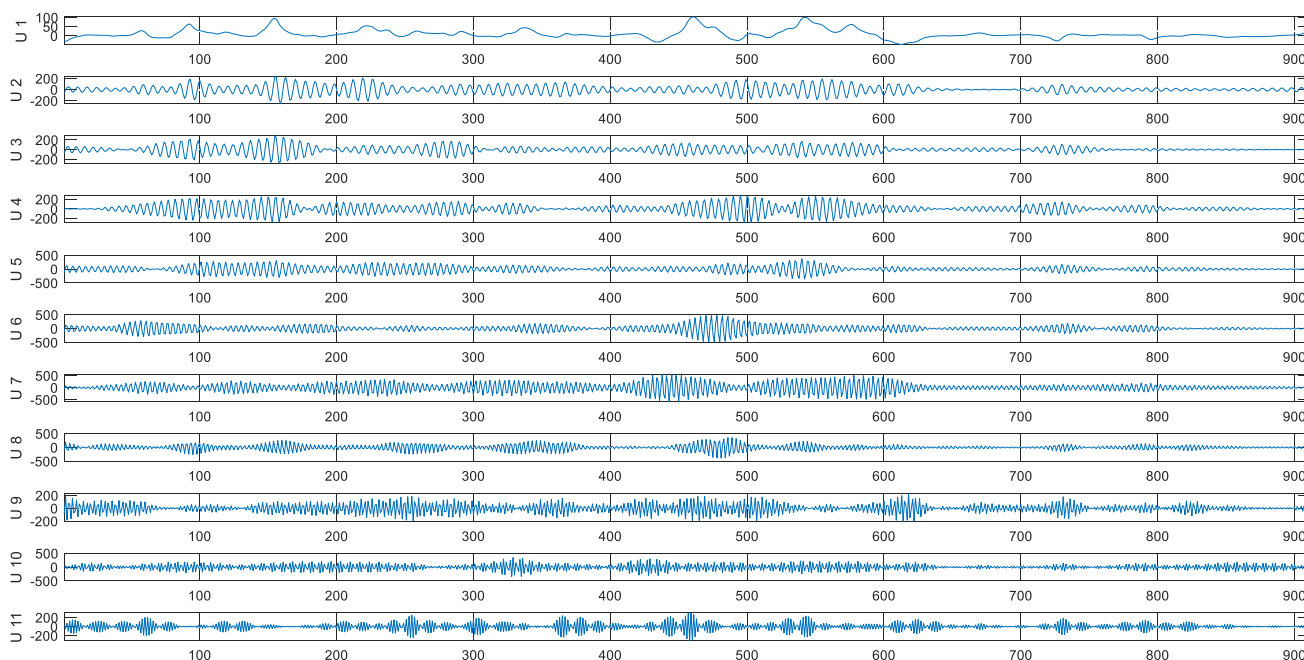


Fig. 7 VMD results of China dataset

Table 2 Input of the prediction model of China dataset

Series	Lags
IMF1	$x_{t-1}, x_{t-2}, x_{t-3}, x_{t-4}, x_{t-5}, x_{t-6}, x_{t-7}, x_{t-8}, x_{t-9}, x_{t-10}, x_{t-11}, x_{t-12}, x_{t-13}$
IMF2	$x_{t-1}, x_{t-2}, x_{t-3}, x_{t-4}, x_{t-5}, x_{t-6}, x_{t-8}, x_{t-9}, x_{t-10}, x_{t-11}, x_{t-12}, x_{t-13}$
IMF3	$x_{t-1}, x_{t-2}, x_{t-3}, x_{t-4}, x_{t-5}, x_{t-6}, x_{t-7}, x_{t-8}, x_{t-10}, x_{t-14}, x_{t-15}$
IMF4	$x_{t-1}, x_{t-2}, x_{t-3}, x_{t-4}, x_{t-5}, x_{t-6}, x_{t-7}, x_{t-8}, x_{t-9}, x_{t-10}, x_{t-12}, x_{t-13}, x_{t-14}, x_{t-15}$
IMF5	$x_{t-1}, x_{t-2}, x_{t-3}, x_{t-4}, x_{t-5}, x_{t-6}, x_{t-7}, x_{t-8}, x_{t-9}, x_{t-10}$
IMF6	$x_{t-1}, x_{t-2}, x_{t-3}, x_{t-4}, x_{t-5}, x_{t-6}, x_{t-7}, x_{t-8}$
IMF7	$x_{t-1}, x_{t-2}, x_{t-3}, x_{t-4}, x_{t-5}, x_{t-6}, x_{t-7}, x_{t-8}, x_{t-9}, x_{t-10}$
IMF8	$x_{t-1}, x_{t-2}, x_{t-3}, x_{t-4}, x_{t-5}, x_{t-6}, x_{t-7}, x_{t-9}, x_{t-10}, x_{t-13}$
IMF9	$x_{t-1}, x_{t-2}, x_{t-3}, x_{t-4}, x_{t-5}, x_{t-6}, x_{t-7}, x_{t-9}, x_{t-12}, x_{t-14}$
IMF10	$x_{t-1}, x_{t-2}, x_{t-3}, x_{t-4}, x_{t-5}, x_{t-6}, x_{t-7}, x_{t-8}, x_{t-9}, x_{t-13}$
R	$x_{t-1}, x_{t-2}, x_{t-3}, x_{t-4}, x_{t-5}, x_{t-7}, x_{t-8}, x_{t-9}, x_{t-10}, x_{t-11}, x_{t-12}, x_{t-13}, x_{t-15}$
U1	$x_{t-1}, x_{t-2}, x_{t-3}, x_{t-4}, x_{t-5}, x_{t-6}, x_{t-7}, x_{t-10}, x_{t-13}, x_{t-15}$
U2	$x_{t-1}, x_{t-2}, x_{t-3}, x_{t-4}, x_{t-5}, x_{t-6}, x_{t-7}, x_{t-8}, x_{t-9}, x_{t-10}, x_{t-11}, x_{t-12}, x_{t-13}, x_{t-14}, x_{t-15}$
U3	$x_{t-1}, x_{t-2}, x_{t-3}, x_{t-4}, x_{t-5}, x_{t-6}, x_{t-7}, x_{t-8}, x_{t-9}, x_{t-10}, x_{t-11}, x_{t-12}, x_{t-15}$
U4	$x_{t-1}, x_{t-2}, x_{t-3}, x_{t-4}, x_{t-5}, x_{t-6}, x_{t-7}, x_{t-8}, x_{t-9}, x_{t-10}, x_{t-11}, x_{t-12}, x_{t-13}, x_{t-15}$
U5	$x_{t-1}, x_{t-2}, x_{t-3}, x_{t-4}, x_{t-5}, x_{t-6}, x_{t-7}, x_{t-8}, x_{t-9}, x_{t-10}, x_{t-11}, x_{t-12}, x_{t-13}, x_{t-15}$
U6	$x_{t-1}, x_{t-4}, x_{t-5}, x_{t-6}, x_{t-7}, x_{t-8}, x_{t-9}, x_{t-10}, x_{t-11}, x_{t-13}, x_{t-14}, x_{t-15}$
U7	$x_{t-1}, x_{t-2}, x_{t-3}, x_{t-4}, x_{t-6}, x_{t-7}, x_{t-8}, x_{t-9}, x_{t-10}, x_{t-11}, x_{t-13}, x_{t-14}$
U8	$x_{t-1}, x_{t-2}, x_{t-3}, x_{t-4}, x_{t-6}, x_{t-7}, x_{t-8}, x_{t-9}, x_{t-10}, x_{t-11}, x_{t-12}, x_{t-13}, x_{t-14}, x_{t-15}$
U9	$x_{t-1}, x_{t-2}, x_{t-3}, x_{t-4}, x_{t-5}, x_{t-6}, x_{t-7}, x_{t-8}, x_{t-9}, x_{t-10}, x_{t-12}, x_{t-13}, x_{t-15}$
U10	$x_{t-1}, x_{t-2}, x_{t-3}, x_{t-4}, x_{t-5}, x_{t-6}, x_{t-7}, x_{t-8}, x_{t-9}, x_{t-10}, x_{t-11}, x_{t-14}$
U11	$x_{t-1}, x_{t-2}, x_{t-3}, x_{t-4}, x_{t-6}, x_{t-7}, x_{t-8}, x_{t-9}, x_{t-10}, x_{t-12}, x_{t-14}, x_{t-15}$
I1	$x_{t-1}, x_{t-2}, x_{t-5}, x_{t-7}, x_{t-8}$
I2	$x_{t-1}, x_{t-2}, x_{t-3}, x_{t-4}, x_{t-5}, x_{t-6}, x_{t-7}, x_{t-8}, x_{t-9}, x_{t-10}, x_{t-12}, x_{t-13}$
I3	$x_{t-1}, x_{t-2}, x_{t-3}, x_{t-4}, x_{t-5}, x_{t-6}, x_{t-11}, x_{t-12}$
I4	$x_{t-1}, x_{t-2}, x_{t-3}, x_{t-4}, x_{t-5}, x_{t-6}, x_{t-7}, x_{t-8}, x_{t-12}, x_{t-13}, x_{t-15}$
I5	$x_{t-1}, x_{t-2}, x_{t-3}, x_{t-4}, x_{t-5}, x_{t-6}, x_{t-7}, x_{t-10}, x_{t-11}, x_{t-12}$
I6	$x_{t-1}, x_{t-2}, x_{t-3}, x_{t-4}, x_{t-5}, x_{t-6}, x_{t-8}, x_{t-10}, x_{t-11}$
I7	$x_{t-1}, x_{t-2}, x_{t-3}, x_{t-4}, x_{t-5}, x_{t-6}, x_{t-7}$
I8	$x_{t-1}, x_{t-2}, x_{t-3}, x_{t-4}, x_{t-5}, x_{t-6}, x_{t-7}, x_{t-8}, x_{t-9}, x_{t-12}$
I9	$x_{t-1}, x_{t-2}, x_{t-3}, x_{t-4}, x_{t-5}, x_{t-6}$
Original	$x_{t-1}, x_{t-2}, x_{t-3}, x_{t-4}, x_{t-6}, x_{t-8}, x_{t-10}, x_{t-15}$

mapping relationships, resulting in large prediction error. If there are too many hidden units, overfitting may occur. Since the nonlinear function in this paper is not complicated, the number of hidden units will be set to 96×3 . In addition, the maximum number of iterations of LSTM is within [1, 1000] (Peng et al. 2018). In order to improve the prediction accuracy and the calculation speed of the model, we choose to carry out 250 iterations. Ren et al. (2021) suggest that the range of initial learning rate is [0.0001, 0.1]. The learning rate controls to update the weight according to the estimated gradient at the end of each batch. Taking into account the influence of the learning rate on the speed and effect of the model learning problem, 0.005 is selected as the initial learning rate. The parameter settings of LSTM are listed in Table 3.

Comparison experiments

For representing the accuracy and stability of EEMD-VMD-LSTM prediction model more intuitively, five groups of comparative experiments are set up in this paper. There are 13 models participating in the comparison, including ARIMA, BP, EMD-BP, EEMD-BP,

Table 3 Parameters of the LSTM network

Parameter	Value
Iterations	250
Initial learning rate	0.005
Number of hidden units	96×3

EEMD-VMD-BP, ELM, EMD-ELM, EEMD-ELM, EEMD-VMD-ELM, LSTM, EMD-LSTM, EEMD-LSTM, and EEMD-VMD-LSTM. Figure 8 shows the ultimate prediction results of the proposed model from January 2, 2021 to June 30, 2021. Table 4 summarizes the error results of each model.

Experiment I: Comparing between single models

In experiment I, the performance of single prediction models without decomposition algorithm is compared. The comparison models include BP, ELM, and LSTM.

Among them, LSTM model has the best forecasting effect. In terms of the three error indicators, LSTM is slightly better than ELM and significantly better than BP. It suggests that in experiment I, comparing with ELM and BP, LSTM can better capture the characteristics of high

complexity and irregularity of the daily carbon emission data, which is more suitable as a prediction model in the short-term carbon emission field.

Experiment II: Comparing between single models and models combined primary decomposition

After combining with EMD and EEMD decomposition algorithms, the prediction precision of the model has been significantly advanced. Among the six hybrid models, even the EMD-BP model with the worst accuracy ($R^2=0.8419$, $MAPE=0.0303$, $RMSE=1154.7480$) is much better than the three single prediction models. It suggests that data preprocessing is crucial and effective to decrease the difficulty of short-term carbon emission prediction. When LSTM is combined with EMD decomposition algorithm, R^2 is advanced by 39.357%, MAPE is decreased by 50.693%,

Fig. 8 The final prediction results of the proposed model in China (EEMD-VMD-LSTM)

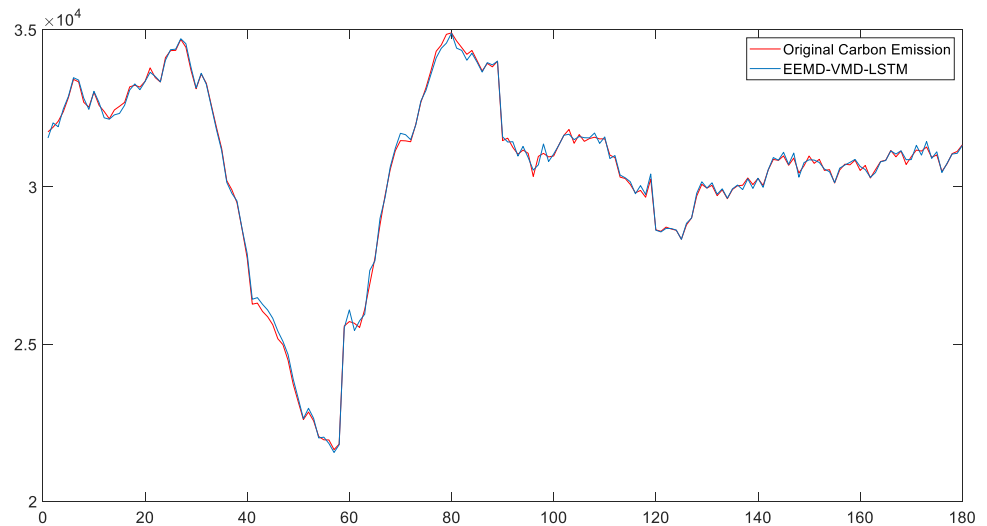


Table 4 The error results of model prediction in China

	Decomposition	Prediction	R^2	MAPE	RMSE
Proposed model	EEMD-VMD	LSTM	0.9983	0.0031	118.1610
Models for comparison		ELM	0.9708	0.0121	495.8792
		BP	0.9618	0.0145	567.4731
		LSTM	0.9862	0.0088	341.0111
	EEMD	ELM	0.9579	0.0165	595.5027
		BP	0.9249	0.0242	795.7480
		LSTM	0.9323	0.0178	755.6394
	EMD	ELM	0.9086	0.0203	877.7854
		BP	0.8419	0.0303	1154.7480
		LSTM	0.6690	0.0361	1670.8170
	Original data	ELM	0.6674	0.0448	1674.8100
		BP	0.6215	0.0523	1786.5330
		ARIMA	0.9240	0.0218	865.2170

and RMSE is reduced by 54.774%, compared to the LSTM model. Similarly, for EEMD-LSTM model, R^2 , MAPE, and RMSE are improved by 47.414%, 75.623%, and 79.590% respectively. It illustrates that different decomposition means having significantly different effects on the improvement of prediction accuracy. In contrast with EMD, EEMD is more suitable for processing the short-term carbon emission data, which is the same as described by the conclusion of Sun and Ren (2021).

Experiment III: Comparing between models combined secondary decomposition and models combined primary decomposition

In this part, there are nine comparison models and the error results of each model are depicted in Fig. 9. It can be seen from the image that EEMD-VMD-LSTM has the highest accuracy, which suggests EEMD-VMD is superior to EEMD and EMD in improving the model prediction ability. The same conclusions are gotten in the experiments of BP and ELM. It seems to show that the secondary decomposition method can better enhance the prediction accuracy than the primary decomposition algorithm. For instance, R^2 , MAPE, and RMSE of EEMD-VMD-LSTM model are respectively improved by 49.223%, 91.413%, and 92.928%

compared with the LSTM model. The gap between EEMD-VMD-LSTM model and LSTM model is obviously larger than that between EEMD-LSTM, EMD-LSTM, and LSTM models. Experiment III indicates that using the secondary decomposition to preprocess the short-term carbon emission data can get a more satisfactory prediction result.

Experiment IV: Comparing between models combined secondary decomposition

In this section, the comparative models include EEMD-VMD-LSTM, EEMD-VMD-ELM, and EEMD-VMD-BP. All the three models add the step of further decomposing IMF1. As for EEMD-VMD-LSTM model ($R^2 = 0.9983$; MAPE = 0.0031; RMSE = 118.1610), compared with EEMD-VMD-ELM and EEMD-VMD-BP, R^2 is advanced by 2.833% and 3.795%, MAPE is decreased by 74.380% and 78.621%, and RMSE is reduced by 76.171% and 79.178%, respectively. Based on R^2 , MAPE, and RMSE, the proposed model has higher prediction accuracy than the other two models.

On the other hand, to further explore the performance of these three models, we repeat the experiment ten times for comparison. As visualized in Fig. 10, among them, EEMD-VMD-LSTM model has the highest stability, and

Fig. 9 Error results of different models

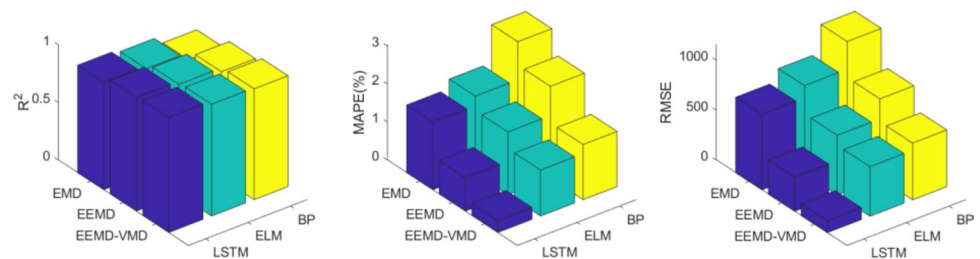
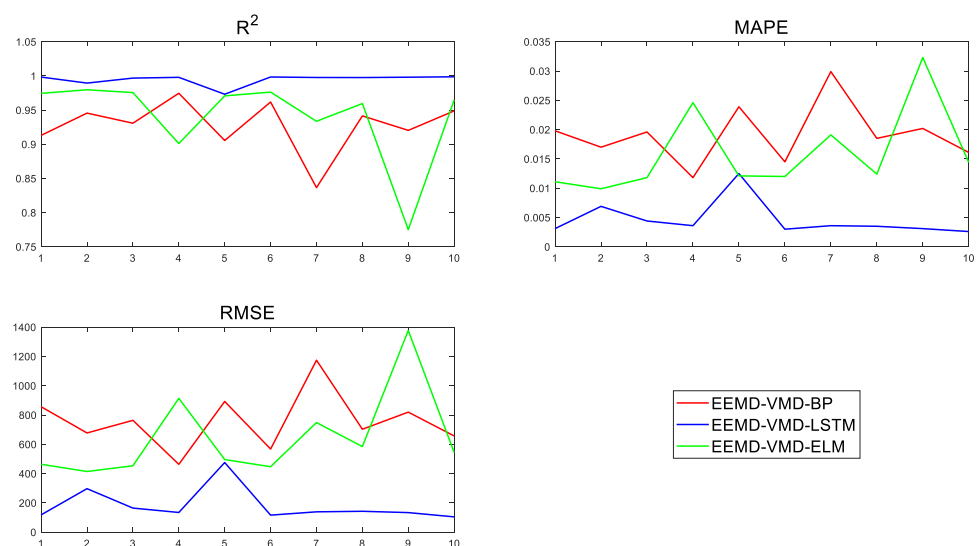


Fig. 10 The performance of models combined secondary algorithm



there is little difference between the results of its ten runs. Meanwhile, the EEMD-VMD-ELM model has the largest volatility, R^2 fluctuates within the range of [0.7751, 0.9797]. In addition, the stability of EEMD-VMD-BP model is between the above two models. This exploration verifies the superiority of EEMD-VMD-LSTM model in terms of robustness.

Experiment V: Comparing between the proposed model and ARIMA model

ARIMA model is a typical autoregressive model. However, it fails to meet desired results in the field of short-term carbon emission prediction ($R^2=0.9240$; MAPE=0.0218; RMSE=865.2170). On the contrary, the EEMD-VMD-LSTM model shows good prediction performance. Compared with the ARIMA model, R^2 , MAPE and RMSE of EEMD-VMD-LSTM model are respectively improved by 8.041%, 85.780%, and 86.343%. It is clear that the proposed model has higher prediction accuracy by comparing with ARIMA model. The comparative analysis further proves that the proposed model is suitable for short-term carbon emission prediction.

Supplementary case

To verify the stability, generalization ability, and applicability of the proposed model in the field of short-term carbon emission, the US daily carbon emissions dataset from January 1, 2019 to June 30, 2021 is added to test the model. Similarly, the sample data from January 1, 2019 to January 1, 2021 are selected as the training set, and the sample data from January 2, 2019 to June 30, 2021 are selected as the testing set. The final prediction results of the proposed model are shown in Fig. 11. The prediction error results of each model are shown in

Table 5. Through the analysis of prediction results and error results, the same conclusions as previous experiments can be obtained.

Results and discussion

Through experiment I, it can be found that the prediction accuracy of LSTM is higher than that of BP and ELM. The same conclusion is obtained in the US sample set. This empirical result is consistent with Ma et al. (2015). The reason may be that LSTM has strong learning ability and time series data analysis ability. Compared with BP and ELM, LSTM can better capture valuable information in daily carbon emission historical data.

The prediction results also show that the instability and volatility of daily carbon emission series will lead to low prediction accuracy. Even the LSTM model with good prediction performance cannot get satisfactory prediction results. However, the method of “decomposition before prediction” can solve the problem well. Before the prediction process, the decomposition algorithm is used to decompose the original data into several low-frequency and more regular subsequences, which will reduce the prediction difficulty. It is also proved by Sun and Zhang (2018). Furthermore, as mentioned by Yin et al. (2017), the IMF1 obtained after decomposition contains the most random components and is still difficult to predict. The prediction accuracy can be further improved by re-decomposing IMF1, which is the same as the conclusion of our empirical analysis. In addition, by combining the secondary decomposition algorithm and deep learning neural network, the prediction accuracy is further improved. Compared with the experimental results of Sun and Ren (2021), the R^2 and MAPE of the proposed model are improved by 5.006% and 66.667%, respectively.

Fig. 11 The final prediction results of the proposed model in the USA

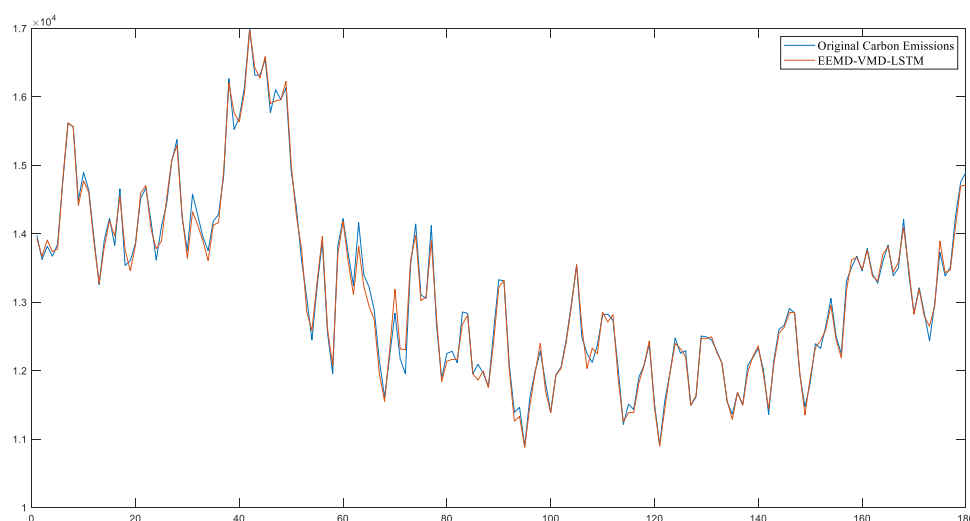


Table 5 The error results of model prediction in the USA

	Decomposition	Prediction	R^2	MAPE	RMSE
Proposed model	EEMD-VMD	LSTM	0.9926	0.0066	110.8699
Models for comparison		ELM	0.9357	0.0187	327.4049
		BP	0.9567	0.0166	268.8085
		LSTM	0.9653	0.0138	240.5428
	EEMD	ELM	0.9103	0.0229	386.8637
		BP	0.9251	0.0227	353.3224
		LSTM	0.9326	0.0198	335.3777
	EMD	ELM	0.8808	0.0272	445.8296
		BP	0.9058	0.0235	396.4291
		LSTM	0.7559	0.0367	638.1003
	Original data	ELM	0.6782	0.0422	732.5987
		BP	0.7206	0.0421	682.6865
		ARIMA	0.8180	0.0379	658.2190

According to the empirical results, it can be found that (1) No matter which data set, the prediction accuracy of EEMD-VMD-LSTM model proposed in this paper is the highest. At the same time, the prediction results of ARIMA model in the two datasets show great differences. The superiority and excellent generalization ability of the proposed model have been further verified in the field of short-term carbon emission by comparing with the typical autoregressive model. (2) After combining with the decomposition algorithm, the prediction accuracy of the model has been significantly improved, which proves that the method of “decomposition before prediction” is feasible and effective. (3) EEMD-VMD improves the prediction accuracy the most. The empirical results also show that in the field of short-term carbon emission, the improvement of prediction accuracy by the three decomposition algorithms presents a fixed order: EEMD-VMD > EEMD > EMD. It enriches the research of decomposition algorithms in the field of carbon emission. (4) In China dataset, ELM predicts better than BP, but the opposite result is obtained in the US dataset. In terms of R^2 , MAPE, and RMSE, LSTM is the best of the single models. In the field of short-term carbon emission, LSTM has better prediction performance and generalization ability compared with BP and ELM.

Moreover, it can be found that China’s daily carbon emissions show obvious seasonal variation characteristics. The lowest value of carbon emissions occurs in spring, while the highest value appears in winter. It may be that anthropogenic carbon emissions and vegetation dynamics are the main reasons for the seasonal change of carbon emissions. In spring, vegetation dynamics dominate the variation of carbon emissions, and the strong carbon sink effect of vegetation significantly reduces carbon emissions. In winter, carbon emissions rise significantly due to heating demand. And China’s daily carbon emissions have increased compared with the previous 2 years, which is still worthy of

our vigilance. Based on the above findings, some policy recommendations have been put forward, which may help decision-makers. In China, coal-based heating in winter will lead to a large amount of carbon dioxide emissions. It seems sensible to promote energy-saving technology development and energy transition in the short term. The government should further increase the research investment in new clean energy and reduce the cost of low-carbon research industry through tax relief. In addition, increasing vegetation planting and urban green coverage is also a good way to achieve low carbon emission reduction.

The proposed model provides a new way to improve the prediction accuracy of daily carbon emissions. EEMD-VMD-LSTM model can not only provide decision-makers with timely dynamic change information of carbon emissions to make response plans rapidly, but also provide direction and reference for carbon emission reduction, which has strong practical value. The generalization ability and robustness of the model can be demonstrated by empirical evidence in China and the USA. The prediction model can be applied to other regions, industries, and departments with high carbon emissions to help decision-makers make reasonable plans to control carbon emissions.

Conclusion

In this paper, a novel model combined with LSTM and secondary decomposition, EEMD-VMD-LSTM, is proposed to predict carbon emissions, which is simultaneously the first application of secondary decomposition in short-term carbon emission forecasting. On the one hand, it supplies a new method for the prediction of short-term carbon emissions. On the other hand, it can provide a valuable reference for the research on carbon emission prediction in the future. Through comparison with the prediction results of the other

12 models, the excellent precision and stability of the proposed model are demonstrated. On the foundation of the empirical analysis results, the conclusions are summarized below:

- (1) The combination models possess distinguished performance compared with the single models. It is proved that data preprocessing is essential, which can greatly reduce the difficulty of prediction. It provides a novel thought for carbon emission research.
- (2) The secondary decomposition is also applicable to the field of carbon emission. Compared with primary decomposition, secondary decomposition has better decomposition ability and prediction ability. It broadens the application scope of secondary decomposition, makes up the gap of secondary decomposition in the carbon emission field, and enriches the study of secondary decomposition.
- (3) LSTM in the field of deep learning shows strong predictive ability. Whether combined with decomposition algorithm or not, LSTM achieves a better prediction effect than ELM and BP, in terms of three error test criteria. It explains that LSTM not only has strong generalization ability but also possesses strong adaptability to different decomposition algorithms. By comparing with ELM and BP, the excellent performance of LSTM in short-term carbon emission forecasting is verified.
- (4) The proposed EEMD-VMD-LSTM model in this paper is suitable for the forecasting of short-term carbon emissions. The combination model has a high reference value, which can benefit the carbon emission reduction process through realizing accurate forecast and supplying short-term carbon emission dynamics. Planners responsible for carbon emission reduction can formulate response measures according to the changing trend of short-term carbon emission more flexibly. Moreover, time series prediction is independent of external variables, which emphasizes the importance of time factors in forecasting and temporarily ignores the influence of external factors. The time span of the training sample data selected in this paper includes the time of COVID-19 outbreak, and the proposed model accurately predicts the future trend of carbon emission. This paper further verifies the advantages of time series prediction by predicting the future trend of carbon emission.

Although this study has achieved some innovations and improvements in the field of carbon emission prediction, there are still some limitations. For instance, the proposed model shows good prediction performance in the field of short-term carbon emissions, but we do not know whether it is suitable for other fields. In this paper, the parameters of LSTM are not optimized. The combination of LSTM and

optimization algorithm is worth exploring by us and more scholars.

Acknowledgements Thanks are due to Kong Feng for valuable discussion.

Author contribution Kong Feng: conceptualization, feasibility analysis. Song Jianbo: methodology, data collection, debugging, original draft writing. Yang Zhongzhi: software, verification.

Data availability The datasets used during and/or analyzed during the current study are available in the CEADs repository, <https://www.ceads.net/>.

Declarations

Ethics approval and consent to participate Not applicable.

Consent for publication Not applicable.

Competing interests The authors declare no competing interests.

References

- Aslam B, Hu J, Ali S, AlGarni TS, Abdullah MA (2021) Malaysia's economic growth, consumption of oil, industry and CO₂ emissions: evidence from the ARDL model. *Int J Environ Sci Technol*. <https://doi.org/10.1007/s13762-021-03279-1>
- Boamah KB, Du J, Adu D, Mensah CN, Dauda L, Khan MAS (2021) Predicting the carbon dioxide emission of China using a novel augmented hypo-variance brain storm optimisation and the impulse response function. *Environ Technol* 42:4342–4354. <https://doi.org/10.1080/09593330.2020.1758217>
- Chang T, Jo S-H, Lu W (2011) Short-term memory to long-term memory transition in a nanoscale memristor. *ACS Nano* 5:7669–7676. <https://doi.org/10.1021/nn202983n>
- Chang Z, Zhang Y, Chen W (2019) Electricity price prediction based on hybrid model of adam optimized LSTM neural network and wavelet transform. *Energy* 187. <https://doi.org/10.1016/j.energy.2019.07.134>
- Ding Y, Zhu Y, Feng J, Zhang P, Cheng Z (2020) Interpretable spatio-temporal attention LSTM model for flood forecasting. *Neurocomputing* 403:348–359. <https://doi.org/10.1016/j.neucom.2020.04.110>
- Dragomiretskiy K, Zosso D (2014) Variational mode decomposition. *IEEE Trans Signal Process* 62:531–544. <https://doi.org/10.1109/TSP.2013.2288675>
- E J, Ye J, He L, Jin H (2021) A denoising carbon price forecasting method based on the integration of kernel independent component analysis and least squares support vector regression. *Neurocomputing* 434:67–79. <https://doi.org/10.1016/j.neucom.2020.12.086>
- Fischer T, Krauss C (2018) Deep learning with long short-term memory networks for financial market predictions. *Eur J Oper Res* 270:654–669. <https://doi.org/10.1016/j.ejor.2017.11.054>
- Gao P, Yue S, Chen H (2021) Carbon emission efficiency of China's industry sectors: from the perspective of embodied carbon emissions. *J Clean Prod* 283. <https://doi.org/10.1016/j.jclepro.2020.124655>
- Gu B, Zhang T, Meng H, Zhang J (2021) Short-term forecasting and uncertainty analysis of wind power based on long short-term

- memory, cloud model and non-parametric kernel density estimation. *Renewable Energy* 164:687–708. <https://doi.org/10.1016/j.renene.2020.09.087>
- Hashim H, Ramlan MR, Shiun LJ, Siong HC, Kamyab H, Majid MZA, Lee CT (2015) An integrated carbon accounting and mitigation framework for greening the industry. *Energy Procedia* 75:2993–2998. <https://doi.org/10.1016/j.egypro.2015.07.609>
- Heydari A, Garcia DA, Keynia F, Bisegna F, Santoli LD (2019) Renewable energies generation and carbon dioxide emission forecasting in microgrids and national grids using GRNN-GWO methodology. *Energy Procedia* 159:154–159. <https://doi.org/10.1016/j.egypro.2018.12.044>
- Hochreiter S, Schmidhuber J (1997) Long short-term memory. *Neural Comput* 9:1735–1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
- Hosseini SM, Saifoddin A, Shirmohammadi R, Aslani A (2019) Forecasting of CO₂ emissions in Iran based on time series and regression analysis. *Energy Rep* 5:619–631. <https://doi.org/10.1016/j.egy.2019.05.004>
- Huang NE, Wu ZH (2008) A review on Hilbert-Huang transform: method and its applications to geophysical studies. *Rev Geophys* 46. <https://doi.org/10.1029/2007RG000228>
- Huang Y, Shen L, Liu H (2019) Grey relational analysis, principal component analysis and forecasting of carbon emissions based on long short-term memory in China. *J Clean Prod* 209:415–423. <https://doi.org/10.1016/j.jclepro.2018.10.128>
- Li H, Jin F, Sun S, Li Y (2021a) A new secondary decomposition ensemble learning approach for carbon price forecasting. *Knowl-Based Syst* 214. <https://doi.org/10.1016/j.knosys.2020.106686>
- Li Y, Dong H, Lu S (2021b) Research on application of a hybrid heuristic algorithm in transportation carbon emission. *Environ Sci Pollut Res Int* 28:48610–48627. <https://doi.org/10.1007/s11356-021-14079-y>
- Liu H, Zhang X (2021) AQI time series prediction based on a hybrid data decomposition and echo state networks. *Environ Sci Pollut Res Int* 28:51160–51182. <https://doi.org/10.1007/s11356-021-14186-w>
- Liu Z, Sun W, Zeng J (2013) A new short-term load forecasting method of power system based on EEMD and SS-PSO. *Neural Comput Appl* 24:973–983. <https://doi.org/10.1007/s00521-012-1323-5>
- Luo X, Li D, Yang Y, Zhang S (2019) Spatiotemporal traffic flow prediction with KNN and LSTM. *J Adv Transp* 2019:1–10. <https://doi.org/10.1155/2019/4145353>
- Ma X, Tao Z, Wang Y, Yu H, Wang Y (2015) Long short-term memory neural network for traffic speed prediction using remote microwave sensor data. *Transp Res Part c: Emerg Technol* 54:187–197. <https://doi.org/10.1016/j.trc.2015.03.014>
- Malik A, Hussain E, Baig S, Khokhar MF (2020) Forecasting CO₂ emissions from energy consumption in Pakistan under different scenarios: the China-Pakistan economic corridor. *Greenh Gases: Sci Technol* 10:380–389. <https://doi.org/10.1002/ghg.1968>
- Memarzadeh G, Keynia F (2020) A new short-term wind speed forecasting method based on fine-tuned LSTM neural network and optimal input sets. *Energy Convers Manage* 213. <https://doi.org/10.1016/j.enconman.2020.112824>
- Mi X-w, Liu H, Li Y-f (2017) Wind speed forecasting method using wavelet, extreme learning machine and outlier correction algorithm. *Energy Convers Manage* 151:709–722. <https://doi.org/10.1016/j.enconman.2017.09.034>
- Park S-Y, Sur C, Lee J-H, Kim J-S (2020) Ecological drought monitoring through fish habitat-based flow assessment in the Gam river basin of Korea. *Ecol Ind* 109. <https://doi.org/10.1016/j.ecolind.2019.105830>
- Peng L, Liu S, Liu R, Wang L (2018) Effective long short-term memory with differential evolution algorithm for electricity price prediction. *Energy* 162:1301–1314. <https://doi.org/10.1016/j.energy.2018.05.052>
- Ren X, Liu S, Yu X, Dong X (2021) A method for state-of-charge estimation of lithium-ion batteries based on PSO-LSTM. *Energy* 234. <https://doi.org/10.1016/j.energy.2021.121236>
- Şentürk AŞ, Zehra K (2021) Yapay Sinir Ağları İle Göğüs Kanseri Tahmini. *El-Cezeri Fen ve Mühendislik Derg* 3. <https://dergipark.org.tr/tr/pub/ecjse/264199>
- Sun W, Huang C (2020) A novel carbon price prediction model combines the secondary decomposition algorithm and the long short-term memory network. *Energy* 207. <https://doi.org/10.1016/j.energy.2020.118294>
- Sun W, Ren C (2021) Short-term prediction of carbon emissions based on the EEMD-PSOBP model. *Environ Sci Pollut Res Int* 28:56580–56594. <https://doi.org/10.1007/s11356-021-14591-1>
- Sun W, Zhang C (2018) Analysis and forecasting of the carbon price using multi-resolution singular value decomposition and extreme learning machine optimized by adaptive whale optimization algorithm. *Appl Energy* 231:1354–1371. <https://doi.org/10.1016/j.apenergy.2018.09.118>
- Wang L, Xue X, Zhao Z, Wang Y, Zeng Z (2020) Finding the decarbonization potentials in the transport sector: application of scenario analysis with a hybrid prediction model. *Environ Sci Pollut Res Int* 27:21762–21776. <https://doi.org/10.1007/s11356-020-08627-1>
- Wang W, Wang J (2021) Determinants investigation and peak prediction of CO₂ emissions in China's transport sector utilizing bio-inspired extreme learning machine. *Environ Sci Pollut Res Int* 28:55535–55553. <https://doi.org/10.1007/s11356-021-14852-z>
- Wen L, Cao Y (2020) Influencing factors analysis and forecasting of residential energy-related CO₂ emissions utilizing optimized support vector machine. *J Clean Prod* 250. <https://doi.org/10.1016/j.jclepro.2019.119492>
- Wu Z, Huang NE (2009) Ensemble empirical mode decomposition: a noise-assisted data analysis method. *Adv Adapt Data Anal* 01:1–41. <https://doi.org/10.1142/S1793536909000047>
- Yang H, O'Connell JF (2020) Short-term carbon emissions forecast for aviation industry in Shanghai. *J Clean Prod* 275. <https://doi.org/10.1016/j.jclepro.2020.122734>
- Yin H, Dong Z, Chen Y, Ge J, Lai LL, Vaccaro A, Meng A (2017) An effective secondary decomposition approach for wind power forecasting using extreme learning machine trained by crisscross optimization. *Energy Convers Manage* 150:108–121. <https://doi.org/10.1016/j.enconman.2017.08.014>
- Zhao B, Yang W (2020) Short-run forecast and reduction mechanism of CO₂ emissions: a Chinese province-level study. *Environ Sci Pollut Res Int*. <https://doi.org/10.1007/s11356-020-09936-1>
- Zhao Y, Zhang B, Han L (2020) Laser self-mixing interference displacement measurement based on VMD and phase unwrapping. *Opt Commun* 456. <https://doi.org/10.1016/j.optcom.2019.124588>

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